Some stochastic and statistical models for marine ecology

Adeline Leclercq Samson (Grenoble, France)

Joint work with

E. Buckwar (Linz), A. Delporte (Grenoble), S. Ditlevsen (Copenhagen),
M.P. Heide-Jørgensen (Greenland), A. Melnykova (Avignon),
P. Pilipovic (Copenhagen), P. Reynaud-Bouret (Nice),
M. Tamborrino (Warwick), O. Tervo (Copenhagen), I. Tubikanec (Klagenfurt)

LABORATOIRE JEAN KUNTZMANN
MATHÉMATIQUES APPLIQUÉES - INFORMATIQUE

UGA Université Grenoble Alpes

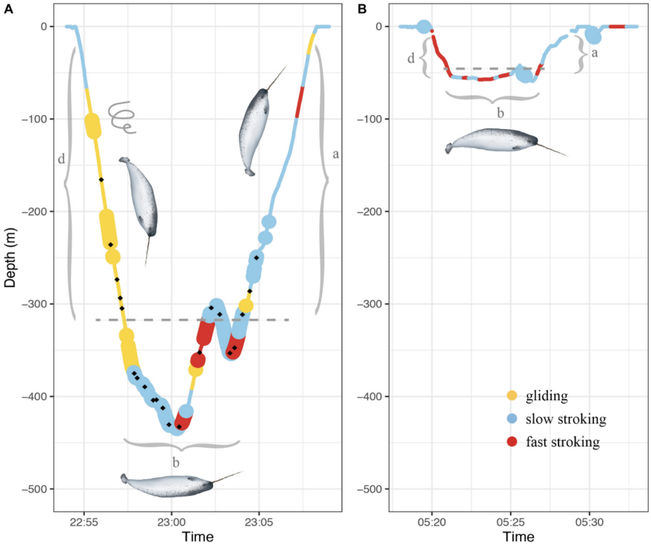
Outline of the talk

- **1. Climate change in Greenland**
  - Effect on endemic whales
  - Design of experimental data
- **2. Sound data**
  - Point process and
  - Memory components
  - Mixed effects and covariates
- **3. Spatial data**
  - Hypoelliptic diffusion
  - Numerical scheme
  - Statistical inference and tests

1. Narwhals in Greenland

- Endemic whales
  - Bowhead whale
  - Beluga whale
  - Narwhal
- Males
  - Long straight tusk - canine tooth
- Seasonal migration
  - Close to coast in summer
  - Offshore in winter

Narwhal dive



Foraging V dive (A) and Non-foraging U dive (B), Buzzes (black dots)

Climate change: a danger for narwhals?

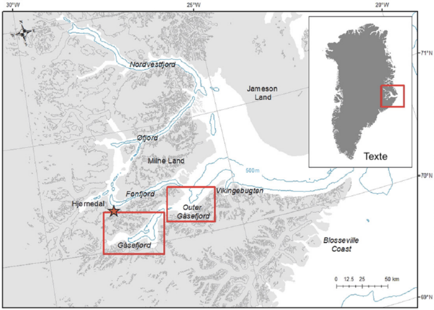
- Ice melt
- Decrease of sea ice coverage
- Increase of anthropogenic activities, of mining activities
- Reactions of narwhals to noise and human presence?

An overhead view of Hudson Resources' planned anorthosite mine at Qaqortorsuaq, Greenland. (Hudson Resources)

Students emerge from the KTI mine-training facility in Sisimiut, Greenland. The country is banking on the emergence of a mining sector that would generate income and jobs (KTI)

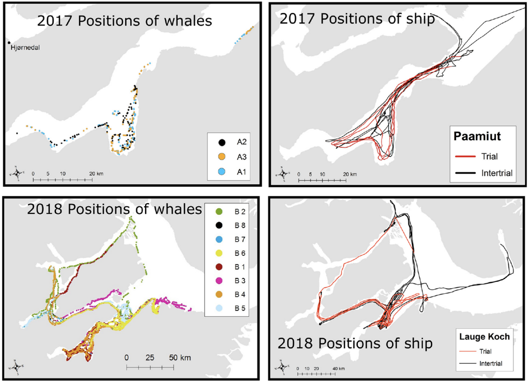
Experimental design

- 16 animals in 2017 and 2018
- Instrumented with satellite tags and Acousonde acoustic-behavioral recorders
- Exposition to airgun pulses and vessel sounds
  - Pre exposure
  - Trials: airgun and ship-noise
  - Intertrial: only ship-noise
  - Post exposure

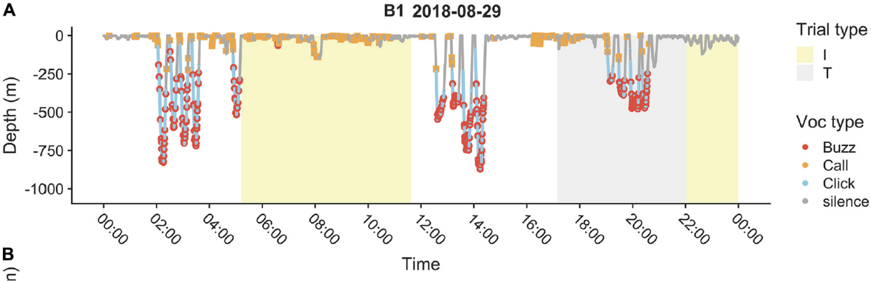


Recorded data

- Data recorded each second
- Sound (click/buzz)
- Depth
- Position of whales
- Position of ship
- Stroke
- Accelerometer



2. Sound data


A
B1 2018-08-29
Depth (m)
0
-250
-500
-750
-1000
00:00
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
18:00
20:00
22:00
00:00
Time
Trial type
I
T
Voc type
Buzz
Call
Click
silence
B

Sound analysis with point process

Point process

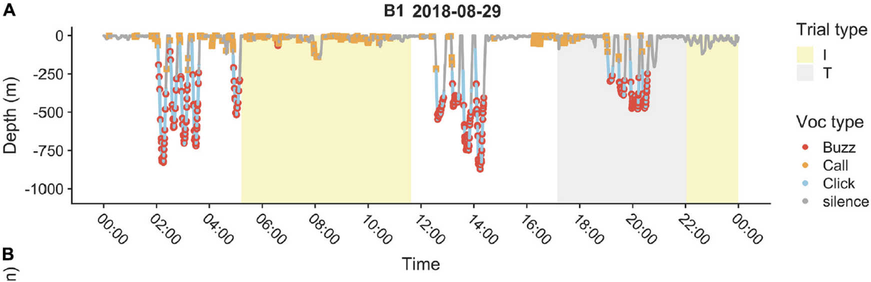
- T_j, time of the jth sound emitted by the animal
- $T_0 = 0 < T_1 < T_2 < \dots$
- $T_n \to \infty$ when $n \to \infty$ (no accumulation)
- $N(t)$ number of sounds emitted in the interval $]0; t]$
- Independence of variables $N(t_{j+1}) - N(t_j)$ for all $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$

Poisson process

- $\lambda(t)$ **rate** or **intensity** function
- For all $(t, s) \in \mathbb{R}_+^2$,

$$N(t+s) - N(t) \sim Poisson\left(\int_t^{t+s} \lambda(u) du\right)$$

Non independence of observations



- When buzzes start to be emitted, a whole sequence of buzzes is emitted
- Dependence of the emission between observations of a given animal

→ Intensity should depend on the past and thus be a random variable

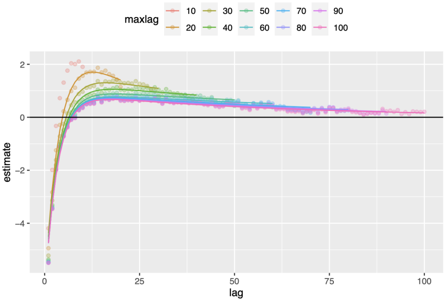
Point process with random intensity function

- $\mathcal{F}_t$ denotes the filtration of the process up to time t
- $\lambda(t)$ is a random variable whose value is determined by $\mathcal{F}_t$ such that
  - $\mathbb{P}(N(t+\Delta) - N(t) = 1 | \mathcal{F}_t) = \lambda(t)\Delta + o(\Delta)$
  - $\mathbb{P}(N(t+\Delta) - N(t) \geq 2 | \mathcal{F}_t) = o(\Delta)$
- Hawkes point process
  - Memory kernel ϕ
  - Random intensity function $\lambda(t)$

$$\lambda(t) = \lambda + \sum_{j=1}^{N(t)} \phi(t - T_j)$$

Estimation of the memory kernel ϕ

- From individual data before exposure
- Several lags p compared with BIC
- Parametric estimation with a sum of two exponential functions



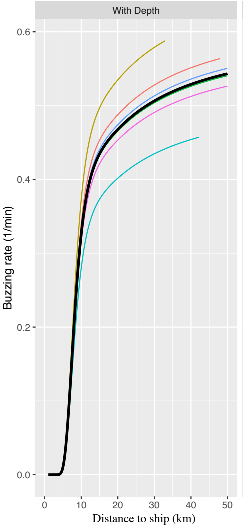
Estimation with exposure covariates and random effects:

$$\log \lambda_i(t) = \lambda + b_i + \sum_{j=1}^{N(t)} \left(\beta_1 e^{-\alpha_1 (t - T_{ij})} + \beta_2 e^{-\alpha_2 (t - T_{ij})} \right) + \gamma_1 D_i(t) + \gamma_2 E_i(t)$$

with

- $D(t)$: Depth
- $E(t)$: Exposure level of seismic activity = 1/distance to ship
- b_i random effect of individual i: $b_i \sim \mathcal{N}(0, \omega^2)$

Estimated buzzing rate



- Variability between individuals
- When ship is close, no buzz
- Buzzing rate without exposure $= 0.55/\text{min}$

Still to be done

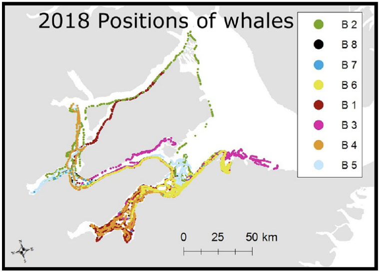
- Hawkes process
  - Non-parametric intensity and random effects
  - Mediation effect
    - Exposure causes animal to dive less
    - Less dives make animals produce less sounds

    - What is the direct of Exposure and its indirect effect through Depth?
    - Is Depth a mediation factor?

3. Spatial position data

Do animal adapt their position due to

- anthropogenic activities?
- warming of ocean temperatures?



Stochastic model of animal movement

[Blackwell 1997, Brillinger et al 2002, Preisler et al 2004, Hooten et al 2010, Michelot et al 2019, Gloaguen, Michelot, Etienne 2018]

Stochastic Differential Equation (SDE) to describe the dynamic of the position:

- X_t position at time t
- SDE on the position X_t

$$dX_t = b(X_t, \theta)dt + \Sigma(X_t, t)dW_t$$

- The drift $b(x, \theta)$ models the direction preference depending on position x
- The diffusion coefficient $\Sigma(X_t, t)$ models the variability around the mean

More complex SDE for movement ecology

Drift depend on spatial maps [Brillinger et al 2002, Gloaguen et al 2018]

$$dX_t = b(X_t, S_t, \theta)dt + \Sigma(X_t, t)dW_t$$

- The drift $b(x, S, \theta)$ models the direction preference depending on position x and spatial covariates S

$$b(x, S, \theta) = \sum_{k=1}^{L} \gamma_k H'(S_k(x), \theta)$$

  - H potential function, H' gradient of the potential
  - S_k different spatial maps (temperature, depth, ice coverage, etc)
  - γ_k weights of the mixture
- How should the noise Σ be chosen?

SDE on velocity $V_t \in \mathbb{R}^2$ [Hamiltonian models]

$$\begin{aligned} dX_t &= V_t dt \\ dV_t &= b(V_t, S_t, \theta)dt + \Sigma dW_t \end{aligned}$$

- Hypoelliptic system
- Only the position X_t is observed
- Nonlinear drift b
- Choice of Σ?

Parametric estimation for multidimensionnal systems

Matricial notations $Y_t \in \mathbb{R}^d$ with either $Y_t = X_t \in \mathbb{R}$ or $Y_t = (X_t, V_t) \in \mathbb{R}^4$:

$$dY_t = b(Y_t, \theta) + \Sigma dW_t$$

Difficult because

- Data: discrete observations $X_{10:n} = (X_{10}, \ldots, X_{1n})$ at times $t_0 = 0 < t_1 = \Delta < \ldots < t_n = n\Delta$
- No explicit transition density of the SDE except if b is linear
- Hypoellipticity: Σ is degenerated but Y_t has a smooth density (noise propagates to $\mathbb{R}^d$)

Different strategies of inference

Based on simulation

- Exact simulation
- Approximated simulation with a numerical scheme

- Estimation methods based on simulation
  - ABC
  - Monte Carlo, Importance Sampling

Based on approximation of the transition density

- What is a "good" approximation? Depends on the properties of the numerical scheme

- Estimation methods based on approximation
  - Contrast estimator [Thieullen Samson 2012, Ditlevsen Samson 2019, Melnykova 2020]
  - MCMC
  - EM and SAEM algorithm [Beskos et al 2005, Gloaguen et al 2018, Ditlevsen, Samson 2014, 2019]

Numerical approximation schemes $\tilde{Y}(t_i)$

Expected properties of a numerical scheme

- Locally Lipschitz conditions on b
- Exact moments up to a certain order
- Mean-square convergence of order p for a step size Δ

$$\max_{t_i} \left(\mathbb{E} \left(\| Y(t_i) - \tilde{Y}(t_i) \|^2 \right) \right)^{1/2} \leq c \Delta^p$$

- Preservation of structural properties: hypoellipticity, ergodicity, amplitudes, frequences, phases of oscillations

Some numerical schemes

Euler-Maruyama [Kloeden, Platen, 1992]

$$\tilde{Y}_{i+1} = \tilde{Y}_i + \Delta b(\tilde{Y}_i) + \sqrt{\Delta}\Sigma\eta_i, \quad \eta_i \sim_{iid} \mathcal{N}(0, I)$$

- Not mean-square convergent
- Does not preserve ergodicity
- Does not preserve hypoellipticity: if $\Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \Delta \end{pmatrix}$, then

$$Var(\tilde{Y}_{i+1}|\tilde{Y}_i) = \sigma^2 \begin{pmatrix} 0 & 0 \\ 0 & \Delta \end{pmatrix}$$

Local linearization [Ozaki 1989, Biscay et al. 1996, Jimenez et al 2015, Melnykova 2020]

- Approximate linear SDE on each interval $[i\Delta, (i+1)\Delta[$
- Mean square convergent but not ergodic

Numerical splitting scheme

[Buckwar, Samson, Tamborrino, Tubikanec, 2021; Pilipovic, Samson, Ditlevsen, 2022]

Introduction of two subsystems.

1. **Subsystem a:** Linear SDE with exact solution

$$dY_t^a = AY_t^a dt \; + \; \Sigma dW_t$$

2. **Subsystem b:** Non-linear (decoupled) ODE with (exact) solution

$$dY_t^b = N(Y_t^b)dt$$

Numerical splitting schemes with time step Δ

- Lie-Trotter

$$\hat{\mathbf{Y}}^{LT} = Y_\Delta^a \circ Y_\Delta^b$$

- Strang

$$\hat{\mathbf{Y}}^{S} = Y_{\Delta/2}^b \circ Y_\Delta^a \circ Y_{\Delta/2}^b$$

Properties of splitting schemes

- Exact first moment up to Δ^3, covariance matrix up to Δ^3
- Mean-square convergence with order 1
- Preservation of noise structure, 1-step hypoellitpticity
- Preservation of Lyapounov structure
- Geometric ergodicity
- Transition density of the scheme highly non-linear

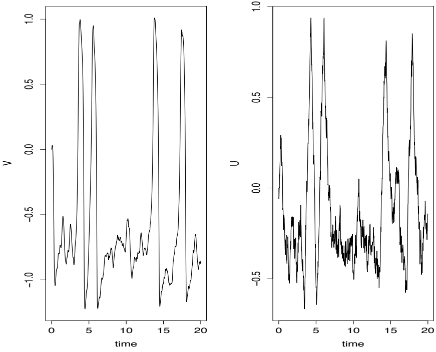
Example on the hypoelliptic Fitzhugh-Nagumo model

[Lindner et al 1999, Gerstner and Kistler, 2002, Lindner et al 2004, Berglund and Gentz, 2006]

$Y_t = X_t \in \mathbb{R}^2$

$$
\begin{aligned}
dX_{1t} &= \frac{1}{\varepsilon}(X_{1t} - X_{1t}^3 - X_{2t} - s)dt, \\
dX_{2t} &= (\gamma X_{1t} - X_{2t} + \beta)\, dt + \sigma dW_t,
\end{aligned}
$$

- ε time scale separation
- s stimulus input
- β position of the fixed point
- γ duration of excitation



Splitting scheme on the FHN model

1. **Subsystem a:** Linear SDE

$$dY_t = \begin{pmatrix} 0 & -\frac{1}{\varepsilon} \\ \gamma & -1 \end{pmatrix} Y_t dt \; + \; \Sigma dW_t$$

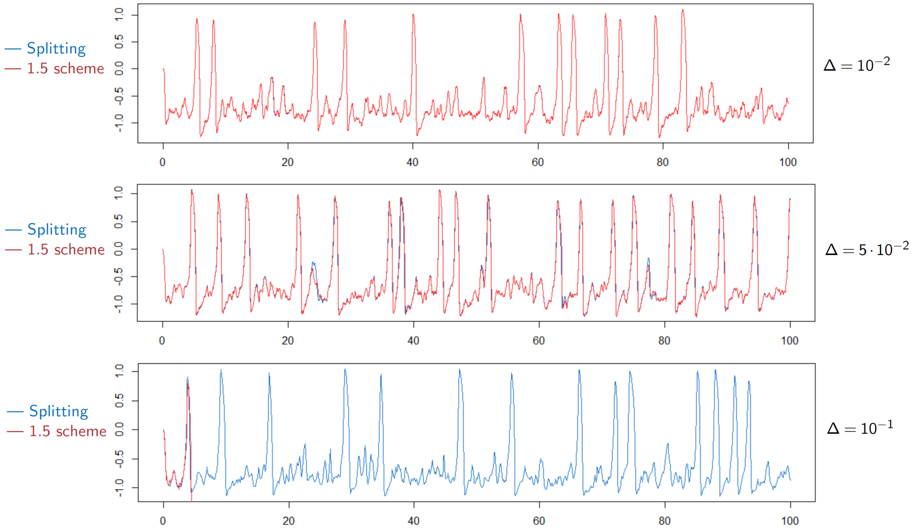
2. **Subsystem b:** Non-linear ODE

$$dY_t = \begin{pmatrix} \frac{1}{\varepsilon}(Y_{1t} - Y_{1t}^3) \\ \beta \end{pmatrix} dt$$

Explicit solution for both systems

Comparison of Splitting and Order 1.5 Strong Taylor Scheme

[Buckwar, Samson, Tamborrino, Tubikanec, 2021]



Estimation for elliptic SDE

$$dY_t = b_\theta(Y_t)dt + \Sigma dW_t, \quad \Sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

Minimum contrast estimator [Genon-Catolot, Jacod, 1993; Kessler 1996; Pilipovic, Samson, Ditlevsen, 2022]

Set $\Gamma = \Sigma'\Sigma$. Contrast function of the numerical scheme

$$(\hat{\theta}, \hat{\Gamma}) = \arg\min \left(\sum_{i=1}^{n-1} (Y_{i+1} - Y_i - \Delta b_\theta(Y_i))' \Gamma^{-1} (Y_{i+1} - Y_i - \Delta b_\theta(Y_i)) + \sum_{i=1}^{n-1} \log\det\Gamma \right)$$

Rate of convergence

- $\hat{\theta}$ asymptotically normal at rate $\sqrt{n\Delta}$
- $\hat{\Gamma}$ asymptotically normal at rate $\sqrt{n}$

Estimation for hypoelliptic SDE with partial observations

[Samson, Thieullen, 2012, Ditlevsen, Samson, 2019, Melnykova 2020, Pilipovic, Samson, Ditlevsen 2023]

Partial observations: V_t is not observed

- $dX_t = V_t dt$ gives $X_i - X_{i-1} = \int_{i-1}^{i} V_t dt \approx \Delta V_i$
- Approximation of V_i by $\tilde{V}_i = \frac{X_i - X_{i-1}}{\Delta}$

Constrats based on Euler-Maruyama or Splitting scheme

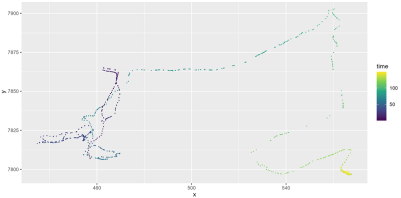
$$(\hat{\theta}, \hat{\Gamma}) = \arg\min \left(\sum_{i=1}^{n-1} \left(\tilde{V}_{i+1} - \tilde{V}_i - \Delta b_\theta(\tilde{V}_i) \right)' \Gamma^{-1} \left(\tilde{V}_{i+1} - \tilde{V}_i - \Delta b_\theta(\tilde{V}_i) \right) + \frac{2}{3} \sum_{i=1}^{n-1} \log \det \Gamma \right)$$

Rate of convergence specific to hypoelliptic systems

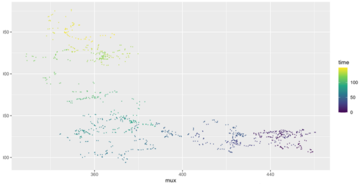
- $\hat{\theta}$ asymptotically normal at rate $\sqrt{n\Delta}$
- $\hat{\Gamma}$ asymptotically normal at rate $\sqrt{n}$

A simple SDE on narwhal GPS data

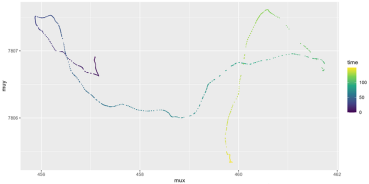
$$
\begin{aligned}
dX_t &= V_t dt \\
dV_t &= -\frac{1}{\tau)}(V_t - \mu)dt + \nu dW_t
\end{aligned}
$$



Asgeir's observed track



$\tau = 0.1h$, $\nu = 20km/h$

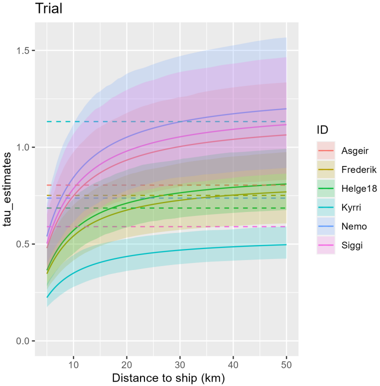


$\tau = 20h$, $\nu = 0.1km/h$

Exposure effects on narwhals movement

$$\begin{aligned} dX_t &= V_t dt \\ dV_t &= -\frac{1}{\tau(Z_t)}(V_t - \mu)dt + \nu dW_t \end{aligned}$$

with $\tau(Z_t) = \tau_0 + spline(Exposure_t)$



Still to be done

- First numerical applications of contrast estimator based on splitting scheme are promising
- Extension to SDE with spatial covariates
  - numerical schemes?
  - asymptotic results of estimators?
- Modeling Depth together with the position to get a 3D system
  - where should we put noise?
  - can we test the presence of noise?

Statistical test on the diffusion coefficient

[Melnykova, Reynaud-Bouret, Samson, work in progress]

$$dX_t = b(X_t)dt + \Sigma dW_t$$

- $X_t \in \mathbb{R}^d$
- Drift $b(X_t) = b(t)$
- Diffusion coefficient Σ diagonal: $\Sigma = diag(\sigma_1, \dots, \sigma_d)$

Statistical test

$$H_0 : \forall k = 1, \dots, d, \sigma_k^2 = \sigma_{k,0}^2$$
$$H_1 : \exists k = 1, \dots, d, \sigma_k^2 > \sigma_{k,0}^2$$

- Control of type I and II errors
- Non-asymptotic framework: n and Δ fixed

Test in dimension $d = 1$

$$dX_t = b_t dt + \sigma dW_t, \quad X_0 = x_0, \quad t > 0$$

Hypothesis

$$H_0 : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma^2 > \sigma_0^2$$

- Increments $\xi_i = \frac{X_{i\Delta} - X_{(i-1)\Delta} - \int_{(i-1)\Delta}^{i\Delta} b_s ds}{\sqrt{\Delta}} \sim \mathcal{N}(0, \sigma^2)$
- Test statistics $S = \frac{1}{n} \sum_{i=1}^{2} \xi_i^2 \sim \frac{\sigma^2}{n} \chi_n^2$

Control of the errors

- $\alpha \in]0;1[$ a fixed constant
- Test Υ which rejects H_0 if

$$S \geq \frac{\sigma_0^2}{n} q_{\chi^2_n, 1-\alpha}$$

- The test Υ is of Type I error α and therefore it is of level α

- Let $\beta \in]0;1[$ be a constant such that $1 - \beta \geq \alpha$
- For all σ^2 such that $\sigma^2 \geq \frac{q_{\chi^2_n, 1-\alpha}}{q_{\chi^2_n, \beta}} \sigma_0^2$ then

$$\mathbb{P}_{\sigma^2}(\Upsilon \text{ accepts } H_0) \leq \beta$$

Test in dimension $d = 2$

Hypothesis

$$H_0 : \det \Sigma\Sigma^T = \det \Sigma_0\Sigma_0^T$$
$$H_1 : \det \Sigma\Sigma^T > \det \Sigma_0\Sigma_0^T,$$

- Increments $\dot{\xi}_{ij} := \frac{X_{(2i+j-2)\Delta} - X_{(2i+j-3)\Delta} - \int_{(2i+j-3)\Delta}^{(2i+j-2)\Delta} b_s ds}{\sqrt{\Delta}} \quad j = 1, 2, \quad i = 1, \ldots, n/2$
- Test statistics $\dot{S} = \frac{1}{n/2} \sum_{i=1}^{n/2} \dot{s}_i$ with $\dot{s}_i = \det[(\dot{\xi}_{i1})^2, (\dot{\xi}_{i2})^2] = \dot{\xi}_{i11}^2 \dot{\xi}_{i22}^2 - \dot{\xi}_{i12}^2 \dot{\xi}_{i21}^2$

1. Sub-gaussian lower bound: $\mathbb{P}\left(\dot{S} - \mathbb{E}\left[\dot{S}\right] \leq -t\right) \leq \exp\left(-\frac{nt^2}{192\sigma_1^4\sigma_2^4}\right)$
2. Chebyshev's bound: $\mathbb{P}\left(\dot{S} - \mathbb{E}\left[\dot{S}\right] \geq t\right) \leq \frac{1}{n/2} \frac{20\sigma_1^4\sigma_2^4}{t^2}$

Control of the errors

- $\alpha \in]0;1[$ a fixed constant
- Test $\dot{\Upsilon}$ which rejects H_0 if

$$\dot{S} \geq 2 \det \Sigma_0 \Sigma_0^T \left(\sqrt{\frac{10}{n\alpha}} + 1 \right)$$

- The test $\dot{\Upsilon}$ is of Type I error α and therefore it is of level α

- Let $\beta \in]0;1[$ be a constant such that $1 - \beta \geq \alpha$.
- If $\det \Sigma \Sigma^T \geq \frac{\det \Sigma_0 \Sigma_0^T \left(\sqrt{\frac{10}{n\alpha}} + 1\right)}{1 - 4\sqrt{-\frac{3}{n} \log \beta}}$ then

$$\mathbb{P}_{\sigma^2} \left(\dot{\Upsilon} \text{ accepts } H_0 \right) \leq \beta$$

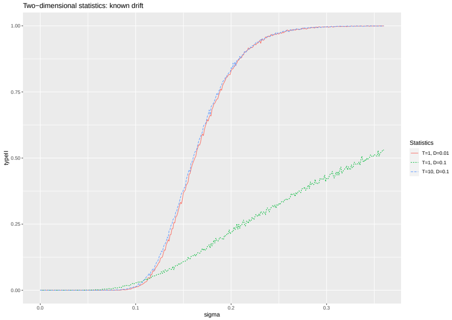
Test in dimension $d \geq 2$

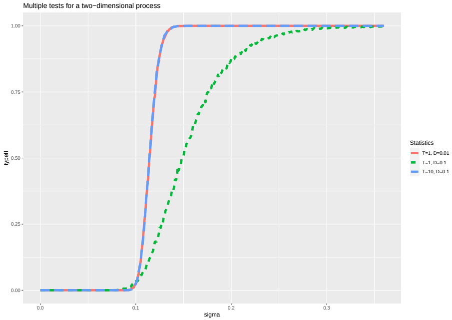
Hypothesis

$$H_0 : \forall k = 1, \dots, d, \sigma_k^2 = \sigma_{k,0}^2$$
$$H_1 : \exists k = 1, \dots, d, \sigma_k^2 > \sigma_{k,0}^2$$

- Test for each component $k = 1, \dots, d$ (1D test)
- Correction for multiplicity (for example Bonferroni)

Power of the tests on a 2 dimensional SDE





Perspectives

- Narwhals data are very rich
  - Time-dependent variables
  - Multi-dimensional analysis
  - Mediation effect
- Hypoelliptic SDE for movement/spatial data
  - Splitting schemes are promising
  - Adaptation to drift with gradient of potential
  - Modeling Depth and test of the noise

MERCI !