



HAL
open science

Learning spatiotemporal piecewise-geodesic trajectories from longitudinal manifold-valued data

Juliette Chevallier, Stéphane Oudard, Stéphanie Allasonnière

► To cite this version:

Juliette Chevallier, Stéphane Oudard, Stéphanie Allasonnière. Learning spatiotemporal piecewise-geodesic trajectories from longitudinal manifold-valued data. *Neural Information Processing Systems*, Dec 2017, Long Beach, United States. hal-01646617v1

HAL Id: hal-01646617

<https://hal.science/hal-01646617v1>

Submitted on 23 Nov 2017 (v1), last revised 12 Dec 2017 (v2)

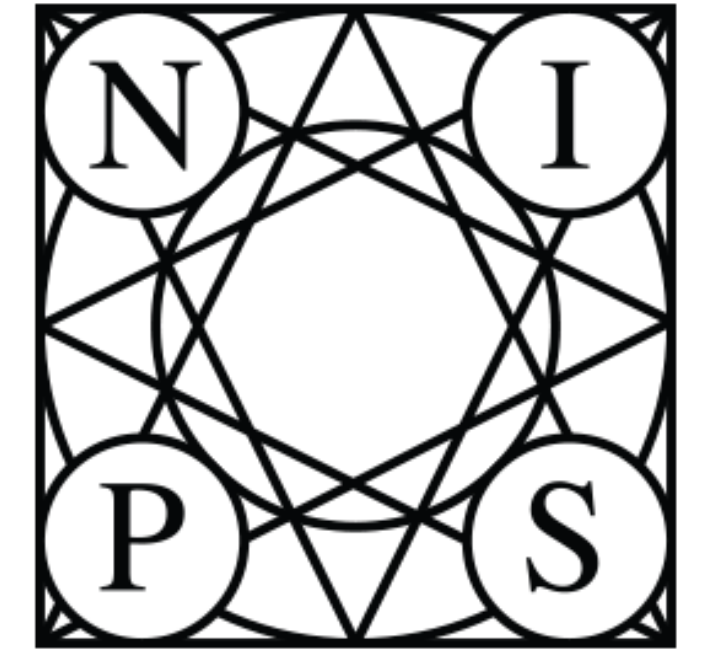
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Learning spatiotemporal piecewise-geodesic trajectories from longitudinal manifold-valued data

Juliette Chevallier, Stéphane Oudard and Stéphanie Allasonnière

CMAP, École polytechnique – Oncology département, HEGP – CRC, Université Paris Descartes



Overview

We introduce a **hierarchical model** which allows to estimate both a group-representative piecewise-geodesic trajectory in the **Riemannian** space of shape and inter-individual variability. Following the approach of [3], we estimate a **representative piecewise-geodesic trajectory** of the global progression and together with **spacial and temporal inter-individual variabilities**. We first introduce our model in its most generic formulation and then make it explicit for RECIST [4] score monitoring, *i.e.* for one-dimension manifolds and piecewise-logistically distributed data.

Medical context – RECIST score

- **New anti-angiogenic therapies.** Patients suffering from the metastatic kidney cancer take a drug each day [1];
- **Tumoral growth.** The RECIST score (Response Evaluation Criteria In Solid Tumors) is a set of published rules that measures the tumoral growth;
- **Patient's response.** The response to a given treatment has generally two distinct phases: first, tumor's size reduces; then, the tumor grows again;
- Moreover, a practical question is to quantify the correlation between both phases and to accurately determine when the patient's response escape to treatment.

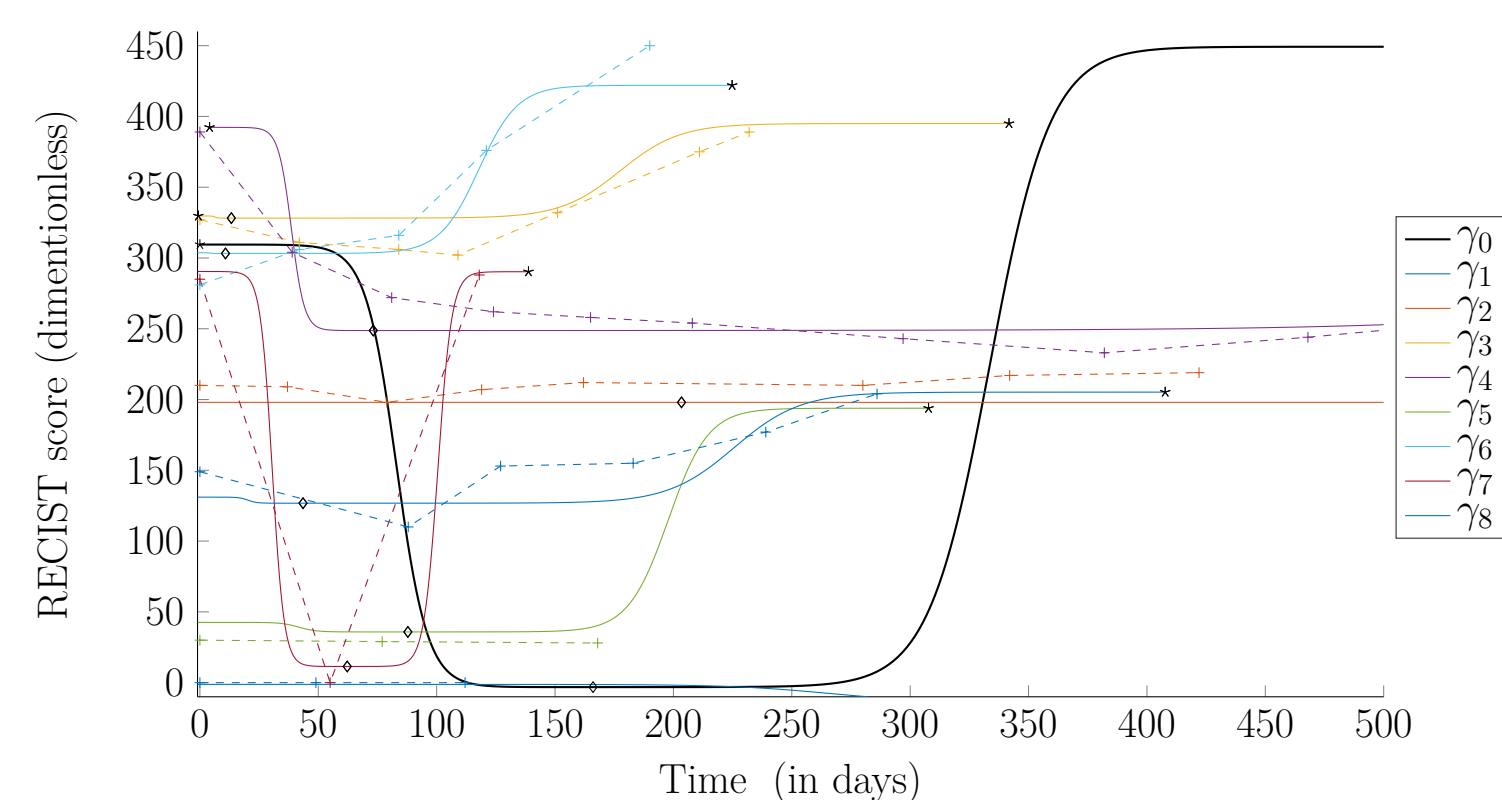


Figure 1: After 600 iterations. First 8 patients among the 176

Longitudinal dataset: Given $n \in \mathbb{N}^*$ individuals and $(k_i)_{i \in [1, n]}$, we observe the sequences $t = (t_{i,j})_{j \in [1, k_i]}^{i \in [1, n]} \in \mathbb{R}^k$ and $y = (y_{i,j})_{j \in [1, k_i]}^{i \in [1, n]} \in \mathbb{R}^{kd}$ where $k = \sum_{i=1}^n k_i$. Data points are seen as noisy samples along trajectories.

Real data consists of RECIST scores of a drove of 176 patients of the HEGP, with an average of 7 visits per subjects (min:3, max:22).

Mixed-effects model for piecewise-geodesically distributed data

Group-representative trajectory γ_0 . Given $m \in \mathbb{N}^*$ and $t_R = (-\infty < t_R^1 < \dots < t_R^{m-1} < +\infty)$, we build γ_0 in order it to be geodesic on each $]t_R^{\ell-1}, t_R^\ell]$.

- 1 Let $M_0 \subset \mathbb{R}^d$ a geodesically complete manifold, $(\bar{\gamma}_0^\ell)_{\ell \in [1, m]}$ a family of geodesics on M_0 and $(\phi_0^\ell)_{\ell \in [1, m]}$ a family of isometries defined on M_0 ;

- 2 $\forall \ell \in [1, m]$ we set $M_0^\ell = \phi_0^\ell(M_0)$ and $\gamma_0^\ell = \phi_0^\ell \circ \bar{\gamma}_0^\ell$;

- 3 A *piecewise-geodesic curve*. We define γ_0 as

$$\gamma_0 = \gamma_0^1 \mathbb{1}_{]-\infty, t_R^1]} + \sum_{\ell=2}^{m-1} \gamma_0^\ell \mathbb{1}_{]t_R^{\ell-1}, t_R^\ell]} + \gamma_0^m \mathbb{1}_{]t_R^{m-1}, +\infty[};$$

- 4 **Boundary conditions.** We impose boundary conditions on the rupture times to ensure continuity.

Individual trajectories $(\gamma_i)_{i \in [1, n]}$. Let $i \in [1, n]$. We build γ_i to derive from γ_0 through spatiotemporal transformations.

- 1 **Time warps $(\psi_i^\ell)_{\ell \in [1, m]}$.** We choose affine time warps constrains by the continuity of each individual paths;

- 2 **Space warps $(\phi_i^\ell)_{\ell \in [1, m]}$** have to be defined in view of applications. We require $\phi_i^\ell \circ \gamma_0^\ell(t_R^\ell) = \phi_i^{\ell+1} \circ \gamma_0^{\ell+1}(t_R^{\ell+1})$;

- 3 Last, $\forall \ell \in [1, m]$ we set $\gamma_i^\ell = \phi_i^\ell \circ \gamma_0^\ell \circ \psi_i^\ell$ and

$$\gamma_i = \gamma_i^1 \mathbb{1}_{]-\infty, t_{R,i}^1]} + \sum_{\ell=2}^{m-1} \gamma_i^\ell \mathbb{1}_{]t_{R,i}^{\ell-1}, t_{R,i}^\ell]} + \gamma_i^m \mathbb{1}_{]t_{R,i}^{m-1}, +\infty[};$$

- 4 **Gaussian noise.** $\forall j \in [1, k_i]$, $y_{i,j} = \gamma_i(t_{i,j}) + \varepsilon_{i,j}$ where $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma^2)$, $\sigma \in \mathbb{R}^+$.

Chemotherapy monitoring: Piecewise-logistic curve model

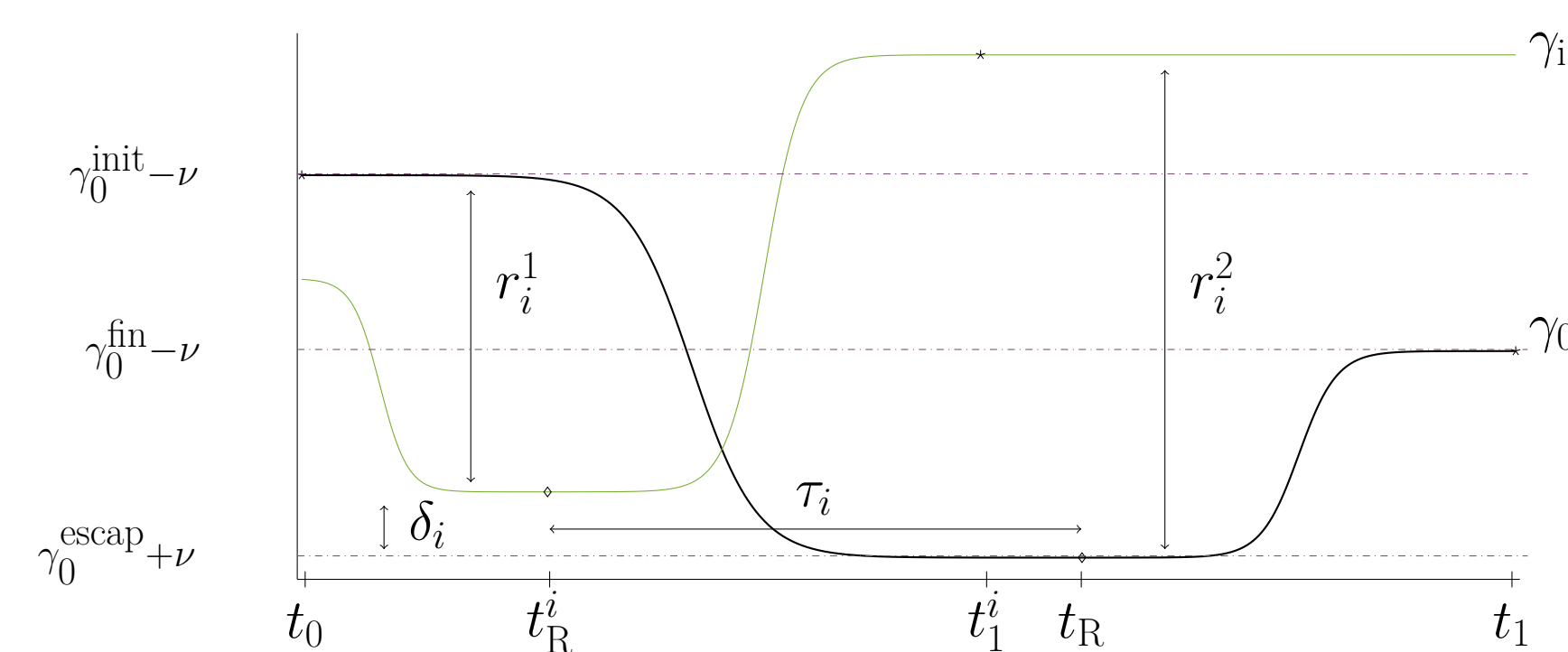


Figure 2: From average to individual path. Boundary conditions and transition from γ_0 to γ_i through spacial and temporal warps.

$$z_{\text{pop}} = (\gamma_0^{\text{init}}, \gamma_0^{\text{escap}}, \gamma_0^{\text{fin}}, t_R, t_1) \text{ and } z_i = (\xi_i^1, \xi_i^2, \tau_i, \rho_i^1, \rho_i^2, \delta_i).$$

Our observations consist of patient's RECIST score over time. So, we set $m = 2$, $d = 1$ and $M_0 =]0, 1[$ equipped with the logistic metric. Let $\nu \in \mathbb{R}$ and $i \in [1, n]$.

- **Representative trajectory γ_0 .** Let $\gamma_0^{\text{init}}, \gamma_0^{\text{escap}}, \gamma_0^{\text{fin}} \in \mathbb{R}$. We map M_0 onto $] \gamma_0^{\text{escap}}, \gamma_0^{\text{init}} [$ and $] \gamma_0^{\text{escap}}, \gamma_0^{\text{fin}} [$ through affine transformations and require that $\gamma_0^1(t_0) = \gamma_0^{\text{init}} - \nu$, $\gamma_0^1(t_R) = \gamma_0^2(t_R) = \gamma_0^{\text{escap}} + \nu$ and $\gamma_0^2(t_1) = \gamma_0^{\text{fin}} - \nu$;

- **Time warps (ψ_i^1, ϕ_i^2) .** We set $\alpha_i^\ell = e^{\xi_i^\ell}$, $\ell \in \{1, 2\}$;

- **Space warps (ϕ_i^1, ϕ_i^2) .** Given $(\rho_i^1, \rho_i^2, \delta_i) \in \mathbb{R}^3$, we set $\phi_i^\ell: x \mapsto e^{\rho_i^\ell} (x - \gamma_0(t_R)) + \gamma_0(t_R) + \delta_i$, $\ell \in \{1, 2\}$.

Parameters estimation with the MCMC-SAEM algorithm

Existence of the MAP

Given the piecewise-geodesic model and the choice of probability distributions for the parameters and latent variables of the model, for any dataset (t, y) , there exists $\hat{\theta}_{\text{MAP}} \in \arg\max_{\theta \in \Theta} q(\theta|y)$.

- **Parameters.** We assume that $z_{\text{pop}} \sim \mathcal{N}(\bar{z}_{\text{pop}}, \Sigma_{\text{pop}})$ where Σ_{pop} is a diagonal matrix with small fixed entries [2] and that $z_i \sim \mathcal{N}(0, \Sigma)$ where $\Sigma \in \mathcal{S}_6(\mathbb{R})$. Let $\theta = (\gamma_0^{\text{init}}, \gamma_0^{\text{escap}}, \gamma_0^{\text{fin}}, \bar{t}_R, \bar{t}_1, \Sigma, \sigma)$;

- **Hierarchical model.** Let $z = (z_{\text{pop}}, z_i)_{i \in [1, n]}$. We have

$$\begin{cases} y | z, \theta \sim \prod_{i=1}^n \prod_{j=1}^{k_i} \mathcal{N}(\gamma_i(t_{i,j}), \sigma^2) \\ z | \theta \sim \mathcal{N}(\bar{z}_{\text{pop}}, \Sigma_{\text{pop}}) \prod_{i=1}^n \mathcal{N}(0, \Sigma) \\ (\Sigma, \sigma) \sim \mathcal{W}^{-1}(V, m_\Sigma) \otimes \mathcal{W}^{-1}(v, m_\sigma) \end{cases}$$

where $V \in \mathcal{S}_6(\mathbb{R})$ and $v, m_\Sigma, m_\sigma \in \mathbb{R}$;

- **Estimation.** We use a Symmetric Random Walk Hastings-Metropolis sampler in a stochastic version of the EM algorithm.

Experimental results

Synthetic data. Experiments are performed for the piecewise-logistic model.

Table 1: Mean (standard deviation) of relative error (expressed as a percentage) for the population parameters \bar{z}_{pop} and the residual standard deviation σ for 50 runs according to the sample size n .

n	50	100	150
$\bar{\gamma}_0^{\text{init}}$	1.63 (1.46)	2.42 (1.50)	2.14 (1.17)
$\bar{\gamma}_0^{\text{escap}}$	9.45 (5.40)	9.07 (5.19)	11.40 (5.72)
$\bar{\gamma}_0^{\text{fin}}$	6.23 (2.25)	7.82 (2.43)	5.82 (2.55)
\bar{t}_R	11.58 (1.64)	13.62 (1.31)	9.24 (1.63)
\bar{t}_1	4.41 (0.75)	5.27 (0.60)	3.42 (0.71)
σ	25.24 (12.84)	10.35 (3.96)	2.83 (2.31)

Real data. Figure 1 illustrates the qualitative performance of the model on the first 8 patients.

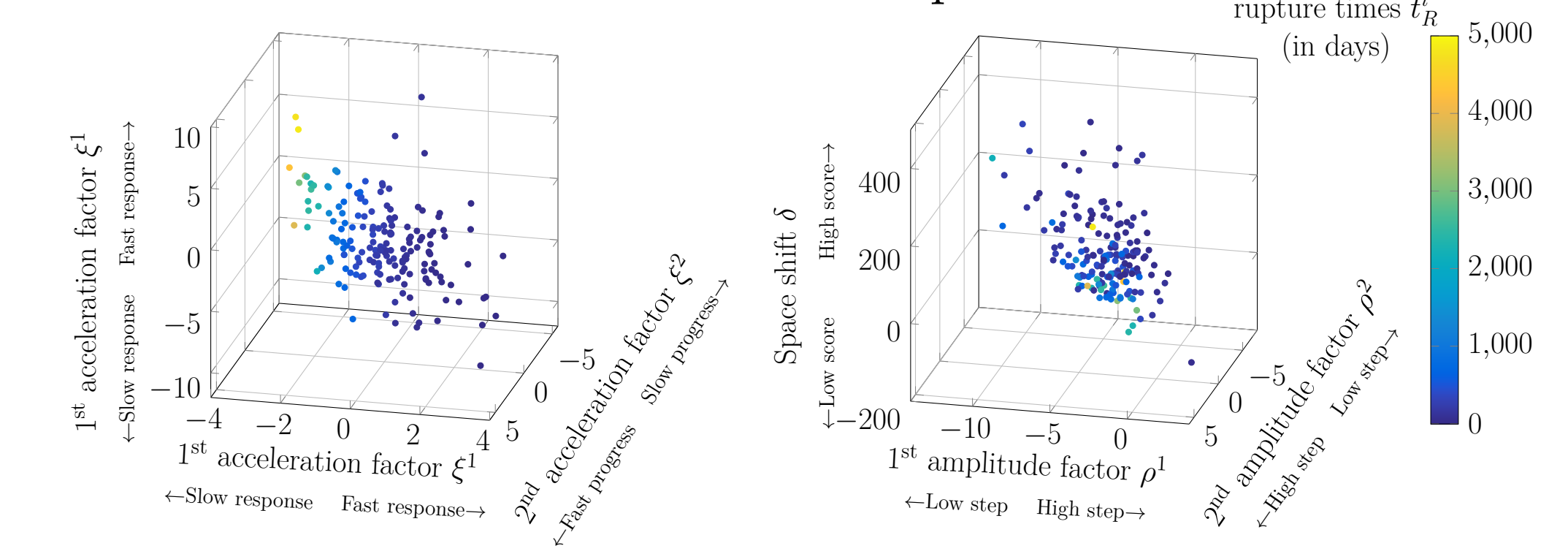


Figure 3: Individual random effects. Figure 3a: ξ_i^1 and ξ_i^2 against τ_i . Figure 3b: ρ_i^1 and ρ_i^2 against δ_i . In both figure, the color corresponds to the individual rupture time t_R^i .

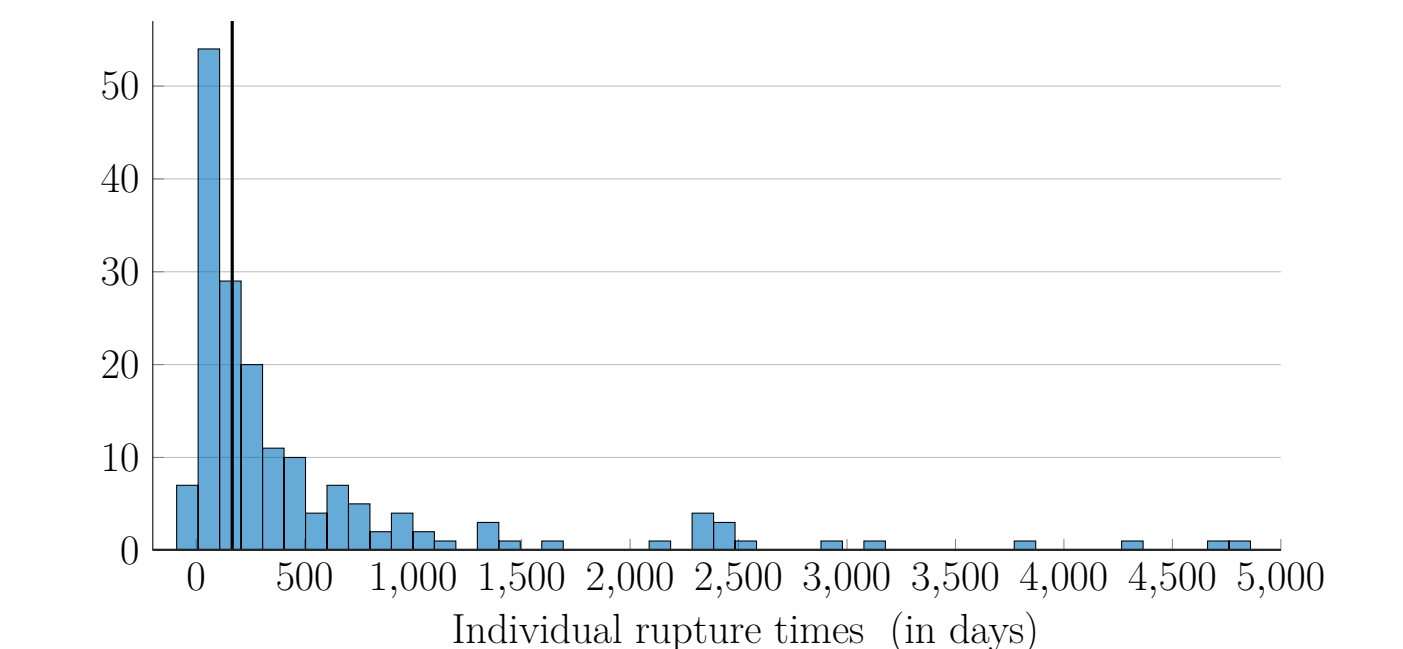


Figure 4: Distribution of the individual rupture times t_R^i . In black bold line, the estimated average rupture time t_R .

References

- [1] Escudier, Porta, Schmidinger, Rioux-Leclercq, Bex, Khoo, Gruenvald, and Horwich, *Renal cell carcinoma: Esmo clinical practice guidelines for diagnosis, treatment and follow-up*, Annals of Oncology **27** (2016), no. suppl 5, v58-v68.
- [2] Kuhn and Lavielle, *Maximum likelihood estimation in nonlinear mixed effects models*, Computational Statistics & Data Analysis **49** (2005), no. 4, 1020-1038.
- [3] Schiratti, Allasonnière, Colliot, and Durrleman, *Learning spatiotemporal trajectories from manifold-valued longitudinal data*, Neural Information Processing Systems 28, 2015.
- [4] Therasse, Arbuck, Eisenhauer, Wanders, Kaplan, Rubinstein, Verweij, Van Glabbeke, van Oosterom, Christian, and Gwyther, *New guidelines to evaluate the response to treatment in solid tumors*, Journal of the National Cancer Institute **92** (2000), no. 3, 205-216.

Acknowledgements

This work was supported by the public grant Investissement d'Avenir, project ANR-11-LBX-0056-LMH, and the Fondation of Medical Research, project DBI20131228564.