Deep Decomposition for Stochastic Normal-Abnormal Transport
Supplementary Material

This supplementary material contains proofs for our representation theorems and additional implementation details for D^2-SONATA. Supp. A and Supp. B give the proofs for Theorem 1 and Theorem 2, respectively. Supp. C introduces our 2D/3D stochastic advection-diffusion PDE package in PyTorch and discusses numerical discretization and stability conditions. And Supp. D provides descriptions of how we generate our normal-abnormal brain advection-diffusion simulation dataset, including how we construct the velocity and diffusion fields with random abnormal patterns, and how we simulate the corresponding normal and abnormal brain advection-diffusion time-series.

A. Theorem 1: Anomaly-decomposed Divergence-free Vector Representation

For any vector field $\mathbf{V} \in L^p(\Omega)^d$ and scalar field A in $\mathbb{R}_{(0,1]}(\Omega)$ on a bounded domain $\Omega \subset \mathbb{R}^d$ with smooth boundary $\partial\Omega$. If $\mathbf{V}$ satisfies $\nabla \cdot \mathbf{V} = 0$, there exist a potential Ψ in $L^p(\Omega)^\alpha$ ($\alpha = 1(3)$ when $d = 2(3)$):

$$\mathbf{V} = \nabla \times (A\,\Psi), \quad (A\,\Psi) \cdot \mathbf{n}\big|_{\partial\Omega} = 0. \tag{12}$$

Conversely, for any $A \in \mathbb{R}_{(0,1]}(\Omega)$, $\Psi \in L^p(\Omega)^\alpha$, $\nabla \cdot \mathbf{V} = \nabla \cdot (\nabla \times (A\,\Psi)) = 0$.
(Here, L^p refers to the space of measurable functions for which the p-th power of the function absolute value is Lebesgue integrable. Specifically, let $1 \leq p < \infty$ and (Ω, Σ, μ) be a measure space. $L^p(\Omega)$ space is the set of all measurable functions whose absolute value raised to the p-th power has a finite integral, i.e., $\|f\|_p \equiv \left(\int_\Omega |f|^p \, \mathrm{d}\mu\right)^{1/p} < \infty$.)

Proof. The above Theorem is a Corollary of the following Theorem A.1, which is originally introduced by Liu et al. [37].

Theorem A.1 (Divergence-free Vector Field Representation by the Curl of Potentials). *For any vector field $\mathbf{V} \in L^p(\Omega)^d$ on a bounded domain $\Omega \subset \mathbb{R}^d$ with smooth boundary $\partial\Omega$. If $\mathbf{V}$ satisfies $\nabla \cdot \mathbf{V} = 0$, there exists a potential Ψ in $L^p(\Omega)^\alpha$ such that ($\alpha = 1$ when $d = 2$, $\alpha = 3$ when $d = 3$)*

$$\mathbf{V} = \nabla \times \Psi, \quad \Psi \cdot \mathbf{n}\big|_{\partial\Omega} = 0, \Psi \in L^p(\Omega)^\alpha. \tag{13}$$

Conversely, for any $\Psi \in L^p(\Omega)^\alpha$, $\nabla \cdot \mathbf{V} = \nabla \cdot (\nabla \times \Psi) = 0$.
(For detailed proof of Theorem A.1, please refer to Supp. A in [37].)

Therefore, it is obvious that given any vector field $\mathbf{V} \in L^p(\Omega)^d$ and scalar field $A \in \mathbb{R}_{(0,1]}(\Omega)$ on a bounded domain $\Omega \subset \mathbb{R}^d$ with smooth boundary $\partial\Omega$. If $\mathbf{V}$ satisfies $\nabla \cdot \mathbf{V} = 0$, according to Theorem A.1, there exists a $\widetilde{\Psi}$ such that:

$$\begin{aligned}
\mathbf{V} &\risingdotseq \nabla \times \widetilde{\Psi} \\
&\risingdotseq \nabla \times \left(A\,(\widetilde{\Psi}/A)\right) \\
&\risingdotseq \nabla \times (A\,\Psi) \qquad (14) \\
&\risingdotseq \nabla A \times \Psi + A\,\nabla \times \Psi \\
&\risingdotseq \nabla A \times \Psi + A\,\overline{\mathbf{V}}, \qquad (15)
\end{aligned}$$

where Eq. (14) corresponds to Eq. (12), and is the explicit expression for the implementation of "anomaly-encoded" velocity vector field $\mathbf{V}$ in this paper. $\overline{\mathbf{V}}$ in Eq. (15) refers to the "anomaly-free" velocity field as defined in Eq. (4) in the main manuscript where the relation between $\mathbf{V}$ and $\overline{\mathbf{V}}$ it is denoted as $\mathbf{V} = A \diamond \overline{\mathbf{V}}$.

□

B. Theorem 2: Anomaly-decomposed Symmetric PSD Tensor Representation

For any $n \times n$ symmetric PSD tensor $\mathbf{D}$ and $A \in \mathbb{R}_{(0,1]}(\Omega)$, there exist an upper triangular matrix with zero diagonal entries, $\mathbf{B} \in \mathbb{R}^{\frac{n(n-1)}{2}}$, and a non-negative diagonal matrix, $\Lambda \in SD(n)$, satisfying:

$$\mathbf{D} = \mathbf{U}\,(A\,\Lambda)\,\mathbf{U}^T, \quad \mathbf{U} = exp(\mathbf{B} - \mathbf{B}^T) \in SO(n). \tag{16}$$

Conversely, for $\forall A \in \mathbb{R}_{(0,1]}(\Omega)$*,* $\forall \mathbf{B} \in \mathbb{R}^{\frac{n(n-1)}{2}}$*, and any diagonal matrix with non-negative diagonal entries,* $\Lambda \in SD(n)$*, Eq. (16) results in a symmetric PSD tensor,* $\mathbf{D}$*.*

Proof. The above Theorem is a Corollary of the following Theorem B.1, which is originally introduced by Liu et al. [37].

Theorem B.1 (Symmetric PSD Tensor Representation by Spectral Decomposition). *For any tensor* $\mathbf{D} \in PSD(n)$*, there exists an upper triangular matrix with zero diagonal entries,* $\mathbf{B} \in \mathbb{R}^{\frac{n(n-1)}{2}}$*, and a diagonal matrix with non-negative diagonal entries,* $\Lambda \in SD(n)$*, satisfying:*

$$\mathbf{D} = \mathbf{U}\Lambda\mathbf{U}^T, \quad \mathbf{U} = exp(\mathbf{B} - \mathbf{B}^T) \in SO(n). \tag{17}$$

Conversely, for any upper triangular matrix with zero diagonal entries, $\mathbf{B} \in \mathbb{R}^{\frac{n(n-1)}{2}}$*, and any diagonal matrix with non-negative diagonal entries,* $\Lambda \in SD(n)$*, Eq. (17) results in a symmetric PSD tensor,* $\mathbf{D} \in PSD(n)$*.*
(For detailed proof of Theorem B.1, please refer to Supp. B in [37].)

Therefore, $\forall n \times n$ symmetric PSD tensors $\mathbf{D}$ and $\forall A \in \mathbb{R}_{(0,1]}(\Omega)$, there exists a $\mathbf{B} \in \mathbb{R}^{\frac{n(n-1)}{2}}$, and a $\widetilde{\Lambda} \in SD(n)$, such that:

$$\begin{aligned}
\mathbf{D} &\risingdotseq \mathbf{U}\widetilde{\Lambda}\mathbf{U}^T \\
&\risingdotseq \mathbf{U}\left(A\,(\widetilde{\Lambda}/A)\right)\mathbf{U}^T \\
&\risingdotseq \mathbf{U}\,(A\,\Lambda)\,\mathbf{U}^T \\
&\risingdotseq A\,\mathbf{U}\Lambda\mathbf{U}^T \qquad (18) \\
&\risingdotseq A\overline{\mathbf{D}}, \quad \mathbf{U} = exp(\mathbf{B} - \mathbf{B}^T) \in SO(n), \qquad (19)
\end{aligned}$$

where Eq. (18) corresponds to Eq. (16), and is the explicit expression for the implementation of "anomaly-encoded" diffusion tensor field $\mathbf{D}$ in this paper. $\overline{\mathbf{D}}$ in Eq. (19) refers to the "anomaly-free" diffusion tensor field as defined in Eq. (4) in the main manuscript where the relation between $\mathbf{D}$ and $\overline{\mathbf{D}}$ it is denoted as $\mathbf{D} = A \circ \overline{\mathbf{D}}$.

□

C. `PyTorch` Deterministic-Stochastic Advection-Diffusion PDE Toolkit: Numerical Derivations

Our stochastic advection-diffusion PDE toolkit is designed to solve, separately or jointly, deterministically or stochastically, advection and diffusion PDEs in 1D/2D/3D.

$$\begin{aligned}
\frac{\partial C(\mathbf{x},t)}{\partial t} = \left.\frac{\partial C(\mathbf{x},t)}{\partial t}\right|_{\text{adv.}} + \left.\frac{\partial C(\mathbf{x},t)}{\partial t}\right|_{\text{diff.}} + \left.\frac{\partial C(\mathbf{x},t)}{\partial t}\right|_{\text{sto.}} &= \underbrace{-\nabla\left(\mathbf{V}(\mathbf{x})\cdot C(\mathbf{x},t)\right)}_{\text{Fluid flow}} + \underbrace{\nabla\cdot\left(\mathbf{D}(\mathbf{x})\,\nabla C(\mathbf{x},t)\right)}_{\text{Diffusion}} + \underbrace{\sigma(\mathbf{x})\partial W(\mathbf{x},t)}_{\substack{\text{Model}\\\text{Uncertainty}}} \\
(\nabla\cdot\mathbf{V} = 0) &= \underbrace{-\mathbf{V}(\mathbf{x})\cdot\nabla C(\mathbf{x},t)}_{\text{Incompressible flow}} + \underbrace{\nabla\cdot\left(\mathbf{D}(\mathbf{x})\,\nabla C(\mathbf{x},t)\right)}_{\text{Diffusion}} + \underbrace{\sigma(\mathbf{x})\partial W(\mathbf{x},t)}_{\substack{\text{Model}\\\text{Uncertainty}}}.
\end{aligned} \tag{20}$$

One can choose to model the velocity field as a constant, a general vector field, or a divergence-free vector field (for incompressible flow). Furthermore, the diffusion field can be modeled as a constant, a non-negative scalar field, or a symmetric positive semi-definite (PSD) tensor field. By controlling$\sigma(\mathbf{x})$, one can specify the variance the Brownian motion $W(\mathbf{x},t)$, and thus the levels of uncertainty of the entire system. The toolkit is implemented as a custom `torch.nn.Module` subclass, such that one can directly use it as a (stochastic) advection-diffusion PDE solver for data simulation or easily wrap it into DNNs or numerical optimization frameworks for inverse (stochastic) PDE problems, i.e., parameters estimation.

C.1. Advection and Diffusion Computation

The computation of advection and diffusion and their stability analysis are similar to `YETI` [37], please refer to Supp. C.1-C.2 in [37] for a detailed discussion.

C.2. (Stochastic) Numerical Integration

As introduced above, after discretizing all the spatial derivatives on the right side of Eq. (20), we obtain a system of ordinary differential equations (ODEs), which can be solved by numerical integration [18].

If the system is modeled stochastically, the stochastic term, $\sigma(\mathbf{x})\partial W(\mathbf{x},t)$, should be included during the integration.

Definition 2 (Brownian motion). *A stochastic process* (W_t) *such that (1)* $W_0 = 0$*; (2)* $(W_t - W_s) \sim \mathcal{N}(0, t-s)$, $\forall t \geq s \geq 0$*; (3) For all disjoint time interval pairs* $[t_1, t_2]$, $[t_3, t_4]$ $(t_1 < t_2 \leq t_3 \leq t_4)$*, the increments* $W_{t_4} - W_{t_3}$ *and* $W_{t_2} - W_{t_1}$ *are independent random variables.*

We model the advection-diffusion process via a stochastic PDE (SPDE), where σ $(\sigma(\mathbf{x}) \in \mathbb{R})$ denotes the variance of the Brownian motion $W_t\,(W(\mathbf{x},t) \in \mathbb{R})$ [4] and represents the epistemic uncertainty for the dynamical system. With this additional stochastic term, the existence and uniqueness of the solution to Eq. (20) still holds:

Theorem 3 (Existence and uniqueness of SPDE [4, 49, 60]). *If the coefficients of the stochastic partial differential equation Eq. (20) with initial condition, satisfy the spatially-varying Lipschitz condition*

$$\begin{aligned} &|\mathbf{V}(\mathbf{x}_1) - \mathbf{V}(\mathbf{x}_2)|^2 + |\mathbf{D}(\mathbf{x}_1) - \mathbf{D}(\mathbf{x}_2)|^2 \\ &\qquad + |\sigma(\mathbf{x}_1) - \sigma(\mathbf{x}_2)|^2 \leq K|\mathbf{x}_1 - \mathbf{x}_2|^2, \end{aligned} \tag{21}$$

and the spatial growth condition

$$|\mathbf{V}(\mathbf{x})|^2 + |\mathbf{D}(\mathbf{x})|^2 + |\sigma(\mathbf{x})|^2 \leq K(1 + \mathbf{x}^2), \tag{22}$$

then there is a continuous adapted solution $C(\mathbf{x},t)$ *satisfying the* L^2 *bound. Moreover, if* $C(\mathbf{x},t)$ *and* $\widetilde{C}(\mathbf{x},t)$ *are both continuous solutions satisfying the* L^2 *bound, then*

$$P\left(C(\mathbf{x},t) = \widetilde{C}(\mathbf{x},t) \text{ for all } t \in [0,T]\right) = 1. \tag{23}$$

For a detailed discussion regarding Theorem 3, please refer to [4, 49, 60]. During implementation, we use the Euler-Maruyama scheme [27, 28], and the discretized version of the stochastic term in Eq. (20) can therefore be written as $\sigma(\mathbf{x})\,W(\mathbf{x},t)/\sqrt{\Delta t}$, where $W \sim \mathcal{N}(0,1)$.

We then use the RK45 method to advance in time (δt) to predict $\widehat{C}^{t+\delta t}$. Note when the input mass transport time-series has relatively large temporal resolution (Δt), the chosen δt should be smaller than Δt to satisfy the stability conditions (Supp. C.1), thereby ensuring stable numerical integration.

D. Normal-Abnormal Brain Advection-Diffusion Dataset

Our brain advection-diffusion simulation dataset is based on the public `IXI` brain dataset[8], from which we use 200 patients with complete collections of T1-/T2-weighted images, magnetic resonance angiography (MRA) images, and diffusion-weighted images (DWI) with 15 directions for the simulation of 3D divergence-free velocity vector and symmetric PSD diffusion tensor fields. All images above are resampled to isotropic spacing ($1\,mm$), rigidly registered intra-subject (according to the MRA image), and brain-extracted using `ITK`[9].

In general, the generation of the "anomaly-free" velocity vector fields and "anomaly-free" diffusion tensor fields are of the same procedures as in [37] (Please refer to Supp. D.1-D.2 in [37]). Here, Supp. D.1 provides the anomaly value fields (A) simulation proposed in this paper, and Supp. D.2 introduces the time-series simulation for normal-abnormal advection-diffusion processes.

D.1. "Anomaly-encoded" Velocity Vector and Diffusion Tensor Fields (Fig. D.6)

For each case, we simulate both normal samples and anomaly-encoded samples. For samples treated as normal, we directly use the simulated $\mathbf{V}$ and $\mathbf{D}$ for the concentration time-series simulation. For anomaly encoding, the originally simulated $\mathbf{V}$ and $\mathbf{D}$ are treated as the "anomaly-free" fields for generating anomaly-encoded fields, $\overline{\mathbf{V}}$ and $\overline{\mathbf{D}}$, via Eq. (4). For the generation of the anomaly value field A, its spatially varying value (within $[0,1]$) is determined by a Gaussian with center uniformly sampled across the entire spatial domain of the sample case.

[8] Available for download: http://brain-development.org/ixi-dataset/.
[9] Code in https://github.com/InsightSoftwareConsortium/ITK.

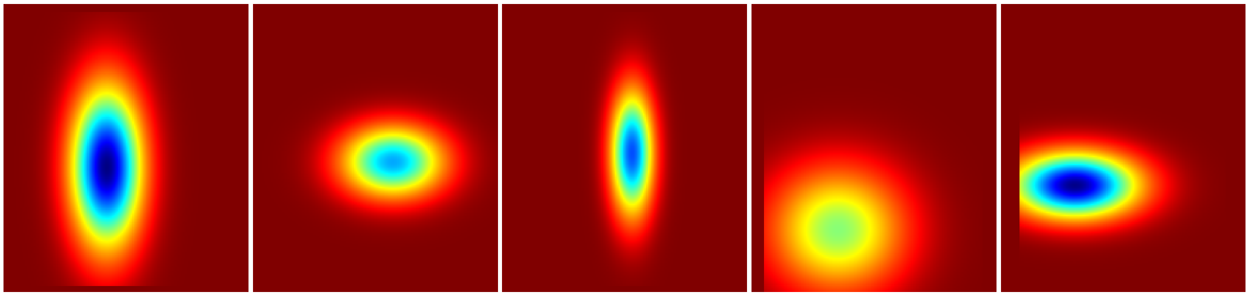

Figure D.6. Examples of different generated anomaly value field patterns applied on five different samples.

D.2. Brain Advection-diffusion Time-series Simulation

For each brain advection-diffusion sample, the initial concentration state is assumed to be given by the MRA image with intensity ranges rescaled to $[0, 1]$. Time-series (length $N_T = 40$, interval $\Delta t = 0.1\,s$) are then simulated given the computed divergence-free velocity fields and symmetric PSD diffusion tensor fields by our advection-diffusion PDE solver (Supp. C). Thus the simulated dataset includes 800 brain advection-diffusion time-series (4 time-series for each of the 200 subjects, based on the four combinations of the simulated two velocity fields and two diffusion fields).