Gradient flows for the stochastic Amari neural field model

Jonas M. Tölle (Augsburg University)

joint work with

Christian Kuehn (TU Munich)

https://arxiv.org/abs/1807.02575

9th International Conference on Stochastic Analysis and its Applications

Bielefeld University

September 3, 2018

Contents

Section 1

The stochastic Amari neural field model

Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy\right] dt + \varepsilon B\, dW_t(x)$$

Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy\right] dt + \varepsilon B\, dW_t(x)$$

- $\mathcal{B} \subset \mathbb{R}^d$ bounded closed domain,

Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy\right] dt + \varepsilon B\, dW_t(x)$$

- $\mathcal{B} \subset \mathbb{R}^d$ bounded closed domain,
- $U : \mathcal{B} \times [0,T] \times \Omega \to \mathbb{R}$, "voltage"

Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy \right] dt + \varepsilon B\, dW_t(x)$$

- $\mathcal{B} \subset \mathbb{R}^d$ bounded closed domain,
- $U : \mathcal{B} \times [0,T] \times \Omega \to \mathbb{R}$, “voltage”
- $\alpha > 0$, decay parameter, $0 < \varepsilon \ll 1$, noise intensity parameter,

Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy\right] dt + \varepsilon B\, dW_t(x)$$

- $\mathcal{B} \subset \mathbb{R}^d$ bounded closed domain,
- $U : \mathcal{B} \times [0, T] \times \Omega \to \mathbb{R}$, "voltage"
- $\alpha > 0$, decay parameter, $0 < \varepsilon \ll 1$, noise intensity parameter,

Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy \right] dt + \varepsilon B\, dW_t(x)$$

- $\mathcal{B} \subset \mathbb{R}^d$ bounded closed domain,
- $U : \mathcal{B} \times [0,T] \times \Omega \to \mathbb{R}$, "voltage"
- $\alpha > 0$, decay parameter, $0 < \varepsilon \ll 1$, noise intensity parameter,
- $w : \mathcal{B} \times \mathcal{B} \to \mathbb{R}$ kernel, modeling neural connectivity,

Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy\right] dt + \varepsilon B\, dW_t(x)$$

- $\mathcal{B} \subset \mathbb{R}^d$ bounded closed domain,
- $U : \mathcal{B} \times [0,T] \times \Omega \to \mathbb{R}$, "voltage"
- $\alpha > 0$, decay parameter, $0 < \varepsilon \ll 1$, noise intensity parameter,
- $w : \mathcal{B} \times \mathcal{B} \to \mathbb{R}$ kernel, modeling neural connectivity,
- $f : \mathbb{R} \to (0, +\infty)$ gain function, modeling neural input,

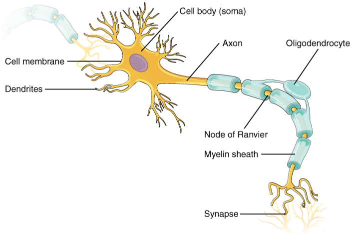
Stochastic PDE for mean-field cortex activity

Amari-type neural field model [Amari, *Biological Cybernetics* (1977)].

$$dU_t(x) = \left[-\alpha U_t(x) + \int_{\mathcal{B}} w(x,y) f(U_t(y))\, dy \right] dt + \varepsilon B\, dW_t(x)$$

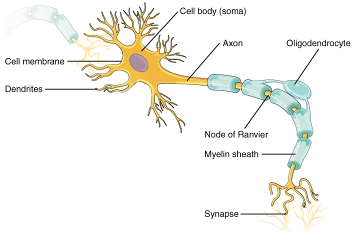
- $\mathcal{B} \subset \mathbb{R}^d$ bounded closed domain,
- $U : \mathcal{B} \times [0,T] \times \Omega \to \mathbb{R}$, "voltage"
- $\alpha > 0$, decay parameter, $0 < \varepsilon \ll 1$, noise intensity parameter,
- $w : \mathcal{B} \times \mathcal{B} \to \mathbb{R}$ kernel, modeling neural connectivity,
- $f : \mathbb{R} \to (0, +\infty)$ gain function, modeling neural input,
- $\{W_t\}_{t \geqslant 0}$ cylindrical Wiener process with values in $H := L^2(\mathcal{B})$, modeled on $(\Omega, \mathcal{F}, \mathbb{P})$; additive noise coefficient $B \in L(H)$.

Neuron

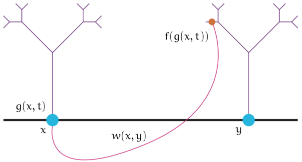


[OpenStax, Anatomy & Physiology (2018)]

Neuron / neural fields



[OpenStax, Anatomy & Physiology (2018)]



$g \sim$ voltage, $f \sim$ gain, $w \sim$ connectivity

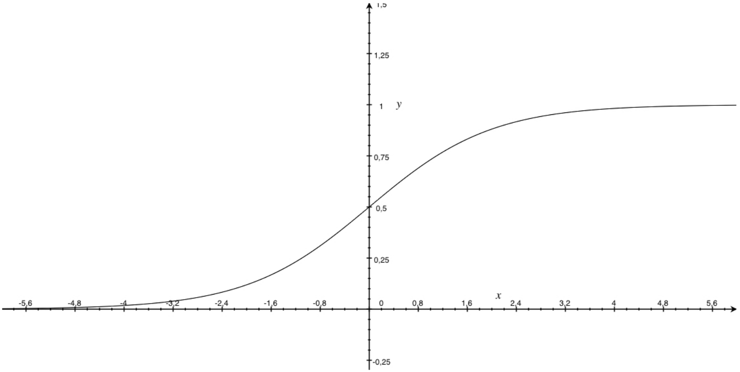
[Coombes, beim Graben, Potthast (2014)

In: Neural Fields. Springer]

Typical gain functions f

Assume that $F : H \to H$, $F(v)(x) := f(v(x))$, $v \in H = L^2(\mathcal{B})$ is a Nemytskii operator.

$$f(s) = (1 + e^{-s})^{-1} \ (Sigmoid)$$

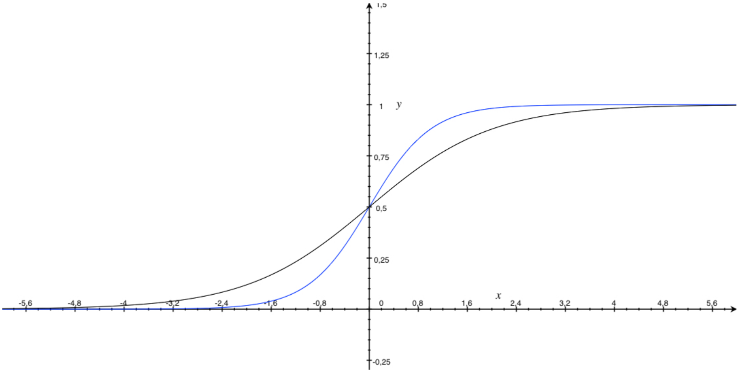


Typical gain functions f

Assume that $F : H \to H$, $F(v)(x) := f(v(x))$, $v \in H = L^2(\mathcal{B})$ is a Nemytskii operator.

$f(s) = (1 + e^{-s})^{-1}$ (*Sigmoid*)

$f(s) = \frac{1}{2}(\tanh(s) + 1)$



The kernel

Let $w : \mathcal{B} \times \mathcal{B} \to \mathbb{R}$ be measurable such that:

Assumption 1

1. $w(x, y) = w(y, x)$ for a.e. $x, y \in \mathcal{B}$,
2. $w \in L^2(\mathcal{B} \times \mathcal{B}) \cap C(\mathcal{B} \times \mathcal{B})$,
3. w satisfies

$$\sum_{i,j=1}^{n} c_i c_j w(x_i, x_j) \geqslant 0$$

for every $n \in \mathbb{N}$, for every $\{x_1, \ldots, x_n\} \subset \mathcal{B}$, and for every $\{c_1, \ldots, c_n\} \subset \mathbb{R}$.

$w(x, y) = J(x - y)$

Then Assumption 1 implies that the linear operator $K \in L(H)$ defined by

$$Kg(x) := \int_{\mathcal{B}} w(x, y) g(y)\, dy, \quad g \in H,$$

is a nonnegative definite, self-adjoint Hilbert-Schmidt operator and, moreover, even of trace-class ($\longrightarrow$ Mercer's therorem), as w is a so-called *Mercer kernel* on a compact subset of $\mathbb{R}^d$.

Assumption 2

Let $J \in C(\mathbb{R}^d)$ such that $w(x, y) = J(x - y)$ for $x, y \in \mathcal{B}$.

Now, w satisfies Assumption 1 (3) e.g. if J is of the form

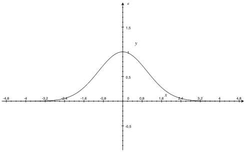
$$J(x) = \int_{\mathbb{R}^d} \cos(\langle y, x \rangle)\, \sigma(dy), \quad x \in \mathbb{R}^d,$$

for some **symmetric** probability measure σ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \longrightarrow$ Bochner's theorem.

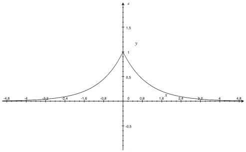
Examples $d = 1$ (characteristic functions)

In this case, J is a real-valued *characteristic function* of a symmetric probability distribution.

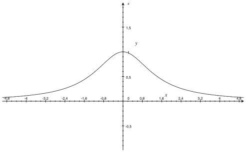
Gaussian



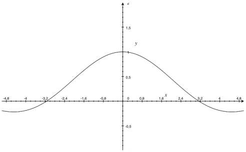
Cauchy



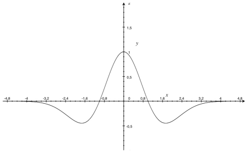
Laplace



Uniform (on $[-1, 1]$)



Maxwell–Boltzmann (*"Mexican hat" wavelet*)



The stochastic PDE

Let $\{W_t\}_{t \geqslant 0}$ be a cylindrical Wiener process with values in $H = L^2(\mathcal{B})$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0}, \mathbb{P})$, $B \in L(H)$.

$$dU_t = [-\alpha U_t + KF(U_t)]\, dt + \varepsilon B\, dW_t, \quad U_0 = u_0 \in H, \; t \in [0, T]. \qquad (*)$$

Fact ([Da Prato, Zabczyk, *Cambridge Univ. Press* (1992)], see also [Kuehn, Riedler, *J. Math. Neuroscience* (2014)])

Let $B \in L_2(H)$. Then there exists a unique mild solution to $(*)$ with $U \in C([0, T]; H)$, $\mathbb{P}$-a.s. having the form

$$U_t = e^{-\alpha t} u_0 + \int_0^t e^{-\alpha(t-s)} KF(U_s)\, ds + \varepsilon \int_0^t e^{-\alpha(t-s)} B\, dW_s \quad \text{in } H.$$

Section 2

A gradient flow formulation

Gradient flows

In fact even for $\varepsilon = 0$, it was not known, whether one can find a gradient structure to rewrite the PDE as

$$\partial_t u = -\nabla_{\mathcal{X}} \mathcal{F}(u), \quad u(\cdot, t) = u(t) \in \mathcal{X},$$

where $\mathcal{X}$ is a suitable function space — Hilbert, Banach, or metric space (see e.g. [Ambrosio, Gigli, Savaré, *Birkhäuser* (2006)]) — and where

$$\mathcal{F} : \mathcal{X} \to \mathbb{R}$$

is a functional, which has often the natural interpretation of an energy, entropy or some other physical notion.

It was already seen in [Kuehn, Riedler, *J. Math. Neuroscience* (2014)] that the naïve guess for $\mathcal{F}$ in $\mathcal{X} := L^2(\mathcal{B})$ fails to produce the desired formulation.

Change of ambient space

Recall: K is trace-class, nonnegative definite, self-adjoint. Hence:

- The spectrum $\sigma(K)$ is discrete with zero being its only accumulation point.
- There exists an orthonormal basis $\{e_i\}$ of eigenvectors in $L^2(\mathcal{B})$ such that the eigenvalues $\lambda_i \in \sigma(K) \setminus \{0\}$ satisfy w.l.o.g.

$$\lim_{i\to\infty} \lambda_i = 0.$$

We have the orthogonal decomposition

$$H = \operatorname{Ker}(K) \oplus S$$

where $S := \operatorname{Ker}(K)^\perp = \overline{\operatorname{span}\{e_i\}_{i\in\mathbb{N}}}$.

On S becomes a separable Hilbert space (denoted by H_{-1}) with norm

$$\|u\|_{-1} := \|K^{-\frac{1}{2}} u\|_H \quad u \in S,$$

where $K^{-\frac{1}{2}}$ is the *Moore-Penrose pseudo-inverse* of $K^{\frac{1}{2}}$.

Gradients

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be any primitive function of f. Set

$$\Phi(u) := \int_{\mathcal{B}} \varphi(u(x))\, dx \quad u \in H,$$

and let

$$\Psi(v) := \frac{\alpha}{2} \|v\|_{-1}^2, \quad u \in S.$$

Lemma

Φ is well-defined, finite for all $u \in H$ and continuous in H. Furthermore, we have that

$$D\Phi(u)h = (F(u), h)_H, \quad u, h \in H,$$

where $D\Phi(u)h$ denotes the Gâteaux-directional derivative of Φ in u in direction h.

Gradients

Furthermore, set $\Theta(u) := \Psi(u) - \Phi|_S(u)$, $u \in S$.

Lemma

For $u, h \in H_{-1}$, we have that

$$D\Theta(u)h = \alpha(u,h)_{-1} - (KF(u), h)_{-1},$$

where $D\Theta(u)h$ denotes the Gâteaux-directional derivative of Φ in u in direction h.

Compare also with the ideas of [Ren, Röckner, Wang, *J. Differential Equations* (2007)] and [Röckner, Wang, *J. Differential Equations* (2008)] $\longrightarrow$ generalized stochastic porous media equation.

However, in our situation, F is **not** assumed monotone.

Inhibition and excitation

Remark

In the case that K is nonpositive definite, we can redefine H_{-1} by replacing K by $-K$ in the definition. Now, by changing the sign for Θ above, we obtain a gradient by a similar procedure. We can interpret the case of nonnegative definite symmetric kernels as domination by *excitation*, while the case of nonpositive definite kernels corresponds to domination of the *inhibition* effects.

Gradient flow formulation

Let $\{W_t\}_{t\geqslant 0}$ be as above. Let $B \in L(H, H_{-1})$. Consider the gradient flow SPDE

$$dV_t = -D\Theta(V_t)\,dt + \varepsilon B\,dW_t, \quad V_0 = v_0 \in H_{-1}.$$

We shall assume below that $B \in L_2(H, H_{-1})$. Under the assumption that B is diagonalized w.r.t. $\{e_i\}$ with eigenvalues $\{b_i\}$, i.e.

$$Be_i = b_i e_i, \quad i \in \mathbb{N},$$

it holds that

Lemma

$$\{b_i^2 \lambda_i^{-1}\} \in \ell^1 \quad \Longleftrightarrow \quad B \in L_2(H, H_{-1}).$$

One possibility is to set $B := K$, which corresponds to the continuum limit of a neural Langevin equation, see [Bressloff, *J. Phys. A* (2012)].

Gradient flow formulation (invariant subspace)

Theorem (Kuehn, T. (2018+), invariance of the subspace H_{-1})

For $B \in L_2(H.H_{-1})$, we have that there exists a unique mild solution $\{V_t\}_{t\geqslant 0}$ in H_{-1} such that, in particular, for $v_0 \in H_{-1}$, we have that $V \in L^2([0,T];H_{-1})$ $\mathbb{P}$-a.s. and

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|V_t\|_{-1}^2\right] \leqslant C\|u_0\|_{-1}^2 + \varepsilon^2 C,$$

In particular, the mild solution in H_{-1} coincides with the mild solution in H for initial data in H_{-1}.

In our work, we also study strong and weak solutions to the above SPDE in H_{-1} via:

- Transformation of the additive noise,
- Dissipativity of the drift.

Invariant measures

Let $\{W_t\}_{t\geqslant 0}$ be a cylindrical Wiener process with values in H. Consider solutions to the equation

$$dX_t^{z,\varepsilon} = (AX_t^{z,\varepsilon} + D\Phi(X_t^{z,\varepsilon}))\,dt + \varepsilon\,dW_t, \quad X_0^{z,\varepsilon} = z \in H_{-1},$$

with $(Au,\cdot)_{-1} := -\alpha(K^{-1}u,\cdot)_H$ (which is the generator of a C_0-semigroup $\{S_t\}_{t\geqslant 0}$ on H_{-1} which is the restriction of a C_0-semigroup $\{S_t^0\}_{t\geqslant 0}$ on H).

Recall that $D\Phi(u)h = (F(u),h)_H,\ u,h \in H$.

Remark

We may w.l.o.g. assume that $t \mapsto \int_0^t S_{t-s}^0\,dW_s \in C([0,T];H_{-1})$ $\mathbb{P}$-a.s. as the semigroup $\{S_t^0\}_{t\geqslant 0}$ is *analytic*.

Invariant measures (existence)

Define the *transition semigroup*

$$P_t^\varepsilon G(z) := \mathbb{E}\left[G(X_t^{z,\varepsilon})\right] \quad t \geqslant 0,\ z \in H_{-1},$$

where $G : H_{-1} \to \mathbb{R}$ is bounded and measurable.

Theorem (by applying results from [Zabczyk, *SPDEs and Appl. II* (1989), Da Prato, Zabczyk, *Cambridge Univ. Press* (1992)])

Assume that $t \mapsto \int_0^t S_{t-s}^0 \, dW_s \in C([0,T]; H_{-1})$ $\mathbb{P}$*-a.s. Then* $\{P_t^\varepsilon\}_{t\geqslant 0}$ *is strongly Markovian and symmetric with respect to its invariant measure, which exists and takes the following form*

$$\mu_\varepsilon(dz) := \frac{1}{Z_\varepsilon} \exp\left[2\varepsilon^{-2}\Phi(z)\right] \gamma_\varepsilon(dz),$$

where $Z_\varepsilon := \int_{H_{-1}} \exp\left[2\varepsilon^{-2}\Phi(z)\right] \gamma_\varepsilon(dz)$ *and* $\gamma_\varepsilon \sim N(0, \Gamma_\varepsilon)$*, where* $\Gamma_\varepsilon := 2\varepsilon^2\alpha^{-1}K$.

Invariant measures (uniqueness)

Theorem (Compare with [Maslowski, *Stoch. Systems and Optim.* (1989)])

Assume that $t \mapsto \int_0^t S^0_{t-s}\, dW_s \in C([0,T]; H_{-1})$ $\mathbb{P}$*-a.s.. Then* μ_ε *is strong Feller in the restricted sense (asymptotic strong Feller) and thus unique, and the semigroup* $\{P_t^\varepsilon\}_{t \geqslant 0}$ *is ergodic.*

Possible applications:

- Large deviation principle / small noise asymptotics,
- Kramers' law, (see e.g. [Berglund, *Markov Processes Related Fields* (2013)]),
- Kolmogorov operators / Fokker-Planck equations.

Final remarks

Possible extensions:

- Situations where K and B do **not** commute,
- the situation of $\mathcal{B} = \mathbb{R}^d$,
- multiplicative noise,
- *indefinite* kernels K (however, with dominating positive or negative spectrum).

S. Amari.
Dynamics of pattern formation in lateral-inhibition type neural fields.
Biol. Cybernet., 27:77–87, 1977.

N. Berglund.
Kramers' law: validity, derivations and generalisations.
Markov Processes and Related Fields, 19(3):459–490, 2013.

P. C. Bressloff.
Spatiotemporal dynamics of continuum neural fields.
J. Phys. A, 45(033001):1–109, 2012.

O. Faugeras and J. Inglis.
Stochastic neural field equations: a rigorous footing.
J. Math. Biol., 71(2):259–300, 2015.

R. Jordan, D. Kinderlehrer, and F. Otto.
The variational formulation of the Fokker–Planck equation.
SIAM Journal on Mathematical Analysis, 29(1):1–17, 1998.

C. Kuehn and M. G. Riedler.
Large deviations for nonlocal stochastic neural fields.
J. Math. Neurosci., 4(1):1–33, 2014.

C. Kuehn and J. M. T.
A gradient flow formulation for the stochastic Amari neural field model.
Preprint, arXiv:1807.02575, submitted, 1–18, 2014.

B. Maslowski.
Strong Feller property for semilinear stochastic evolution equations and applications.
In *Stochastic Systems and Optimization*, pages 210–224. Springer, Berlin, Heidelberg, 1989.

H. Wilson and J. Cowan.
A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue.
Biol. Cybern., 13(2):55–80, 1973.

J. Zabczyk.
Symmetric solutions of semilinear stochastic equations.
In *SPDEs and Applications II*, pages 237–256. Springer, Berlin, Heidelberg, 1989.

Thank you for your attention!