Shapes analysis for time series

Thibaut Germain*, Samuel Gruffaz, Charles Truong, Alain Durmus, & Laurent Oudre

NEURAL INFORMATION PROCESSING SYSTEMS

Objective:

This study aims to analyze **inter-individual variability** within a time series dataset characterized by **irregular sampling** intervals and **variable sequence lengths**.

Methodology:

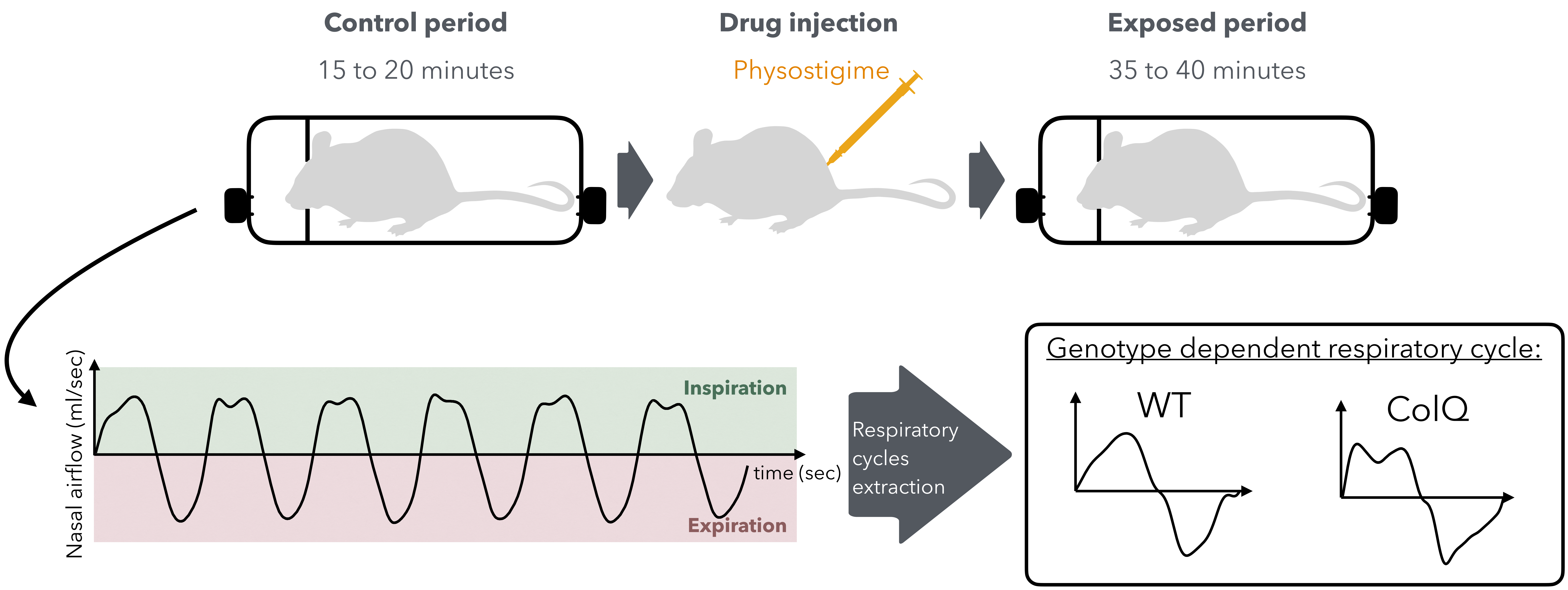
Unsupervised representation learning with an emphasis on capturing **shape structure**.

* Contact author, email: thibaut.germain@ens-paris-saclay.fr

école normale supérieure paris—saclay | université PARIS-SACLAY | CMAP | CENTRE BORELLI

An example: mice ventilation analysis

The following experiment was performed for mice of different genotypes (ColQ or WT).



Methodology

Input : $\left(s_j : \mathrm{I}_j \mapsto \mathbb{R}^d\right)_{j \in [N]}$

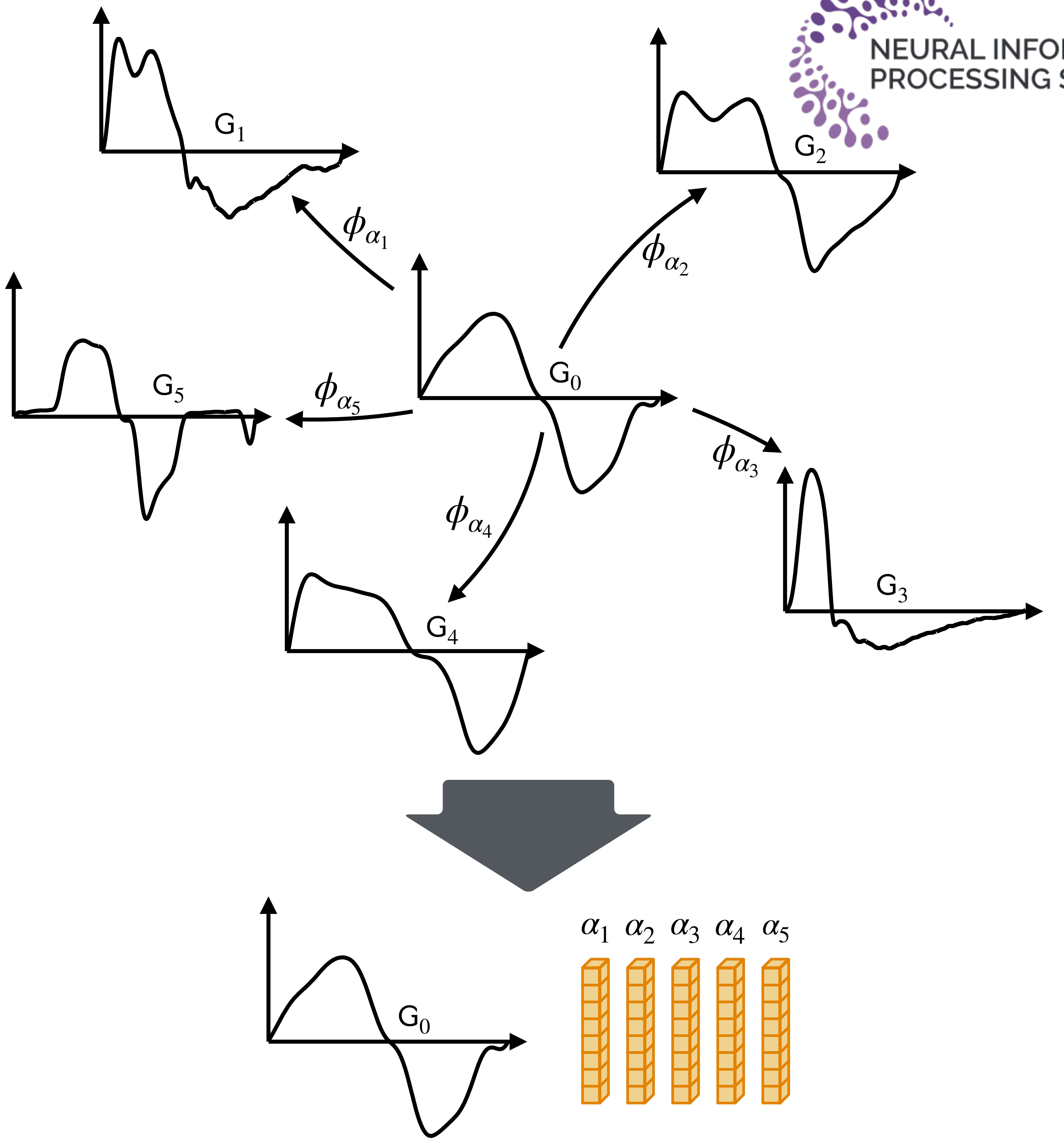
1. The shape of s_j is $\mathrm{G}_j = \{(t, s_j(t) : t \in \mathrm{I}_j\}$
2. G_j is represented as the deformation of a reference graph $\mathrm{G}_0 = \{(t, s_0(t) : t \in \mathrm{I}\}$, i.e. :

$$\mathrm{G}_j \approx \phi_{\alpha_j} \cdot \mathrm{G}_0 = \{\phi_{\alpha_j}\left(t, s_0(t)\right) : t \in \mathrm{I}\} ,$$

   where $\phi_{\alpha_j} : \mathbb{R}^{d+1} \mapsto \mathbb{R}^{d+1}$ is a diffeomorphism parametrized by $\alpha_j \in \mathbb{R}^m$.

3. Learned by solving the empirical Fréchet mean:

$$\underset{\mathrm{G}_0, (\alpha_j)_j \in [N]}{\arg\min} \frac{1}{N} \sum_{j \in [N]} \left(d^2_{\mathrm{G}}(\phi_{\alpha_j} \cdot \mathrm{G}_0, \mathrm{G}_j) + \lambda d^2_{\Phi}(Id, \phi_{\alpha_j}) \right)$$



Output : a graph of reference $\mathrm{G}_0 \subset \mathbb{R}^{d+1}$, deformation parameters $(\alpha_j)_{j \in [N]} \in (\mathbb{R}^m)^N$

Overview

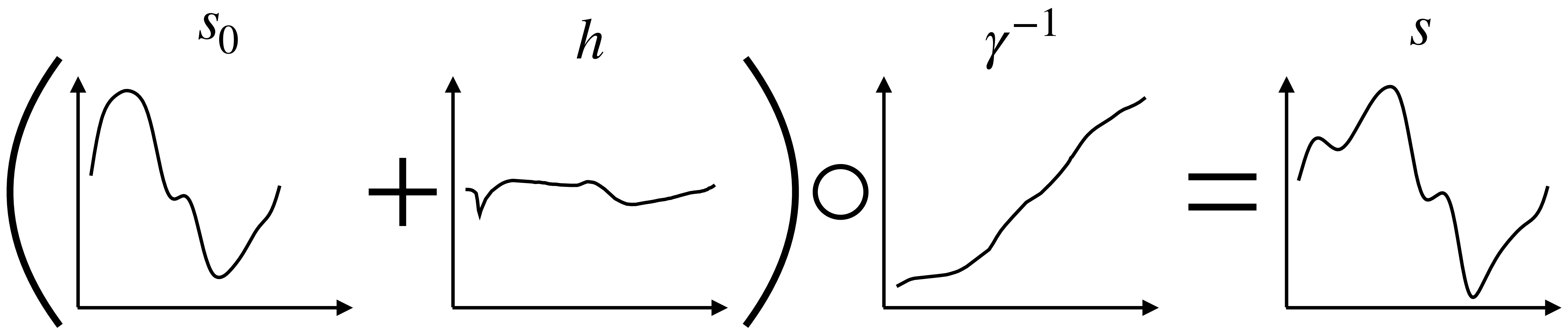
The optimisation problem.
Solved by gradient descent:

$$\underset{G_0,(\alpha_j)_j\in[N]}{\arg\min} \frac{1}{N}\sum_{j\in[N]}\Big(\underbrace{d^2_{G}(\phi_{\alpha_j}\cdot G_0, G_j)}_{(a)} + \lambda\ \underbrace{d^2_{\Phi}(Id,\phi_{\alpha_j})}_{(b)}\Big)$$

(a) Distance measuring the similarity between $\phi_{\alpha_j} \cdot G_0$ and G_j embedded as varifold measure [1]. This distance is similar to Maximum Mean Discrepancy (MMD) and is presented in <u>*Section 4.1*</u>.

(b) Distance comparing ϕ_{α_j} and Id to privilege minimal diffeomorphic deformation and prevent overfitting. Presented in <u>*Section 3*</u>.

Deformation requirements. Diffeomorphic deformations are built with the LDDMM framework [2], and presented in <u>*Section 3*</u>. Our contributions, presented in <u>*Section 4*</u>, impose that for any $s_0 : I \mapsto \mathbb{R}^d$ and $s : J \mapsto \mathbb{R}^d$, the diffeomorphism ϕ mapping s_0 to s is the combination of a distortion $h : I \mapsto \mathbb{R}$ and a time parametrization $\gamma^{-1} : J \mapsto I$ such that: $\phi \cdot G(s_0) = G((s_0 + h) \circ \gamma^{-1}) = G(s)$



[1] Kaltenmark, I., Charlier, B., & Charon, N. (2017). A general framework for curve and surface comparison and registration with oriented varifolds.
[2] Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. Computing large deformation metric mappings via geodesic flows of diffeomorphisms.

Key elements from LDDMM [1]

Generating diffeomorphisms. Assuming $v \in \mathsf{L}^2([0,1], \mathsf{V})$ a time-varying velocity field in $\mathbb{R}^n$, where V is an RKHS with some regularity assumptions [1]. For any $x_0 \in \mathbb{R}^n$, the differential system:

$$\frac{\mathrm{d}X(\tau)}{\mathrm{d}\tau} = v_\tau(X(\tau)) \quad \text{with} \quad X(0) = x_0$$

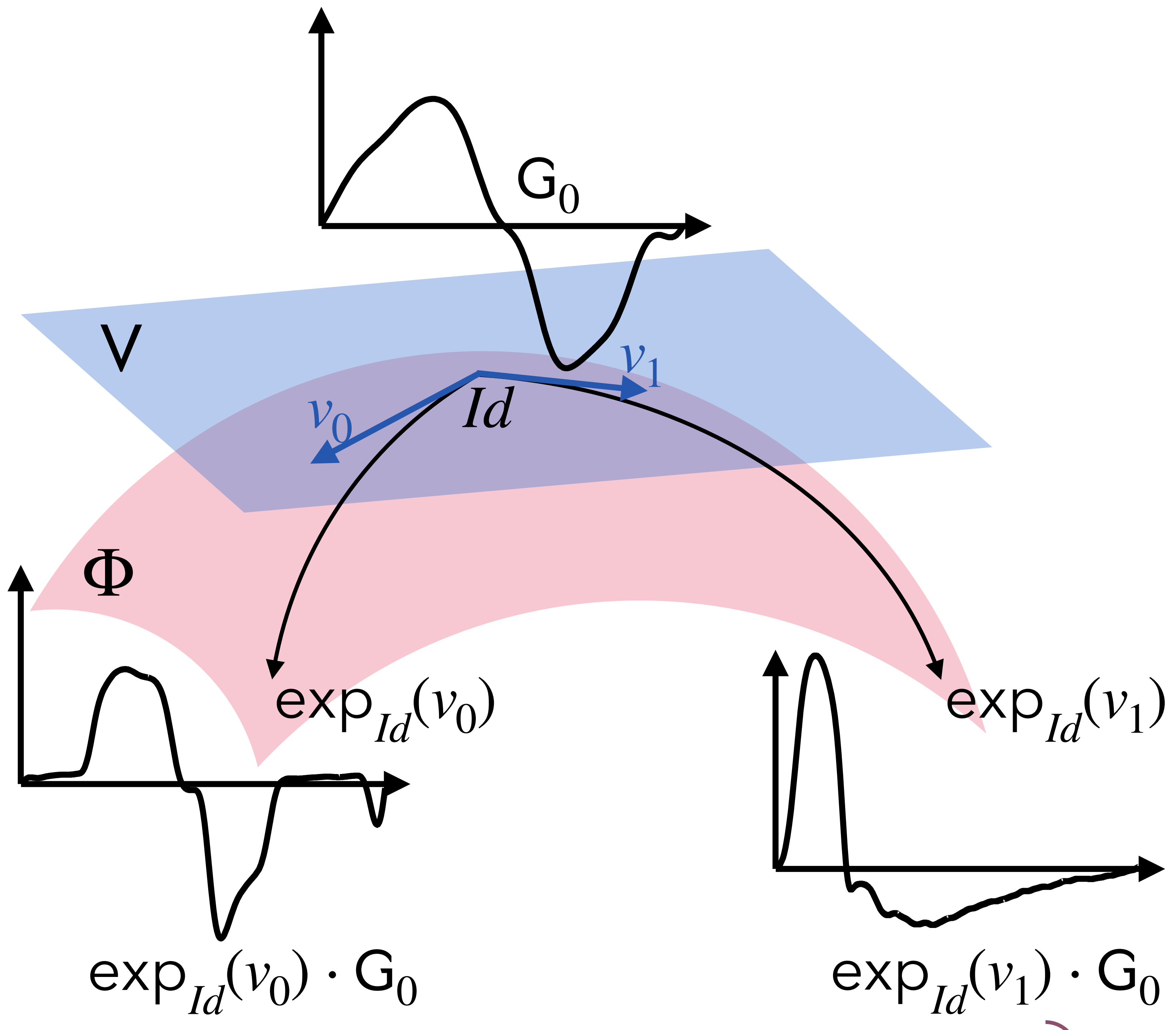
has a unique solution defined for all $\tau \in [0,1]$. The flow application: $\phi_v^\tau : x_0 \in \mathbb{R}^n \mapsto X(\tau) \in \mathbb{R}^n$ solution of (1) at time $\tau \in [0,1]$ is a diffeomorphism. Our interest is in the group of diffeomorphisms: $\Phi \triangleq \{\phi_v^1 \mid v \in \mathsf{L}^2([0,1], \mathsf{V})\}$.

Geodesic shooting. Geodesic flow from Id with initial velocity field $v_0 \in \mathsf{V}$ can be defined [2]. By denoting $\tau \mapsto \rho_{v_0}(\tau)$ the geodesic starting from Id with initial conditions $v_0 \in \mathsf{V}$, the exponential map is:

$$\exp_{Id} : v_0 \in \mathsf{V} \mapsto \rho_{v_0}(1) \in \Phi \quad \text{and} \quad d_\Phi^2(Id, \exp_{Id}(v_0)) = \|v_0\|_{\mathsf{V}}^2$$

In practice, v_0 is parametrized by K, the kernel associated with V, the sampled time series graph $\mathsf{G}_0 \in (\mathbb{R}^{d+1})^{N_0}$ and the parameters $\alpha_0 \in (\mathbb{R}^{d+1})^{N_0}$ such that:

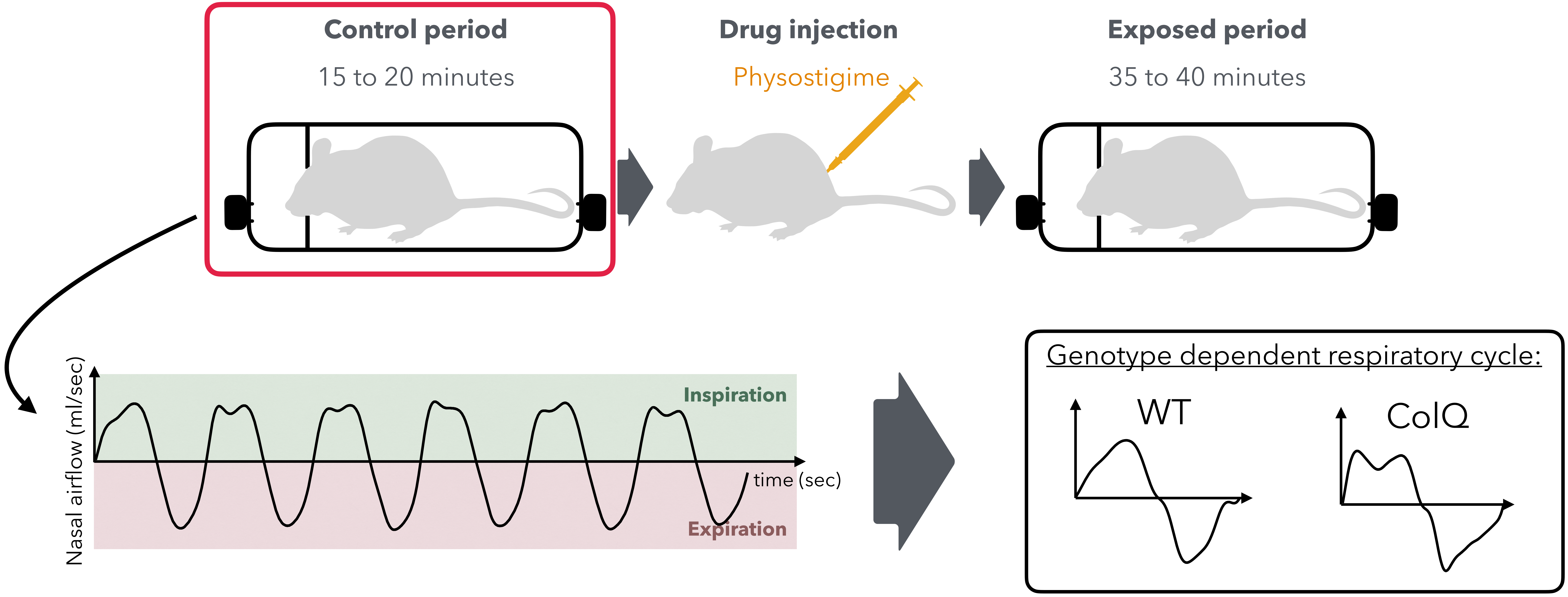
$$v_0 : x \in \mathbb{R}^{d+1} \mapsto \sum_{k \in [N_0]} K(g_0^k, x)\alpha_0^k \in \mathbb{R}^{d+1}$$



[1]. Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. Computing large deformation metric mappings via geodesic flows of diffeomorphisms.
[2] Miller, M. I., Trouvé, A., and Younes, L.. Geodesic shooting for computational anatomy.

An example: mice ventilation analysis before drug injection

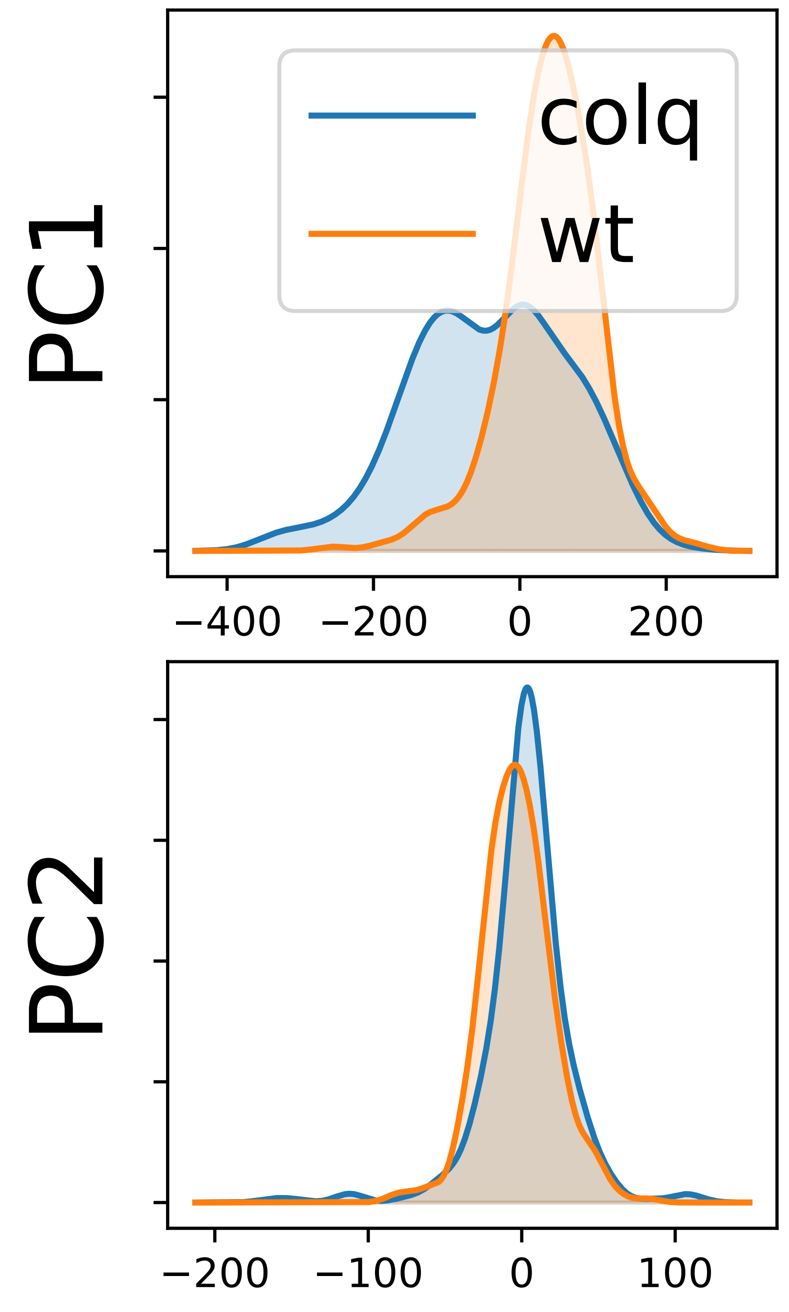
For mice of different genotype (ColQ or WT):



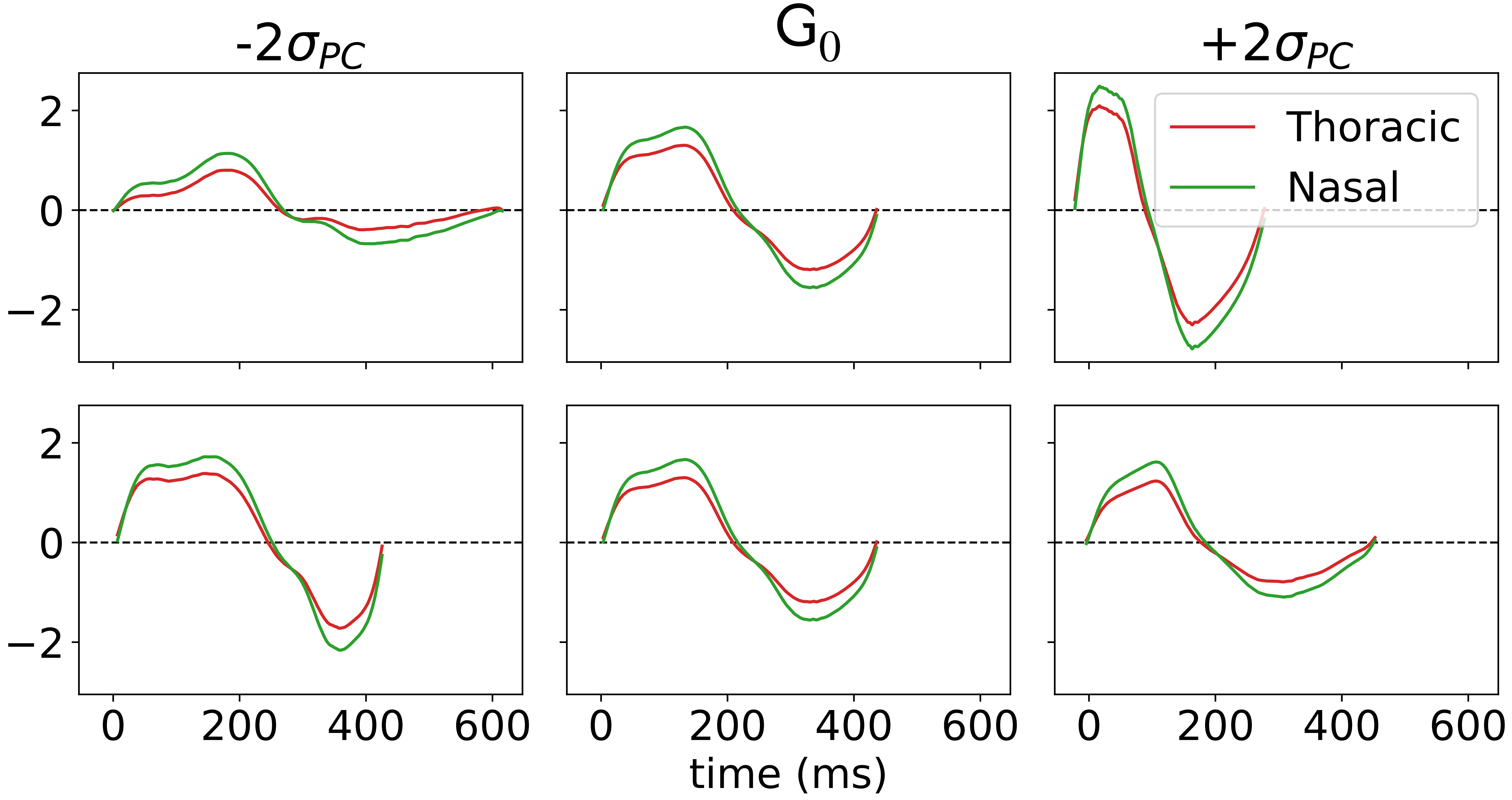
Mice ventilation analysis before drug injection

Kernel PCA is applied on the initial velocity field parameters $(\mathrm{G}_0, \alpha_i)_{i \in [N]}$, resulting in the K principal axis of deformations with the initial velocity field $(\mathrm{G}_0, \alpha_j^{pc})_{j \in [K]}$.

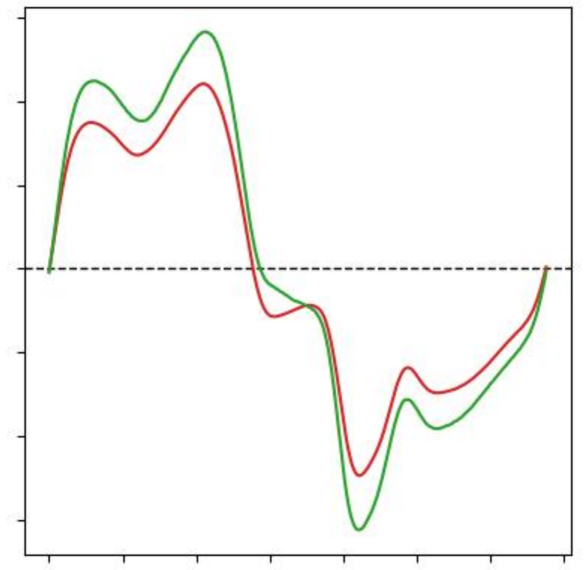
PC densities



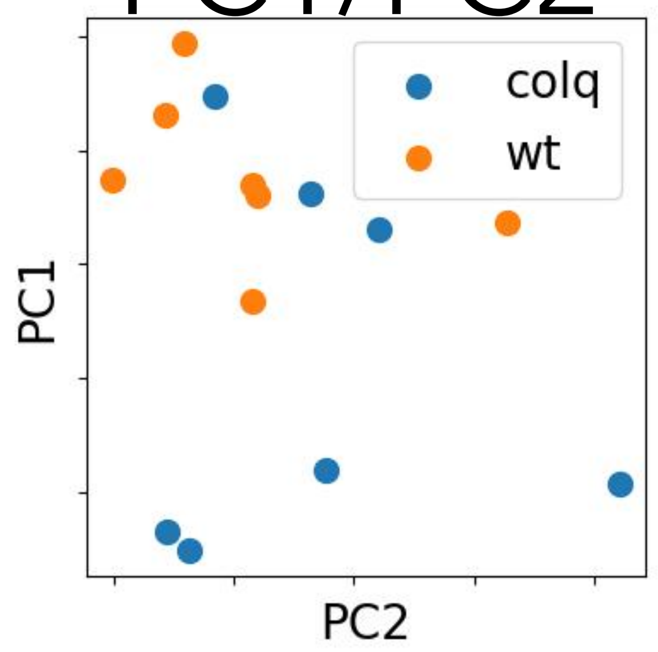
Deformations by flowing along PCs:



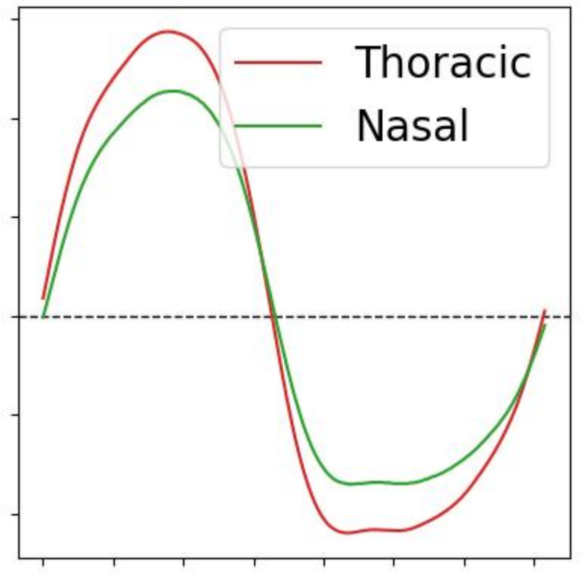
ColQ



PC1/PC2



WT



Thank you !

More details are provided in the next slides

I. Building diffeomorphisms with LDDMM [1]

NEURAL INFORMATION PROCESSING SYSTEMS

Generating diffeomorphisms. Assuming $v \in \mathrm{L}^2([0,1], \mathrm{V})$ a time-varying velocity field in $\mathbb{R}^n$, where V is an RKHS with some regularity assumptions [1]. For any $x_0 \in \mathbb{R}^n$, the differential system:

$$\frac{\mathrm{d}X(\tau)}{\mathrm{d}\tau} = v_\tau(X(\tau)) \quad \text{with} \quad X(0) = x_0\,, \tag{1}$$

has an unique solution defined for all $\tau \in [0,1]$. The flow application: $\phi_v^\tau : x_0 \in \mathbb{R}^n \mapsto X(\tau) \in \mathbb{R}^n$ solution of (1) at time $\tau \in [0,1]$ is a diffeomorphism.

A metric group. The $\Phi \triangleq \{\phi_v^1 \mid v \in \mathrm{L}^2([0,1], \mathrm{V})\}$ is metrizable such that for any $\phi \in \Phi$:

$$d_\Phi^2(Id, \phi) = \inf_{v \in \mathrm{L}^2([0,1],\mathrm{V})} \left\{ \int_0^1 \|v_\tau\|_\mathrm{V}^2 \mid \phi_v^1 = \phi \right\}, \tag{2}$$

the infimum is reached with a v^* and it conserves its norm along its geodesic path i.e.: $\|v_\tau^*\|_\mathrm{V} = \|v_0^*\|_\mathrm{V}$, $\forall \tau \in [0,1]$.

An exponential map. Geodesic flow from Id with initial velocity field $v_0 \in \mathrm{V}$ can be derived from (2) [2]. By denoting $\tau \mapsto \rho_{v_0}(\tau)$ the geodesic starting from Id with initial conditions $v_0 \in \mathrm{V}$, the exponential map is:

$$\exp_{Id} : v_0 \in \mathrm{V} \mapsto \rho_{v_0}(1) \in \Phi \quad \text{and} \quad d_\Phi^2(Id, \exp_{Id}(v_0)) = \|v_0\|_\mathrm{V}^2$$

[1]. Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. Computing large deformation metric mappings via geodesic flows of diffeomorphisms.
[2] Miller, M. I., Trouvé, A., and Younes, L.. Geodesic shooting for computational anatomy.

école normale supérieure paris—saclay

université PARIS-SACLAY

CMAP

CENTRE BORELLI

I. Computing the Exponential map [1]

Let denote $K : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n^2}$ the kernel of the RKHS V.

Given N_0 control points $\mathsf{X}_0 = (x_k^0)_{k\in[N_0]} \in (\mathbb{R}^n)^{N_0}$, and momentums $\alpha = (\alpha_k^0)_{k\in[N_0]} \in (\mathbb{R}^n)^{N_0}$, the initial velocity field is,

$$v_0 : x \in \mathbb{R}^n \mapsto \sum_{i\in[N_0]} K(x_k^0, x)\alpha_k^0 \in \mathbb{R}^n .$$

Then, for any $\tau \in [0,1]$,

$$v_\tau : x \in \mathbb{R}^n \mapsto \sum_{i\in[N_0]} K(x_k(\tau), x)\alpha_k(\tau) \in \mathbb{R}^n,$$

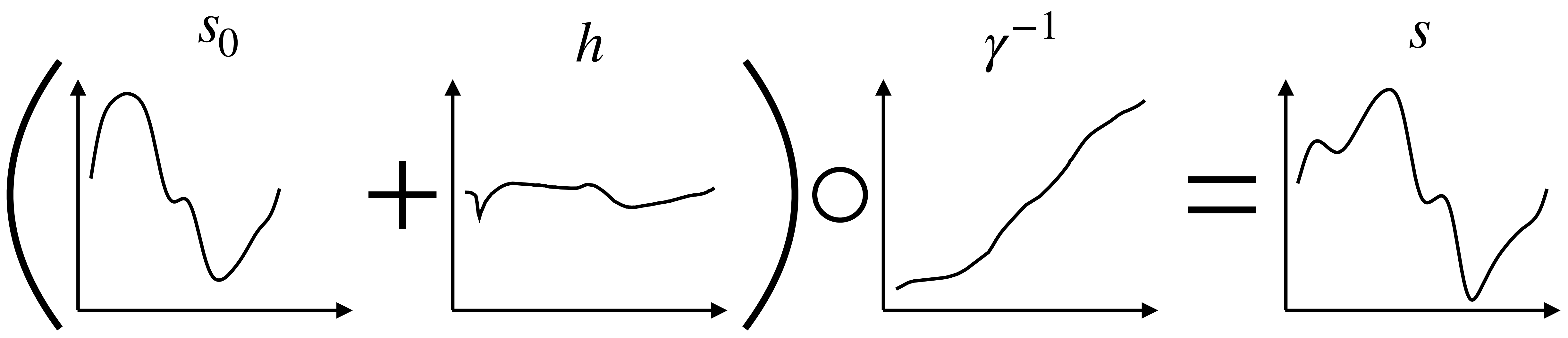
governed by the geodesic equations:

$$(\mathrm{E}) \begin{cases} \frac{\mathrm{d}x_k(\tau)}{\mathrm{d}\tau} = v_\tau(x_k(\tau)) \\ \frac{\mathrm{d}\alpha_k(\tau)}{\mathrm{d}\tau} = -\sum_{l\in N_0} \mathrm{d}_{x_k(\tau)} K(x_k(\tau), x_l(\tau))\alpha_l(\tau)^\top \alpha_k(\tau) \end{cases} \quad \text{with } \forall k \in [N_0] \quad \begin{cases} x_k(0) = x_k^0 \\ \alpha_k(0) = \alpha_k^0 \end{cases}$$

[1] Miller, M. I., Trouvé, A., and Younes, L. Geodesic shooting for computational anatomy.

II. Time series deformation representation

Intuition. Let $s_0 : \mathrm{I} \mapsto \mathbb{R}^d$ and $s : \mathrm{J} \mapsto \mathbb{R}^d$, the diffeomorphic deformation ϕ mapping s_0 to s should be seen as distortion $h : \mathrm{I} \mapsto \mathbb{R}$ and a time parametrization $\gamma^{-1} : \mathrm{J} \mapsto \mathrm{I}$ such that: $\phi \cdot \mathrm{G}(s_0) = \mathrm{G}((s_0 + h) \circ \gamma^{-1}) = \mathrm{G}(s)$



Theorem. For any continuously differentiable time series $s_0 : \mathrm{I} \mapsto \mathbb{R}^d$ and $s : \mathrm{J} \mapsto \mathbb{R}^d$, there exists deformations $\Psi_\gamma : (t, x) \in \mathbb{R}^{d+1} \mapsto (\gamma(t), x) \in \mathbb{R}^{d+1}$ with $\gamma \in \mathrm{D}(\mathbb{R})$, and $\Pi_f : (t, x) \in \mathbb{R}^{d+1} \mapsto (t, f(t, x)) \in \mathbb{R}^{d+1}$ with $f \in \mathrm{C}^1(\mathbb{R}^{d+1}, \mathbb{R}^d)$, such that:

$$\phi_{\gamma,f} \cdot \mathrm{G}(s_0) = \mathrm{G}(s) \quad \text{with} \quad \phi_{\gamma,f} = \Psi_\gamma \circ \Pi_f ,$$

Moreover, for any $\bar{\gamma} \in \mathrm{D}(\mathbb{R})$, and $\bar{f} \in \mathrm{C}^1(\mathbb{R}^{d+1}, \mathbb{R}^d)$, $\phi_{\bar{\gamma},\bar{f}} \cdot \mathrm{G}(s_0)$ is the graph of a continuously differentiable time series.

Remark. For any time series $s_0 : \mathrm{I} \mapsto \mathbb{R}^d$ and deformation $\phi_{\bar{\gamma},\bar{f}}$, the time parametrization and the distortion are:

$$\begin{cases} \gamma^{-1} : t \in \bar{\gamma}(\mathrm{I}) \mapsto \bar{\gamma}^{-1}(t) \in \mathrm{I} \\ h : t \in \mathrm{I} \mapsto \bar{f}(t, s_0(t)) - s_0(t) \end{cases}$$

II. A kernel for time series deformations

Gaussian kernel. For any $n \in \mathbb{N}^*$ and $\sigma > 0$, the one-dimensional Gaussian kernel is defined as,

$$K_\sigma^{(n)} : (x, y) \in \mathbb{R}^b \times \mathbb{R}^n \mapsto \exp(-\|x - y\|^2/\sigma).$$

The proposed kernel. We consider the kernel K_G defined for any $(t, x), (t', x') \in (\mathbb{R}^{d+1})^2$,

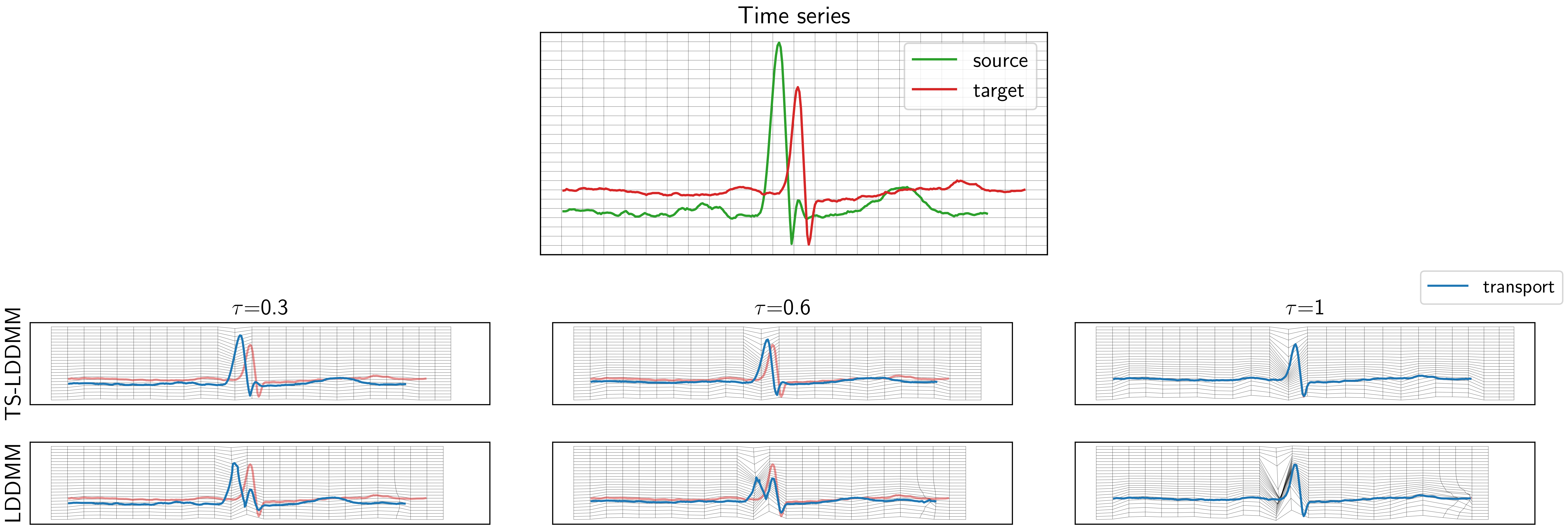
$$K_\mathrm{G}((t,x),(t',x')) = \begin{pmatrix} c_0 K_\mathrm{time} & 0 \\ 0 & c_1 K_\mathrm{space} \end{pmatrix} \quad \text{with} \quad \begin{cases} K_\mathrm{time} = K^{(1)}_{\sigma_{T,0}}(t, t') \\ K_\mathrm{space} = K^{(1)}_{\sigma_{T,1}}(t, t') K^{(d)}_{\sigma_x}(x, x') I_d \end{cases}$$

parametrized by $\sigma_{T,0}, \sigma_{T,1}, \sigma_x > 0$ and the constants $c_0, c_1 > 0$.

Lemma. For any initial velocity field $v_0 \in \mathrm{V}$, the RKHS associated to K_G, the diffeomorphic deformations learned by geodesic shooting ensures a time series graph structure along its geodesic path, i.e.

For any $\tau \in [0,1]$, there exist $\gamma_\tau \in \mathrm{D}(\mathbb{R})$ and $f_\tau \in \mathrm{C}^1(\mathbb{R}^{d+1}, \mathbb{R}^d)$ such that $\exp_{Id}(\tau v_0) = \Psi_{\gamma_\tau} \circ \Pi_{f_\tau}$.

II. Difference between LDDMM and TS-LDDMM



- Large Deformation Diffeomorphic Metric Mapping (LDDMM) with a RBF kernel
- TS-LDDMM an adaptation of LDDMM to time series (our contributions).

III. The varifold distance between time series graph [1]

Let $\mathsf{G} = (g_j)_{j\in[T]} \in (\mathbb{R}^{d+1})^T$ be a sampled time series graph. The approximate varifold representation of G is the measure,

$$\mu_{\mathsf{G}} = \sum_{j\in[T-1]} l_j \delta_{(x_j,\vec{v_j})} \quad \text{with} \quad \begin{cases} l_j = \|g_{j+1} - g_j\| \\ x_j = (g_j + g_{j+1})/2 \\ \vec{v_j} = (g_{j+1} - g_j)/\|g_{j+1} - g_j\| \end{cases}$$

Assuming that test functions belong to the dual of an RKHS W with kernel $k = k_{pos} \otimes k_{dir} : \mathbb{R}^{d+1} \times \mathbb{S}^d \mapsto \mathbb{R}$, such that,

$$\langle \delta_{(x_1,\vec{v_1})}, \delta_{(x_2,\vec{v_2})} \rangle_{\mathsf{W}^*} = k_{pos}(x_1, x_2) k_{dir}(\vec{v_1}, \vec{v_2}).$$

The similarity between time series graphs $\mathsf{G}_1 = (g_j^1)_{j\in[T_1]}$ and $\mathsf{G}_2 = (g_j^2)_{j\in[T_2]}$ is given by,

$$d_{\mathsf{G}}^2(\mathsf{G}_1, \mathsf{G}_2) = \|\mu_{\mathsf{G}_1} - \mu_{\mathsf{G}_2}\|_{\mathsf{W}^*}^2 = \langle \mu_{\mathsf{G}_1}, \mu_{\mathsf{G}_1} \rangle_{\mathsf{W}^*} - 2\langle \mu_{\mathsf{G}_1}, \mu_{\mathsf{G}_2} \rangle_{\mathsf{W}^*} + \langle \mu_{\mathsf{G}_2}, \mu_{\mathsf{G}_2} \rangle_{\mathsf{W}^*}$$

with,

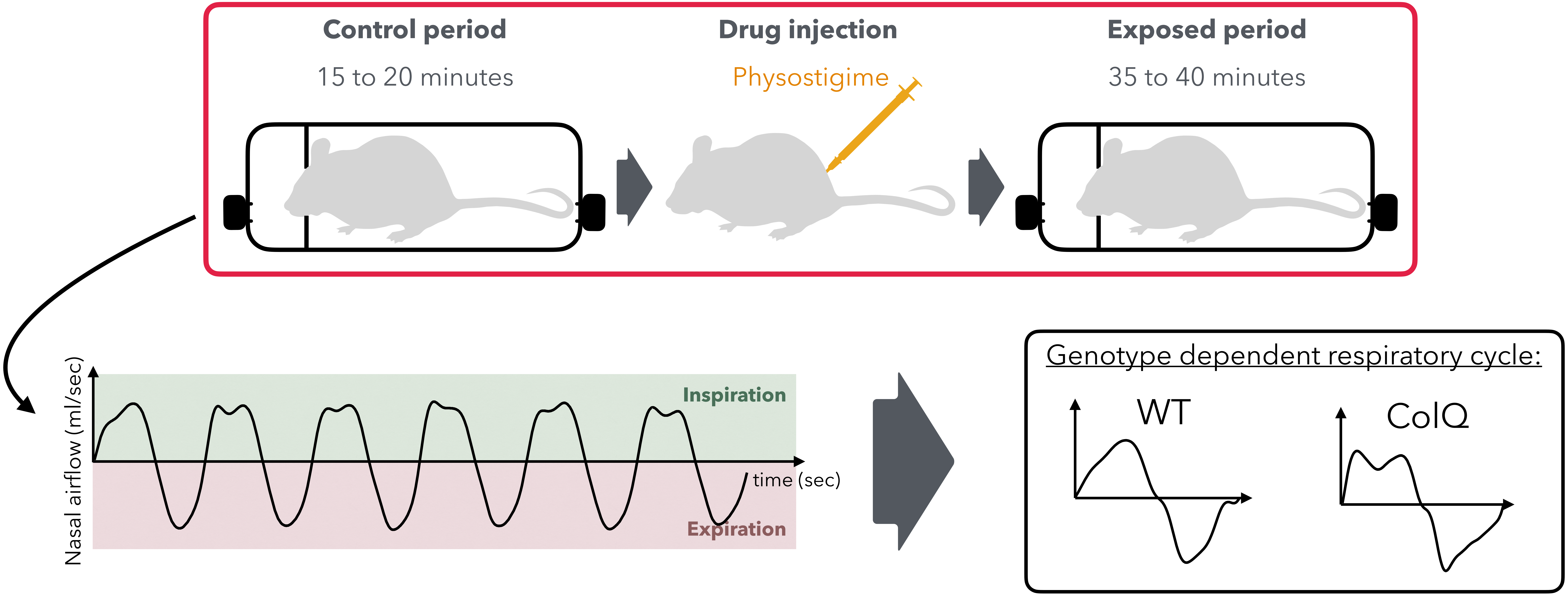
$$\langle \mu_{\mathsf{G}_1}, \mu_{\mathsf{G}_2} \rangle_{\mathsf{W}^*} = \sum_{i\in[T_1-1]} \sum_{j\in[T_2-1]} l_i^1 l_j^2 k_{pos}(x_i^1, x_j^2) k_{dir}(\vec{v}_i^{\,1}, \vec{v}_j^{\,2})$$

[1] Kaltenmark, I., Charlier, B., & Charon, N. (2017). A general framework for curve and surface comparison and registration with oriented varifolds.

Further experiments

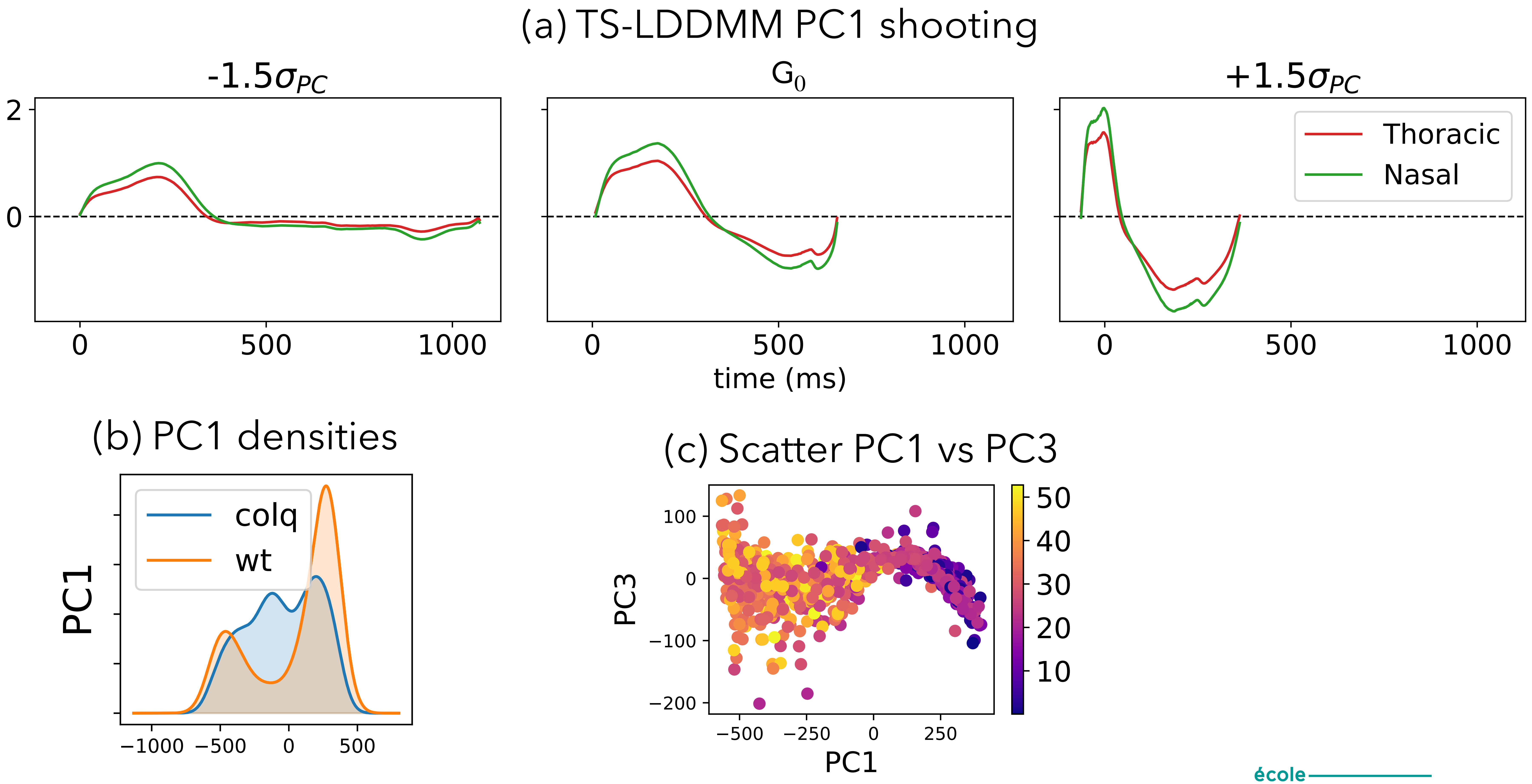
Mice ventilation analysis before/after drug injection

For mice of different genotype (ColQ or WT):



Mice ventilation analysis before/after drug injection

Kernel PCA is applied on the initial velocity field parameters $(\mathrm{G}_0, \alpha_i)_{i\in[N]}$, resulting in the K principal axis of deformations with the initial velocity field $(\mathrm{G}_0, \alpha_j^{pc})_{j\in[K]}$.



Benchmark on classification task on 15 UCR/UEA datasets:

Robustness to irregular sampling, comparison with state-of-the-art in deep learning.

Comparison of average f1score (macro) and ranks as the sample dropping rate increases. **First** & <u>second</u> best performers. TS-LDDMM is the best performer on three out of four dropping rates.

Methods	Regular		30 % dropped		50 % dropped		70 % dropped	
	F1-score	Rank	F1-score	Rank	F1-score	Rank	F1-score	Rank
RNN (1999)	0.64 ± 0.21	6.2	0.53 ± 0.23	6.6	0.48 ± 0.21	7.2	0.44 ± 0.21	6.07
LSTM (1997)	0.61 ± 0.29	6.0	0.57 ± 0.29	6.27	0.53 ± 0.25	6.07	0.51 ± 0.29	5.27
GRU (2014)	0.71 ± 0.26	4.2	0.68 ± 0.28	4.27	0.66 ± 0.28	3.73	<u>0.59 ± 0.28</u>	<u>3.67</u>
MTAN (2021)	0.59 ± 0.28	7.13	0.58 ± 0.28	5.8	0.54 ± 0.29	5.33	0.51 ± 0.28	5.0
MIAM (2022)	0.48 ± 0.35	6.93	0.42 ± 0.33	8.27	0.47 ± 0.31	6.93	0.35 ± 0.31	7.6
ODE-LSTM (2020)	0.63 ± 0.24	6.0	0.57 ± 0.25	6.53	0.51 ± 0.24	7.27	0.45 ± 0.23	6.73
Neural SDE (2019)	0.48 ± 0.28	7.67	0.47 ± 0.26	7.47	0.45 ± 0.27	7.13	0.45 ± 0.25	6.0
Neural LNSDE (2024)	0.7 ± 0.27	<u>3.87</u>	0.68 ± 0.29	<u>4.0</u>	<u>0.67 ± 0.25</u>	<u>3.53</u>	$\mathbf{0.66 \pm 0.23}$	**2.47**
LDDMM (2008)	<u>0.72 ± 0.2</u>	4.53	<u>0.7 ± 0.21</u>	4.2	0.57 ± 0.25	5.0	0.4 ± 0.25	7.13
TS-LDDMM (ours)	$\mathbf{0.83 \pm 0.18}$	**2.93**	$\mathbf{0.8 \pm 0.18}$	**2.07**	$\mathbf{0.7 \pm 0.26}$	**3.33**	0.51 ± 0.27	5.67

Benchmark on classification task on 15 UCR/UEA datasets

Regular sampling, comparison with state-of-the-art in Functional Data Analysis.

Comparison of average f1score (macro) between methods from shape analysis and functional data analysis.
First & <u>second</u> best performers.

	Dataset	Shape-FPCA (2024)	TCLR (2024)	LDDMM (2008)	TS-LDDMM (ours)
Univariate	ArrowHead	0.18	0.75	<u>0.84</u>	**0.91**
	BME	0.16	<u>1.00</u>	0.82	**1.00**
	ECG200	0.40	0.67	**0.81**	<u>0.79</u>
	FacesUCR	0.08	<u>0.73</u>	0.69	**0.86**
	GunPoint	0.93	<u>0.97</u>	0.83	**1.00**
	PhalangesOutlinesCorrect	0.39	**0.63**	<u>0.53</u>	0.52
	Trace	0.55	<u>1.00</u>	0.46	**1.00**
Multivariate	ArticularyWordRecognition	–	–	<u>0.98</u>	**1.00**
	Cricket	–	–	<u>0.77</u>	**0.93**
	ERing	–	–	<u>0.95</u>	**0.98**
	Handwriting	–	–	<u>0.22</u>	**0.44**
	Libras	–	–	<u>0.56</u>	**0.60**
	NATOPS	–	–	<u>0.82</u>	**0.82**
	RacketSports	–	–	**0.83**	<u>0.79</u>
	UWaveGestureLibrary	–	–	<u>0.72</u>	**0.81**