Credibility interval Design for Covid19 Reproduction Number from Nonsmooth Langevin-type Monte Carlo sampling

H. Artigas[1,†], B. Pascal[2], G. Fort[4,†], P. Abry[3,‡], N. Pustelnik[3,5]

† Partly funded by the *Fondation Simone et Cino Del Duca, Institut de France*, ‡ Partially supported by Grant 80PRIME-2021 CNRS

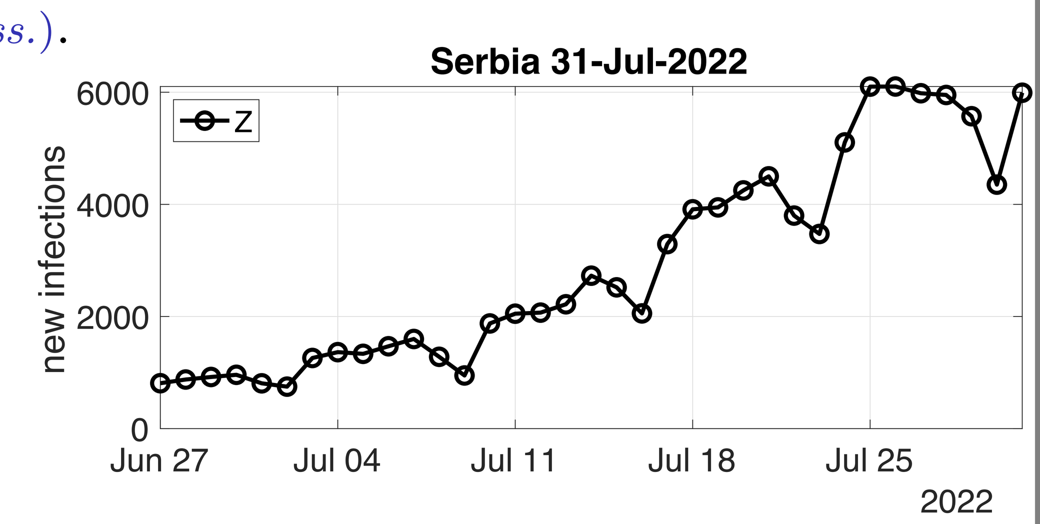
1. Ecole Polytechnique, Paris, France, 2. CNRS, Université de Lille, CRIStAL, France, 3. CNRS, ENS de Lyon, Laboratoire de Physique, France, 4. CNRS, Institut de Mathématique de Toulouse, France, 5. UC Louvain, Belgium

Aims and contributions

Monitoring the Covid19 pandemic is critical to design sanitary policies and to quantify their effectiveness. Reliable estimates of the pandemic *reproduction number* R_t were obtained from a nonsmooth convex optimization procedure designed to fit epidemiology requirements and to be robust to the low quality of the data (outliers, pseudo-seasonalities, ...) (Abry et al., 2020, *PlosOne* ; Pascal et al., 2022, *IEEE Trans. Sig. Process.*).

Purposes:
- grounded monitoring tools, based on sound epidemiological models,
- robust to low quality of the data (outliers data, abnormally small counts on week-ends, ...),
- accompanied by a confidence level, quantifying the uncertainty.

Methods:
- Bayesian paradigm leading to a nonsmooth posterior distribution,
- new Monte Carlo sampler combining Langevin dynamics and proximal operators,
- credibility interval estimate of the reproduction number R_t from Markov chains Monte Carlo samplers.



data from Johns Hopkins University

Bayesian model

(Cori et al., 2013, *Am. Journal of Epidemiology* ; Abry et al., 2020, *PlosOne* ; Pascal et al., 2022, *IEEE Trans. Sig. Process.* ; Fort et al., 2022, *preprint*)

Cori's model with outliers

T observations $\mathbf{Z} = (\mathsf{Z}_1, \ldots, \mathsf{Z}_T)^\top \in \mathbb{N}^T$, Z_t: new infections on day t

$$\mathbb{P}(\mathsf{Z}_t | \mathsf{Z}_1, \ldots, \mathsf{Z}_{t-1}) = \mathrm{Poiss}\left(p_t(\theta)\right) \text{ with } p_t(\theta) = \mathsf{R}_t \left(\Phi \mathsf{Z}\right)_t + \mathsf{O}_t, \quad (\Phi \mathsf{Z})_t = \sum_{u=1}^{\tau_\phi} \Phi_u \mathsf{Z}_{t-u},$$

parameterized by $\theta = (\mathsf{R}_1, \ldots, \mathsf{R}_T, \mathsf{O}_1, \ldots, \mathsf{O}_T) \in (\mathbb{R}_+)^T \times \mathbb{R}^T$

R_t: effective reproduction number ; O_t: models a possible outlier value

Φ: serial interval function $\Rightarrow$ random delay between primary and secondary cases

Negative log-likelihood from Poisson model (by convention $0 \ln 0 = 0$)

$$f(\theta) := \begin{cases} -\sum_{t=1}^T \left(\mathsf{Z}_t \ln p_t(\theta) - p_t(\theta)\right) & \text{if } \theta \in \mathcal{D} = \{\theta \,|\, \forall t,\ p_t(\theta) \geq 0\}, \\ +\infty & \text{otherwise} \end{cases}$$

Negative a priori log-distribution

- reproduction number: $\mathsf{R}_t - 2\mathsf{R}_{t-1} + \mathsf{R}_{t-2} \sim \mathrm{Laplace}(\lambda_{\mathrm{R}})$
- outliers $\mathsf{O}_t \sim \mathrm{Laplace}(\lambda_{\mathrm{O}})$

$$\Longrightarrow \quad g(\theta) = \lambda_{\mathrm{R}} \|\mathsf{D}_2 \mathbf{R}\|_1 + \lambda_{\mathrm{O}} \|\mathbf{O}\|_1,$$

$$\mathsf{D}_2 = \underbrace{\frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & -2 & 1 & 0 & \ldots & 0 \\ \ldots & & & & & & \ldots \\ 0 & \ldots & & & 1 & -2 & 1 \end{bmatrix}}_{\text{Laplacian}}$$

A posteriori distribution $\pi(\theta) \propto \exp\left(-f(\theta) - g(\theta)\right) \mathbb{1}_{\mathcal{D}}(\theta)$

f, g convex, $\quad f$ smooth, g nonsmooth

Monte Carlo sampler

Langevin dynamics

(Kent, 1978, *Adv Appl Probab*)

$\triangleright$ $\theta^{n+\frac{1}{2}} = \mu(\theta^n) + \sqrt{2\gamma} \xi^{n+1}, \quad \xi^{n+1} \sim \mathcal{N}_{2T}(0, \mathsf{I})$

$\mu(\theta)$ adapted to $\pi(\theta) = \exp(-f(\theta) - g(\theta)) \mathbb{1}_{\mathcal{D}}(\theta)$

$\triangleright$ θ^{n+1} via Acceptance-Rejection step

Main challenge:

$$-\ln \pi = f + g = f + h(\mathsf{A} \cdot)$$

closed-form expression of $\mathrm{prox}_{\gamma h}$ but not of $\mathrm{prox}_{\gamma h(\mathsf{A}\cdot)}$

Proximal-gradient dual strategy (Artigas et al., 2022, *EUSIPCO* ; Fort et al., 2022, *preprint*)

1) extend A into **invertible** $\overline{\mathsf{A}}$, and h in $\bar{h}$ such that $\bar{h}(\overline{\mathsf{A}}\theta) = h(\mathsf{A}\theta)$

2) work with the **dual** variable: $\underset{\theta \in \mathbb{R}^{2T}}{\mathrm{argmax}} \ln \pi(\theta) = \underset{\theta \in \mathbb{R}^{2T}}{\mathrm{argmin}}\ f(\theta) + \bar{h}(\overline{\mathsf{A}}\theta) = \mathsf{A}^{-1} \underset{\tilde{\theta} \in \mathbb{R}^{2T}}{\mathrm{argmin}}\ f(\overline{A}^{-1}\tilde{\theta}) + \bar{h}(\tilde{\theta})$

$$\boxed{\Longrightarrow \quad \mu(\theta) = \underbrace{\overline{\mathsf{A}}^{-1}}_{\text{back to } \theta} \underbrace{\mathrm{prox}_{\gamma \bar{h}} \left(\overline{\mathsf{A}}\theta - \gamma \overline{A}^{-\top} \nabla f(\theta)\right)}_{\text{proximal-gradient on } \tilde{\theta} = \overline{A}\theta}}$$

Two strategies to extend $\mathsf{A} = \begin{pmatrix} \mathsf{D}_2 & 0 \\ 0 & \mathsf{I} \end{pmatrix} \in \mathbb{R}^{(2T-1) \times 2T}$ into $\overline{\mathsf{A}} = \begin{pmatrix} \overline{\mathsf{D}} & 0 \\ 0 & \mathsf{I} \end{pmatrix} \in \mathbb{R}^{2T \times 2T}$:

`Invert`

$$\overline{\mathsf{D}}_2 := \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 & \cdots & 0 \\ & \mathsf{D}_2 & & & \end{bmatrix}$$

`Ortho`

$$\overline{\mathsf{D}}_o := \begin{bmatrix} v_1^\top \\ v_2^\top \\ \mathsf{D}_2 \end{bmatrix} \qquad \begin{array}{c} v_1, v_2 \in \mathbb{R}^{2T} \\ v_1 \perp v_2,\ v_1, v_2 \in (\mathsf{D}_2^\top)^\perp \end{array}$$

Metropolis-Hastings

Data: $\overline{\mathsf{D}} = \overline{\mathsf{D}}_2$ or $\overline{\mathsf{D}} = \overline{\mathsf{D}}_o$ $\gamma_{\mathrm{R}}, \gamma_{\mathrm{O}} > 0$, $N_{\max} \in \mathbb{N}_\star$, $\theta^0 = (\mathbf{R}^0, \mathbf{O}^0) \in \mathcal{D}$

Result: A $\mathcal{D}$-valued sequence $\{\theta^n = (\mathbf{R}^n, \mathbf{O}^n), n \in 0, \ldots, N_{\max}\}$

for $n = 0, \ldots, N_{\max} - 1$ **do**

Sample indep. $\xi_{\mathrm{R}}^{n+1}, \xi_{\mathrm{O}}^{n+1} \sim \mathcal{N}_T(0, \mathsf{I})$;

$\mathbf{R}^{n+\frac{1}{2}} = \mu_{\mathrm{R}}(\theta^n) + \sqrt{2\gamma_{\mathrm{R}}} \overline{\mathsf{D}}^{-1} \overline{\mathsf{D}}^{-\top} \xi_{\mathrm{R}}^{n+1}$

$\mathbf{O}^{n+\frac{1}{2}} = \mu_{\mathrm{O}}(\theta^n) + \sqrt{2\gamma_{\mathrm{O}}}\, \xi_{\mathrm{O}}^{n+1}$

$\theta^{n+\frac{1}{2}} = (\mathbf{R}^{n+\frac{1}{2}}, \mathbf{O}^{n+\frac{1}{2}})$

$\theta^{n+1} = \theta^{n+\frac{1}{2}}$ with probability

$$1 \wedge \frac{\pi(\theta^{n+\frac{1}{2}})}{\pi(\theta^n)} \frac{q_{\mathrm{R}}(\theta^{n+\frac{1}{2}}, \theta_{\mathrm{R}}^n)}{q_{\mathrm{R}}(\theta^n, \theta_{\mathrm{R}}^{n+\frac{1}{2}})} \frac{q_{\mathrm{O}}(\theta^{n+\frac{1}{2}}, \theta_{\mathrm{O}}^n)}{q_{\mathrm{O}}(\theta^n, \theta_{\mathrm{O}}^{n+\frac{1}{2}})},$$

$q_{\mathrm{R/O}}$: Gaussian kernels

and $\theta^{n+1} = \theta^n$ otherwise.

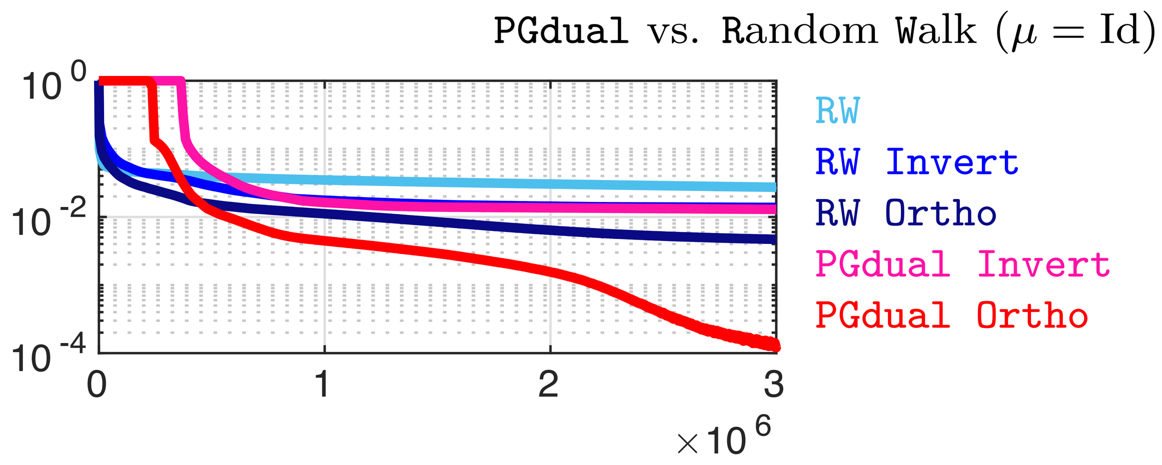
end

Algorithm 1: Proximal-Gradient dual: `PGdual Invert` and `PGdual Ortho`

Numerical experiment

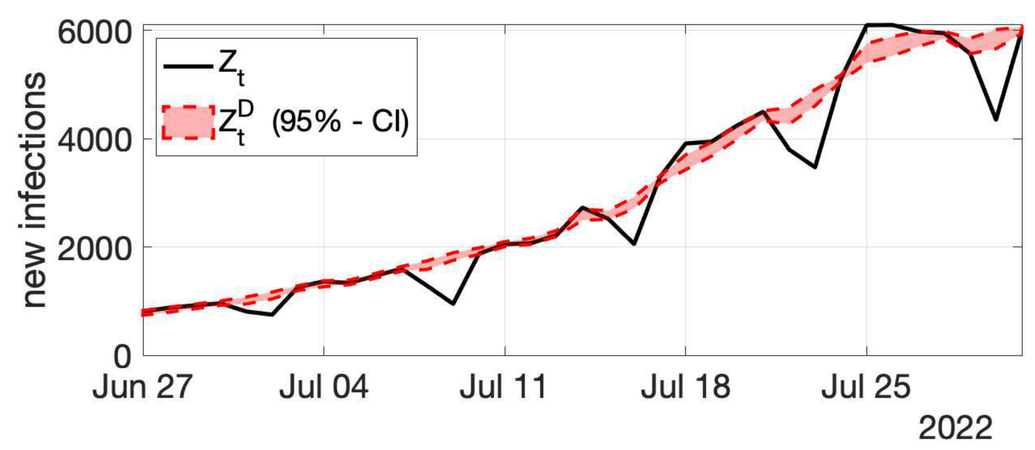
Convergence of the chain $\{\theta^n, n \in \mathbb{N}\}$

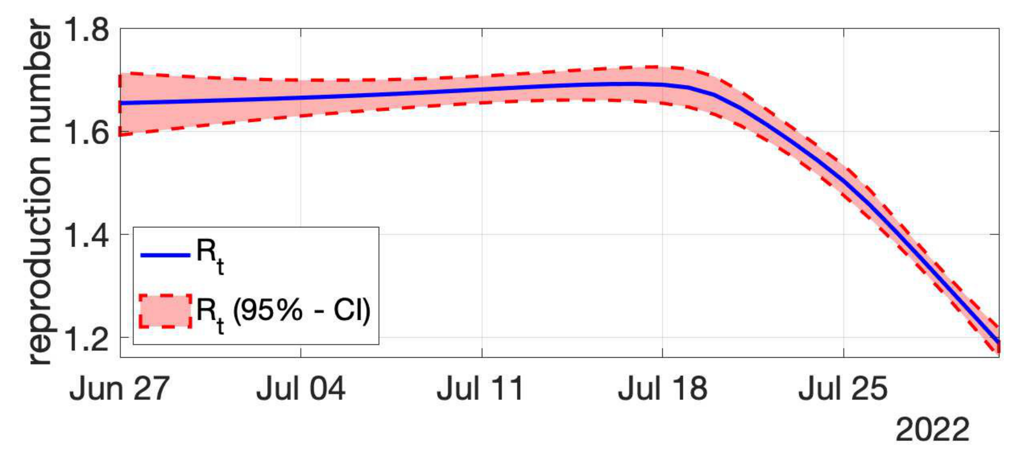
`PGdual` vs. Random Walk ($\mu = \mathrm{Id}$)



$(\ln \pi(\theta^n) - \max \ln \pi) \,/\, \left(\ln \pi(\theta^0) - \max \ln \pi\right)$

Denoised $\mathbf{Z}^{\mathrm{D}} = \mathbf{Z} - \widehat{\mathbf{O}}$ and $\widehat{\mathbf{R}}$





Perspectives

- hyperparameter tuning $(\gamma_{\mathrm{R/O}}, \lambda_{\mathrm{R}}, \lambda_{\mathrm{O}}, \ldots)$;
- handle missing counts: $\mathsf{Z}_{\mathrm{wknd}}$ not reported.