ESAIM: Probability and Statistics Will be set by the publisher
URL: `http://www.emath.fr/ps/`

SEMIPARAMETRIC REGRESSION ESTIMATION USING NOISY NONLINEAR NON INVERTIBLE FUNCTIONS OF THE SIGNAL

Élisabeth Gassiat[1] and Benoît Landelle[1, 2]

Abstract. We investigate a semiparametric regression model where one gets noisy non linear non invertible functions of the signal. We focus on the application to bearings-only tracking. We first investigate the least squares estimator and prove its consistency and asymptotic normality under mild assumptions. We study the semiparametric likelihood process and prove local asymptotic normality of the model. This allows to define the efficient Fisher information as a lower bound for the asymptotic variance of regular estimators and to prove that the maximum likelihood estimator is regular and asymptotically efficient. Simulations are presented to illustrate our results.

Résumé. On s'intéresse à un modèle de régression semi-paramétrique à partir de fonctions bruitées non-linéaires non-inversibles du signal. On détaille l'application à la trajectographie passive. On étudie d'abord l'estimateur des moindres carrés dont on prouve la consistance et la normalité asymptotique sous des hypothèses raisonnables. On s'intéresse ensuite à l'étude de la vraisemblance dans le cadre semi-paramétrique et on prouve la normalité asymptotique locale de ce modèle. Ceci permet de définir l'information de Fisher efficace comme borne inférieure de la variance asymptotique d'estimateurs réguliers et de prouver que l'estimateur du maximum de vraisemblance est régulier et asymptotiquement efficace. Des simulations illustrent ces résultats.

1991 Mathematics Subject Classification. 62J02, 62F12, 62F25, 62P30.

...

1. Introduction

In bearings-only tracking (BOT), one gets information about the trajectory of a target only via bearing measurements obtained by a moving observer. This is a highly ill-posed problem which requires, so that one be able to propose solutions, the choice of a trajectory model. The literature on the subject is very large, and many algorithms have been proposed to track the target, see for instance [3], [5], [2], [13]. All these algorithms are designed for particular classes of models for the trajectory of the target. In [8], it is proven that the least squares estimator may be very sensitive to some small deterministic perturbations, in which case the algorithms are highly non robust. However, it has been also claimed in [8] that stochastic perturbations do not essentially alter the performances of the estimator. The aim of this paper is to develop an estimation theory for a semiparametric

Keywords and phrases: Nonlinear regression, Semiparametric models, Bearings-only Tracking, Inverse models, Mixed Effects models

[1] Laboratoire de Mathématiques, Université Paris-Sud 11 and CNRS
[2] Thales Optronique, Élancourt (France)

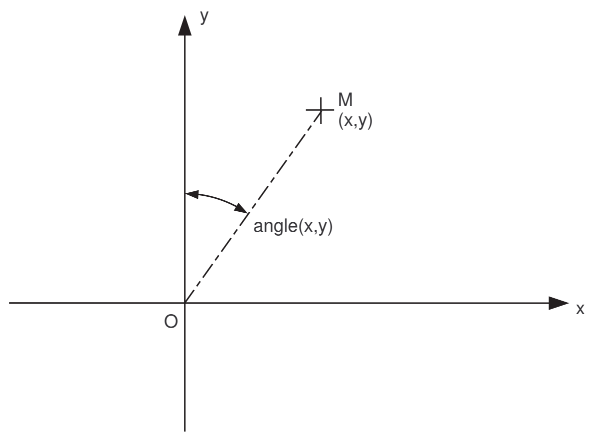

Figure 1. Angle function

model that applies to BOT. The model we study is the following :

$$\begin{cases} X_k = S_\theta(t_k) + \zeta_k \ , \\ Y_k = \Psi(X_k, t_k) + \varepsilon_k \ , \end{cases} \tag{1}$$

where $(t, \theta) \mapsto S_\theta(t)$ is a known map from $[0, 1] \times \Theta$ to $\mathbb{R}^s$ and Θ is the parameter set (in general, a subset of a finite dimensional euclidian space), the map $(x, t) \mapsto \Psi(x, t)$ is a known function from $\mathbb{R}^s \times [0, 1]$ to $\mathbb{R}$ and, in general, for a fixed t in [0,1], the map $x \mapsto \Psi(x, t)$ is non invertible, $\{t_k\}_{k\geq 0}$ is the known sequence of observation times in $[0, 1]$, $\{\zeta_k\}_{k\geq 0}$ is a sequence of random variables taking values in $\mathbb{R}^s$, $\{\varepsilon_k\}_{k\geq 0}$ is a sequence of centered i.i.d. random variables with known density g with respect to Lebesgue measure on $\mathbb{R}$, with known variance σ^2 and independent of the sequence $\{\zeta_k\}_{k\geq 0}$. The process $\{X_k\}_{k\geq 0}$ is referred to as the *signal* process and is not observed while $\{Y_k\}_{k\geq 0}$ is the *observation* process. We aim at estimating θ using only the observations $\{Y_k\}_{k\geq 0}$.

In case of BOT, the signal $\{X_k\}_{k\geq 0}$ is the trajectory of the target, given by its euclidian coordinates $(s = 2)$ at times $\{t_k\}_{k\geq 0}$, S_θ is the parametric trajectory the target is assumed to follow up to some parameter θ, for instance uniform linear motion, or a sequence of uniform linear and circular motions, $\{\zeta_k\}_{k\geq 0}$ is a noise sequence to take into account the fact that the model is only an idealization of the true trajectory and to allow stochastic departures of the trajectory model and $\{\varepsilon_k\}_{k\geq 0}$ is the observation noise. Since the observer is moving, if $O = \{O(t)\}_{t\in[0,1]}$ is its trajectory, the function $(x, t) \mapsto \Psi(x, t)$ is the angle, with respect to some fixed direction, of $x - O(t)$ that is, for $x = (x_1, x_2)$ different from $O(t)$,

$$\Psi(x, t) = \text{angle}[x_1 - O_1(t), x_2 - O_2(t)] \ , \tag{2}$$

where the angle function is defined, see figure 1, by

$$\text{angle}(x, y) \stackrel{\text{def}}{=} \begin{cases} \arctan(x/y) + \pi \times \text{sgn}(x) \cdot \mathbb{1}_{y<0} & \text{if } x \neq 0 \text{ and } y \neq 0, \\ \frac{\pi}{2} \times \text{sgn}(x) & \text{if } x \neq 0 \text{ and } y = 0, \\ \frac{\pi}{2} \times [1 - \text{sgn}(y)] & \text{if } x = 0 \text{ and } y \neq 0, \end{cases}$$

In such a case, for any z and fixed t, the set $\{x : \Psi(x, t) = z\}$ is infinite. Our aim here is to understand how it is possible to estimate the parameter θ in model (1), what are the limitations in the statistical performances, to propose estimation procedures, to build confidence regions for θ and to discuss their optimality under the weakest possible assumptions on the sequence $\{\zeta_k\}_{k\geq 0}$. Indeed, we would like to apply the results to BOT under realistic assumptions, for which it is not a strong assumption to assume that the observation noise $\{\varepsilon_k\}_{k\geq 0}$ consists of i.i.d. random variables with known distribution, but the trajectory noise $\{\zeta_k\}_{k\geq 0}$ may be quite complicated and unknown. To begin with, we will assume that the variables $\{\zeta_k\}_{k\geq 0}$ are i.i.d. with unknown distribution.

As such, the model may be viewed as a regression model with two variables, in which one of the variables is random, is not observed and follows itself a regression model. One could think that it looks like an inverse problem, or that the model may be understood as a state space model, or a mixed effects model, but in a nonstandard way, so that we have not been able to find results in the literature that apply to this setting. Throughout the paper, observations $\{Y_k\}_{k\geq 0}$ are assumed to follow model (1) with true (unknown) parameter θ^* and the observation times are $t_k = k/n$ with $k \in \{1,\dots,n\}$. All norms $\|\cdot\|$ are euclidian norms.

In Section 2, we consider least squares estimation and prove consistency and asymptotic normality in this setting, see Theorems 1 and 2. This allows to introduce basic considerations and set some assumptions. We prove that the results apply to BOT for linear observable trajectory models and when the trajectory noise has an isotropic distribution, see Theorem 3. Then, in Section 3 we study the likelihood process to set local asymptotic normality and efficiency in the parametric setting where the density of the noise $\{\zeta_k\}_{k\geq 0}$ is known, and define the efficient Fisher information in the semiparametric setting where the density of the noise $\{\zeta_k\}_{k\geq 0}$ is unknown. This also gives an estimation criterion which may be used even if the trajectory noise is correlated. In Section 4, we propose strategies for semiparametric estimation and discuss possible extension of the results to possibly dependent trajectory noise $\{\zeta_k\}_{k\geq 0}$. Section 5 is devoted to simulations. In each section, particular attention is given to the application of the results to BOT.

2. Least squares estimation

In sections 2 and 3, we will use :

Assumption 1. *$\{\zeta_k\}_{k\geq 0}$ is a sequence of i.i.d. random variables.*

To be able to obtain a consistent estimator of θ, we require that, in the absence of noise (both observation noise and trajectory noise), the observation at all times is sufficient to retrieve the parameter. This is the observability assumption :

Assumption 2. *If $\theta \in \Theta$ is such that $\Psi[S_\theta(t), t] = \Psi[S_{\theta^*}(t), t]$ a.e. for all $t \in [0,1]$, then $\theta = \theta^*$.*

If the observation noise is centered, in the absence of trajectory noise, the fact that only $\Psi[S_\theta(t), t]$ is observed with additive noise is not an obstacle to the estimation of θ under Assumption 2. But with trajectory noise, only the distribution of $\Psi[S_\theta(t) + \zeta_1, t]$ may be retrieved from noisy data. In case the marginal distribution of the ζ_k's is known, this may be enough, but in case it is unknown, one has to be aware of some link between the distribution of $\Psi[S_\theta(t) + \zeta_1, t]$ and θ. We thus introduce the following assumption, which will be proved to hold in some BOT situations.

Assumption 3. *For all t in $[0,1]$, let us assume that*

$$\mathbb{E}\{\Psi[S_{\theta^*}(t) + \zeta_1, t]\} = \Psi[S_{\theta^*}(t), t] \ .$$

Let us now define the least squares criterion and the least squares estimator (LSE) by

$$M_n(\theta) \overset{\text{def}}{=} n^{-1} \sum_{k=1}^{n} \{Y_k - \Psi[S_\theta(t_k), t_k]\}^2 \ , \qquad \bar{\theta}_n \overset{\text{def}}{=} \arg\min_{\theta\in\Theta} M_n(\theta),$$

where $\arg\min_{\theta\in\Theta} M_n(\theta)$ is any minimizer of M_n.

2.1. **Consistency**

We assume that Θ is a compact subset of $\mathbb{R}^d$ and we will use :

Assumption 4. *The maps $t \mapsto \mathbb{E}\left\{\Psi[S_{\theta^*}(t)+\zeta_1, t]^2\right\}$ and $(t,\theta) \mapsto \Psi[S_\theta(t), t]$ define finite continuous functions on respectively $[0,1]$ and $[0,1]\times\Theta$. Moreover,*

$$\lim_{M\to\infty} \sup_{t\in[0,1]} \mathbb{E}\left(\Psi[S_{\theta^*}(t)+\zeta_1,t]^2 \mathbb{1}\left\{\Psi[S_{\theta^*}(t)+\zeta_1,t]^2 > M\right\}\right) = 0 \ .$$

Theorem 1. *Under assumptions 1 , 2, 3 and 4, $\bar{\theta}_n$ converges in probability to θ^* as n tends to infinity.*

The proof is a consequence of general results in M-estimation. We begin with a simple Lemma:

Lemma 1. *Under Assumption 1, if F is a real function on $\mathbb{R}^s \times [0,1]$ such that*

$$\sup_{t\in[0,1]} \mathbb{E}|F(\zeta_1,t)| < \infty \ , \qquad \lim_{M\to\infty} \sup_{t\in[0,1]} \mathbb{E}\{|F(\zeta_1,t)|\mathbb{1}_{|F(\zeta_1,t)|>M}\} = 0 \ ,$$

and $\mathbb{E}[F(\zeta_1,\cdot)]$ is Riemann-integrable, then

$$n^{-1}\sum_{k=1}^n F(\zeta_k,t_k) = \int_0^1 \mathbb{E}[F(\zeta_1,t)]\,\mathrm{d}t + \mathrm{o}_{\mathbb{P}}(1) \ .$$

Proof. First of all, by the integrability assumption,

$$\lim_{n\to\infty} n^{-1}\sum_{k=1}^n \mathbb{E}[F(\zeta_k,t_k)] = \int_0^1 \mathbb{E}[F\,(\zeta_1,t)]\,\mathrm{d}t \ .$$

Then

$$\begin{aligned} n^{-1}\sum_{k=1}^n &\{F(\zeta_k,t_k) - \mathbb{E}[F(\zeta_k,t_k)]\} \\ &= n^{-1}\sum_{k=1}^n \left\{F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|>M} - \mathbb{E}\left[F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|>M}\right]\right\} \\ &+ n^{-1}\sum_{k=1}^n \left\{F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|\le M} - \mathbb{E}\left[F(\zeta_k,t_k)\mathbb{1}_{|F(\zeta_k,t_k)|\le M}\right]\right\} \ . \end{aligned}$$

The variance of the second term is upper bounded by $2M^2/n$ so that the second term tends to 0 in probability as n tends to infinity and the absolute value of the first term has expectation upper bounded by $2\sup_{t\in[0,1]}\mathbb{E}\left[|F(\zeta_1,t)|\mathbb{1}_{|F(\zeta_1,t)|>M}\right]$, which may be made smaller than any positive ϵ for big enough M, which proves the lemma. □

Proof of Theorem 1. Let us define

$$M(\theta) \stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left\{\Psi[S_{\theta^*}(t)+\zeta_1,t] - \Psi[S_\theta(t),t]\right\}^2 \,\mathrm{d}t + \sigma^2 \ .$$

Direct calculations yield

$$
\begin{aligned}
&M(\theta) - M(\theta^*)\\
&\quad = \int_0^1 \mathbb{E}\left(\{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_\theta(t), t]\}^2 - \{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_{\theta^*}(t), t]\}^2 \right) \mathrm{d}t \;,\\
&\quad = \int_0^1 \{\Psi[S_{\theta^*}(t), t] - \Psi[S_\theta(t), t]\} \times \{2\mathbb{E}\left(\Psi[S_{\theta^*}(t) + \zeta_1, t]\right) - \Psi[S_{\theta^*}(t), t] - \Psi[S_\theta(t), t]\} \;\mathrm{d}t.
\end{aligned}
$$

By Assumption 3, it follows that

$$
M(\theta) - M(\theta^*) = \int_0^1 \{\Psi[S_\theta(t), t] - \Psi[S_{\theta^*}(t), t]\}^2 \;\mathrm{d}t \;,
$$

so that $M(\theta)$ has a unique minimum at θ^* by Assumption 2. Also, under Assumption 4, the map $\theta \mapsto M(\theta)$ is uniformly continuous from Θ to $\mathbb{R}$. Now, for any $\theta \in \Theta$,

$$
\begin{aligned}
M_n(\theta) = n^{-1} \sum_{k=1}^n \varepsilon_k^2 + 2n^{-1} \sum_{k=1}^n \varepsilon_k \{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_\theta(t_k), t_k]\}&\\
+ n^{-1} \sum_{k=1}^n \{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_\theta(t_k), t_k]\}^2 \;.& \qquad (3)
\end{aligned}
$$

It follows, by law of large numbers, that $n^{-1}\sum_{k=1}^n \varepsilon_k^2 = \sigma^2 + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1)$. The variance of the second expression in the right-hand side of (3) is given by

$$
\mathbb{V}\left(n^{-1} \sum_{k=1}^n \varepsilon_k \{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_\theta(t_k), t_k]\} \right) = \sigma^2 n^{-2} \sum_{k=1}^n \mathbb{E}\{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_\theta(t_k), t_k]\}^2 \;,
$$

and converges to 0, so that, by Tchebychev's inequality,

$$
2n^{-1} \sum_{k=1}^n \varepsilon_k \{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_\theta(t_k), t_k]\} = \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \;.
$$

Applying Lemma 1 yields

$$
n^{-1} \sum_{k=1}^n \{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_\theta(t_k), t_k]\}^2 = \int_0^1 \mathbb{E}\{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_\theta(t), t]\}^2 \;\mathrm{d}t + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \;.
$$

Thus for any $\theta \in \Theta$, $M_n(\theta)$ converges in probability to $M(\theta)$. Using the compacity of Θ and the second part of Assumption 4, it is possible to strengthen this pointwise convergence to a uniform one:

$$
\sup_{\theta \in \Theta} |M_n(\theta) - M(\theta)| = \mathrm{o}_{\mathbb{P}_{\theta^*}}(1). \qquad (4)
$$

Indeed, for any θ and θ' in Θ,

$$M_n(\theta) - M_n(\theta') = 2n^{-1} \sum_{k=1}^{n} \varepsilon_k \left\{ \Psi[S_{\theta'}(t_k), t_k] - \Psi[S_\theta(t_k), t_k] \right\} \\ + n^{-1} \sum_{k=1}^{n} \left\{ 2\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_\theta(t_k), t_k] - \Psi[S_{\theta'}(t_k), t_k] \right\} \left\{ \Psi[S_{\theta'}(t_k), t_k] - \Psi[S_\theta(t_k), t_k] \right\} ,$$

so that, for any $\delta > 0$,

$$\sup_{\substack{(\theta,\theta')\in\Theta^2 \\ \|\theta-\theta'\|\le\delta}} |M_n(\theta) - M_n(\theta')| \le \omega(\delta)\, n^{-1} \sum_{k=1}^{n} \left\{ 2|\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k]| + 2 \sup_{(\theta,t)\in\Theta\times[0,1]} |\Psi[S_\theta(t), t]| + 2|\varepsilon_k| \right\} ,$$

where $\omega(\cdot)$ is the uniform modulus of continuity of $(t,\theta) \mapsto \Psi[S_\theta(t), t]$. The right-hand side of the inequality converges in probability by Lemma 1 to a constant times $\omega(\delta)$, so that equation (4) follows from compacity of Θ. Theorem 1 now follows from [14, Theorem 5.7]. $\square$

2.2. Asymptotic normality

Asymptotic normality of the least squares estimator will follow using usual arguments under further regularity assumptions.

Assumption 5. *There exists a neighborhood U of θ^* such that for all $t \in [0,1]$, the map $\theta \mapsto \Psi[S_\theta(t), t]$ possesses two derivatives on U that are continuous as functions of (θ, t) over $U \times [0,1]$.*

Let us define, for θ in U,

$$\begin{aligned} I_R(\theta) &\overset{\text{def}}{=} \int_0^1 \nabla_\theta \Psi[S_\theta(t), t] \nabla_\theta \Psi[S_\theta(t), t]^{\mathrm{T}} \, \mathrm{d}t , \\ I_\Psi(\theta) &\overset{\text{def}}{=} \int_0^1 \mathbb{E}\left(\left\{ \Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_\theta(t), t] \right\}^2 \right) \nabla_\theta \Psi[S_\theta(t), t] \nabla_\theta \Psi[S_\theta(t), t]^{\mathrm{T}} \, \mathrm{d}t . \end{aligned}$$

Then :

Theorem 2. *Under Assumptions 1 , 2, 3, 4 and 5, if $I_R(\theta^*)$ is non singular,*

$$\sqrt{n}\left(\bar{\theta}_n - \theta^*\right) = I_R^{-1}(\theta^*) n^{-1/2} \sum_{k=1}^{n} \left\{ \Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] - \Psi[S_{\theta^*}(t_k), t_k] + \varepsilon_k \right\} \nabla_\theta \Psi[S_{\theta^*}(t_k), t_k] + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) .$$

In particular, $\sqrt{n}\left(\bar{\theta}_n - \theta^*\right)$ converges in distribution, as $n \to \infty$, to $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$ where

$$I_M^{-1}(\theta^*) = I_R^{-1}(\theta^*) \left[I_\Psi(\theta^*) + \sigma^2 I_R(\theta^*) \right] I_R^{-1}(\theta^*) .$$

Let us notice that, for a null sequence $\{\zeta_k\}_{k\ge0}$, we retrieve the usual Fisher information matrix for the parametric regression model.

Proof. The proof follows Wald's arguments. On the set $\left\{\bar{\theta}_n \in U\right\}$ which can be assumed to be convex and which has probability tending to 1 according to Theorem 1,

$$\nabla_\theta M_n(\bar{\theta}_n) = 0 = \nabla_\theta M_n(\theta^*) + \int_0^1 \nabla_\theta^2 M_n[\theta^* + s(\bar{\theta}_n - \theta^*)] \, \mathrm{d}s \, (\bar{\theta}_n - \theta^*) .$$

Direct calculations yield, for any $\theta \in U$,

$$\nabla_\theta M_n(\theta) = -2n^{-1} \sum_{k=1}^{n} \left\{ \Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k - \Psi[S_\theta(t_k), t_k] \right\} \nabla_\theta \Psi[S_\theta(t_k), t_k] \;,$$

and

$$\nabla^2_\theta M_n(\theta) = 2n^{-1} \sum_{k=1}^{n} \nabla_\theta \Psi[S_\theta(t_k), t_k] \nabla_\theta \Psi[S_\theta(t_k), t_k]^{\mathrm{T}} \\ - 2n^{-1} \sum_{k=1}^{n} \left\{ \Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k - \Psi[S_\theta(t_k), t_k] \right\} \nabla^2_\theta \Psi[S_\theta(t_k), t_k] \;. \quad (5)$$

Notice that, using Assumption 3, $\nabla_\theta M_n(\theta^*)$ is a centered random variable, and that, using Assumptions 4, 5, the variance of $\nabla_\theta M_n(\theta^*)$ converges to $4\left[I_\Psi(\theta^*) + \sigma^2 I_R(\theta^*)\right]$ as $n \to \infty$. Also using Assumptions 3, 4, 5, and applying Lemma 1, $\nabla^2_\theta M_n(\theta)$ converges in probability to $2I_R(\theta)$ as $n \to \infty$. Using Assumption 5, there exists an increasing function ω satisfying $\lim_{\delta \to 0} \omega(\delta) = 0$ such that, for all θ and θ' in U with $\|\theta - \theta'\| \le \delta$,

$$\left\| \nabla^2_\theta M_n(\theta) - \nabla^2_\theta M_n(\theta') \right\| \le \omega(\delta) \times n^{-1} \sum_{k=1}^{n} \left\{ |\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k| + 2 \right\} \;.$$

It follows that, on the set $\left\{\bar\theta_n \in U\right\}$, for all s in $[0,1]$,

$$\left\| \nabla^2_\theta M_n[\theta^* + s(\bar\theta_n - \theta^*)] - \nabla^2_\theta M_n(\theta^*) \right\| \le \omega\left(\|\bar\theta_n - \theta^*\|\right) \times n^{-1} \sum_{k=1}^{n} \left\{ |\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k| + 2 \right\} \;.$$

By Lemma 1, $n^{-1}\sum_{k=1}^{n} |\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k| = \mathrm{O}_{\mathbb{P}_{\theta^*}}(1)$ so that, using the consistency of $\bar\theta_n$, Lemma 1 and Assumption 5,

$$\int_0^1 \nabla^2_\theta M_n[\theta^* + s(\bar\theta_n - \theta^*)] \,\mathrm{d}s = 2I_R(\theta^*) + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \;.$$

Finally, we obtain

$$[I_R(\theta^*) + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1)] \sqrt{n}\left(\bar\theta_n - \theta^*\right) = n^{-1/2} \sum_{k=1}^{n} \{\Psi[S_{\theta^*}(t_k) + \zeta_k, t_k] + \varepsilon_k - \Psi[S_{\theta^*}(t_k), t_k]\} \nabla_\theta \Psi[S_{\theta^*}(t_k), t_k] + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) \;.$$

Using Assumption 5, the convergence in distribution to $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$ is a consequence of the Lindeberg-Feller Theorem and Slutzky's Lemma. □

Notice that, if $\hat I_M$ is a consistent estimator of $I_M(\theta^*)$, by Slutsky's Lemma, $n^{1/2} \hat I_M^{1/2}\left(\bar\theta_n - \theta^*\right)$ converges in distribution to the centered standard gaussian distribution in $\mathbb{R}^d$, which allows to build confidence regions with asymptotic known level. If the distribution of the trajectory noise $\{\zeta_k\}_{k \ge 0}$ is known, one may use $\hat I_M = I_M\left(\bar\theta_n\right)$. If the distribution of the noise is unknown, one could use bootstrap procedures to build confidence regions based on the empirical distribution of $\bar\theta_n$ using bootstrap replicates.

Another possibility occurs if one has a majoration

$$\mathbb{E}\left(\left\{ \Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_{\theta^*}(t), t] \right\}^2 \right) \le A^2, \quad (6)$$

where A denotes a known constant. Indeed, in such a case, $I_\Psi(\theta^*)$ is upper bounded (in the natural ordering of positive symetric matrices) by $A^2 I_R(\theta^*)$, so that $I_M^{-1}(\theta^*)$ is upper bounded by $(A^2 + \sigma^2) I_R^{-1}(\theta^*)$, and one may use $(A^2 + \sigma^2) I_R^{-1}\left(\bar\theta_n\right)$ as variance matrix to obtain conservative confidence regions.

2.3. **Application to BOT**

To apply the results to BOT, one has to see whether Assumptions 1, 2, 3, 4 and 5 hold and if $I_R(\theta^*)$ is non singular. If $y > 0$, the angle measure is given by

$$\text{angle}(x, y) = \arctan(x/y) \ . \tag{7}$$

To ensure that the angle measure is given by (7), we will use the following hypothesis :

Assumption 6. *For all θ in Θ and t in $[0, 1]$,*

$$S_{\theta,2}(t) - O_2(t) > 0 \ .$$

This hypothesis is usually made in BOT litterature. Under assumption 6, the observation is given by

$$\Psi(x, t) = \arctan[x_1 - O_1(t)]/[x_2 - O_2(t)] \ . \tag{8}$$

The bearing exact measurements of the non noisy possible trajectory stay inside an interval with length π. This may be seen as an assumption on the manoeuvres of the observer. Assumption 2 is the observability assumption which holds for models such as uniform linear motion if the observer does not move itself along uniform linear motion, or a sequence of uniform linear and circular motions, if the observer does not move along uniform linear motion or circular motion in the same time intervals as the target. Various observability properties are proved in [7]. Assumptions 4 and 5 hold as soon as the trajectory model $S_\theta(t)$ is twice differentiable for all t in $[0, 1]$ as a function of θ and the denominator in (8) may not be 0.The fact that $I_R(\theta^*)$ is non singular is equivalent to the observability assumptions for linear models. Let us introduce such models.

Let $(e_1, \ldots, e_p)$ be a family of continuous functions on $[0, 1]$, $\theta = (a_1, \ldots, a_p, b_1, \ldots, b_p)$ in Θ a subset of $\mathbb{R}^{2p}$ and

$$S_\theta = (S_{\theta,1}, S_{\theta,2}) = \left(\sum_{k=1}^{p} a_k e_k, \sum_{k=1}^{p} b_k e_k \right) \ . \tag{9}$$

Then

Proposition 1. *Under model (9) satisfying Assumption 6, Assumption 2 holds if and only if $I_R(\theta^*)$ is non singular.*

Proof. Let $\theta^* = (a_1^*, \ldots, a_p^*, b_1^*, \ldots, b_p^*)$. For t in $[0, 1]$, let

$$m(\theta, t) = [S_{\theta,1}(t) - O_1(t)]/[S_{\theta,2}(t) - O_2(t)] \ .$$

Thanks to Assumption 6, simple algebra gives that $\Psi[S_\theta(t), t] = \Psi[S_\theta^*(t), t]$ if and only if

$$\sum_{k=1}^{p} (a_k - a_k^*) e_k(t) - \sum_{k=1}^{p} (b_k - b_k^*) e_k(t) m(\theta^*, t) = 0 \ ,$$

so that Assumption 2 holds if and only if the family of functions $\big(e_1, \ldots, e_p, e_1 \times m(\theta^*, \cdot), \ldots, e_p \times m(\theta^*, \cdot)\big)$ is linearly independent in the space of continuous functions on $[0, 1]$. Also, for i in $\{1, \ldots, p\}$ and t in $[0, 1]$, it follows by direct calculations

$$\frac{\partial}{\partial a_i} \arctan m(\theta^*, t) = - \left\{ [1 + m(\theta^*, t)^2][S_{\theta,1}(t) - O_1(t)] \right\}^{-1} e_i(t) m(\theta^*, t) \ ,$$

and

$$\frac{\partial}{\partial b_i} \arctan m(\theta^*, t) = \left\{ [1 + m(\theta^*, t)^2][S_{\theta,1}(t) - O_1(t)] \right\}^{-1} e_i(t) \ ,$$

so that $I_R(\theta^*)$ is non singular if and only if the family of functions $\big(e_1, \dots, e_p, e_1 \times m(\theta^*, \cdot), \dots, e_p \times m(\theta^*, \cdot)\big)$ is linearly independent in the space of continuous functions on $[0, 1]$, which ends the proof. $\square$

Thus under model (9), if the trajectory of the observer is such that Assumptions 2 and Assumption 6 hold, then Assumptions 4 and 5 hold and $I_R(\theta^*)$ is non singular.

What remains to be seen is whether Assumption 3 holds and it is the case under a simple assumption on the distribution of the trajectory noise :

Assumption 7. *ζ_1 has an isotropic distribution in $\mathbb{R}^2$ and a density with compact support such that, for all $t \in [0, 1]$,*

$$S_{\theta^*,2}(t) + \zeta_{1,2} - O_2(t) > 0 \; .$$

We introduce some prior knowledge on the trajectory and on the variance of the trajectory noise to be able to obtain conservative confidence regions.

Assumption 8. *The trajectory $t \mapsto S_{\theta^*}(t)$ is such that, for all $t \in [0, 1]$,*

$$\|S_{\theta^*}(t) - O(t)\| \geq R_{\min} \; ,$$

and a constant number A^2 such that

$$\pi^2 \left(1 + \pi^{-2/3}\right)^3 \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{R_{\min}^2} \leq A^2$$

is known.

This condition makes sense since in the context of passive tracking one usually assumes that the distance between target and observer is quite large. Let us denote by C_α a region with coverage $1 - \alpha$ for the standard gaussian distribution in $\mathbb{R}^d$.

Theorem 3. *If the trajectory model $(t, \theta) \mapsto S_\theta(t)$ and the move of the observer are such that Assumptions 2, 4, 5 and 8 hold and $I_R(\theta^*)$ is non singular, or if the trajectory model is (9) satisfying Assumption 6 and Assumption 2 holds, if moreover Assumption 1 and 7 hold, then, for any $\alpha > 0$,*

$$\liminf_{n \to \infty} \mathbb{P}_{\theta^*} \left[\sqrt{n}(A^2 + \sigma^2)^{-1/2} I_R^{1/2}\left(\bar{\theta}_n\right)\left(\bar{\theta}_n - \theta^*\right) \in C_\alpha\right] \geq 1 - \alpha \; .$$

Proof. Under Assumption 7, let the density of ζ_1 be $F(\|\zeta_1\|)$. Recall that the trajectory of the observer is $\{O(t)\}_{t \in [0,1]}$. Let $\beta(t) = \Psi[S^*_\theta(t), t] = \arctan[S_{\theta^*,1}(t) - O_1(t)]/[S_{\theta^*,2}(t) - O_2(t)]$. It follows by trigonometric considerations, see figure 2, that

$$\begin{aligned}
&\mathbb{E}\{\Psi[S_{\theta^*}(t) + \zeta_1, t]\} \\
&\qquad = \iint_{\mathbb{R} \times (-\pi, \pi)} \arctan\left\{\frac{S_{\theta^*,1}(t) - O_1(t) + r \sin\alpha}{S_{\theta^*,2}(t) - O_2(t) + r\cos\alpha}\right\} F(r)\, r \,\mathrm{d}r\, \mathrm{d}\alpha \; , \\
&\qquad = \beta(t) + \iint_{\mathbb{R} \times (-\pi, \pi)} \arctan\left\{\frac{r \sin[\alpha - \beta(t)]}{\|S_{\theta^*}(t) - O(t)\| + r \cos[\alpha - \beta(t)]}\right\} F(r)\, r \,\mathrm{d}r\, \mathrm{d}\alpha \; .
\end{aligned}$$

Let

$$G_{\theta^*,t}(r, \alpha) = \arctan\left\{\frac{r\sin\alpha}{\|S_{\theta^*}(t) - O(t)\| + r\cos\alpha}\right\} \; .$$

Then,

$$\mathbb{E}\{\Psi[S_{\theta^*}(t) + \zeta_1, t]\} = \Psi[S_{\theta^*}(t), t] + \iint_{\mathbb{R} \times (-\pi, \pi)} G_{\theta^*,t}(r, \alpha) F(r)\, r \,\mathrm{d}r\, \mathrm{d}\alpha \; .$$

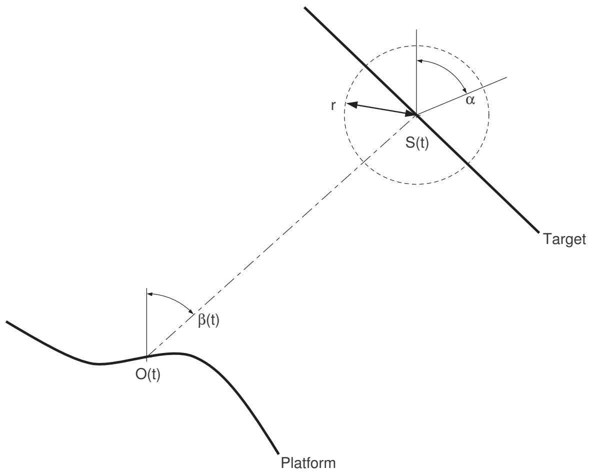

FIGURE 2. Isotropic noise

But for any $r > 0$, for any $\alpha \in (-\pi, \pi)$, $G_{\theta^*,t}(r, -\alpha) = G_{\theta^*,t}(r, \alpha)$ so that

$$\mathbb{E}\{\Psi[S_{\theta^*}(t) + \zeta_1, t]\} = \Psi[S_{\theta^*}(t), t] \ .$$

Now,

$$\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_{\theta^*}(t), t] = \int_0^1 \nabla_x \Psi[S_{\theta^*}(t) + h\zeta_1, t]^{\mathrm{T}} \zeta_1 \, \mathrm{d}h \ ,$$

and direct calculations provide $\|\nabla_x \Psi[x, t]\| = \|x - O(t)\|^{-1}$. Thus, for any $a \in]0, 1[$,

$$\begin{aligned}
&\mathbb{E}\Big(\{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_{\theta^*}(t), t]\}^2\Big) \\
&\qquad \le \pi^2 \mathbb{P}\left(\|\zeta_1\| \ge a\|\Psi[S_{\theta^*}(t), t] - O(t)\|\right) + \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{(1-a)^2 \|\Psi[S_{\theta^*}(t) - O(t)\|^2} \ , \\
&\qquad \le \pi^2 \mathbb{P}\left(\|\zeta_1\| \ge a R_{\min}\right) + \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{(1-a)^2 R_{\min}^2} \ ,
\end{aligned}$$

since $|\Psi(u) - \Psi(v)| \le \pi$ for any real numbers u and v and by using the triangular inequality and Assumption 8. But Tchebychev's inequality leads to

$$\mathbb{E}\left(\{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_{\theta^*}(t), t]\}^2\right) \le \frac{\mathbb{E}\left(\|\zeta_1\|^2\right)}{R_{\min}^2} \left(\frac{\pi^2}{a^2} + \frac{1}{(1-a)^2}\right) \ , \tag{10}$$

which is minimum for $a = \frac{1}{1+\pi^{-2/3}}$ leading to $\left(\frac{\pi^2}{a^2} + \frac{1}{(1-a)^2}\right) = \pi^2\left(1 + \pi^{-2/3}\right)^3$ and

$$\mathbb{E}\left(\{\Psi[S_{\theta^*}(t) + \zeta_1, t] - \Psi[S_{\theta^*}(t), t]\}^2\right) \le A^2 \ .$$

To conclude one may apply the concluding remark of Section 2.2 to obtain asymptotic conservative confidence regions for θ. $\square$

3. Likelihood and efficiency

Let $\mathcal{F}$ be the set of probability densities f on $\mathbb{R}^s$ such that for all t in $[0,1]$, for all θ in Θ,

$$\int_{\mathbb{R}^s} \Psi[S_\theta(t)+\zeta,t]f(\zeta)\,\mathrm{d}\zeta = \Psi[S_\theta(t),t]\ . \tag{11}$$

We will replace Assumptions 1 and 3 by :

Assumption 9. *$\{\zeta_k\}_{k\geq 0}$ is a sequence of i.i.d. random variables with density $f^*\in\mathcal{F}$.*

The normalized log-likelihood is the function on $\Theta\times\mathcal{F}$ given by

$$J_n(\theta,f) \stackrel{\text{def}}{=} n^{-1}\sum_{k=1}^{n}\log\left(\int_{\mathbb{R}^s} g\left\{Y_k-\Psi[S_\theta(t_k)+u,t_k]\right\}\,f(u)\,\mathrm{d}u\right). \tag{12}$$

Define

$$G\left[(\zeta,\varepsilon),t;\theta\right] \stackrel{\text{def}}{=} \log\left(\int_{\mathbb{R}^s} g\left\{\Psi[S_{\theta^*}(t)+\zeta,t]+\varepsilon-\Psi[S_\theta(t)+u,t]\right\}\,f(u)\,\mathrm{d}u\right)\ .$$

As soon as for any (θ,f) in $\Theta\times\mathcal{F}$, it is possible to apply Lemma 1 to $G\left[(\cdot),\cdot;\theta\right]$ with the sequence $\{\zeta_k,\varepsilon_k\}_{k\geq 0}$, $J_n(\theta,f)$ converges in probability to

$$J(\theta,f)=\int_0^1\int_{\mathbb{R}^s}\int_{\mathbb{R}}\log\left(\int_{\mathbb{R}^s} g\left\{\Psi[S_{\theta^*}(t)+\zeta,t]+\varepsilon-\Psi[S_\theta(t)+u,t]\right\}\,f(u)\,\mathrm{d}u\right)\,g(\varepsilon)f^*(\zeta)\,\mathrm{d}\varepsilon\,\mathrm{d}\zeta\,\mathrm{d}t. \tag{13}$$

Let

$$p_{(\theta,f)}\left(z,t\right) \stackrel{\text{def}}{=} \int_{\mathbb{R}^s} g\left\{z-\Psi[S_\theta(t)+u,t]\right\}\,f(u)\,\mathrm{d}u$$

be the density, for fixed t, of the random variable

$$Z \stackrel{\text{def}}{=} \Psi[S_\theta(t)+U,t]+V\ ,$$

where U is a random variable in $\mathbb{R}^s$ with density f independent of the real-valued random variable V with density g. Thus, $p_{(\theta^*,f^*)}\left(\cdot,t_k\right)$ is the probability density of Y_k. Then, the change of variable $z=\Psi[S_{\theta^*}(t)+\zeta,t]+\varepsilon$ in

$$\int_{\mathbb{R}}\log\left(\int_{\mathbb{R}^s} g\left\{\Psi[S_{\theta^*}(t)+\zeta,t]+\varepsilon-\Psi[S_\theta(t)+u,t]\right\}\,f(u)\,\mathrm{d}u\right)\,g(\varepsilon)\,\mathrm{d}\varepsilon$$

leads to

$$J(\theta,f)=\iint_{\mathbb{R}\times(0,1)} p_{(\theta^*,f^*)}\left(z,t\right)\log p_{(\theta,f)}\left(z,t\right)\,\mathrm{d}z\,\mathrm{d}t\ .$$

Thus, for any (θ,f) in $\Theta\times\mathcal{F}$,

$$J(\theta^*,f^*)\geq J(\theta,f),$$

and $J(\theta^*,f^*)=J(\theta,f)$ if and only if $p_{(\theta,f)}\left(z,t\right)=p_{(\theta^*,f^*)}\left(z,t\right)\quad(t,z)$-a.e., that is the probability distribution of $Z=\Psi[S_\theta(t)+U,t]+V$ is the same as that of $\Psi[S_{\theta*}(t)+U^*,t]+V$, where U^* is a random variable in $\mathbb{R}^s$ with density f^* independent of the random variable V. But if $f\in\mathcal{F}$ and $f^*\in\mathcal{F}$, taking expectations leads to the fact that, t-a.e., $\Psi[S_\theta(t),t]=\Psi[S_{\theta*}(t),t]$, so that $\theta=\theta^*$ if Assumption 2 holds. In other words, $J(\theta,f)$ is maximum only for $\theta=\theta^*$.

Following the same lines as for the LSE, we may thus easily obtain that, if the probability density f^* is known, the parametric maximum likelihood estimator is consistent and asymptotically gaussian. Define the parametric maximum likelihood estimator as

$$\tilde{\theta}_n \stackrel{\text{def}}{=} \arg\max_{\theta\in\Theta} J_n(\theta, f^*) ,$$

where $\arg\max_{\theta\in\Theta} J_n(\theta, f^*)$ is any maximizer of $J_n(\cdot, f^*)$. If for any θ in Θ, there exists a small open ball containing θ such that Lemma 1 applies to $\sup_{\theta\in U} G[(\cdot), \cdot; \theta]$, it is possible, as in [14, Theorem 5.14], to strengthen the convergence of $J_n(\theta, f^*)$ to $J(\theta, f^*)$ in a uniforme one. The consistency of $\tilde{\theta}_n$ follows :

Theorem 4. *Under assumptions 2 and 9, if moreover Lemma 1 applies to $\sup_{\theta\in U} G[(\cdot), \cdot; \theta]$, then the estimator $\tilde{\theta}_n$ is consistent.*

We will use the notation Y_t for

$$Y_t \stackrel{\text{def}}{=} \Psi[S_{\theta^*}(t) + \zeta_1, t] + \varepsilon_1 ,$$

to simplify the writing of some integrals. We shall introduce the assumptions we need to prove the asymptotic distribution of $\tilde{\theta}_n$:

Assumption 10. *The following conditions are assumed :*

- *For all (z, t) in $\mathbb{R} \times [0, 1]$, the function $\theta \mapsto p_{(\theta,f^*)}(z, t)$ is twice continuously differentiable;*
- *For any θ in Θ, $t \mapsto \mathbb{E}\left\{\left\|\nabla_\theta \log p_{(\theta,f^*)}(Y_t, t)\right\|^2\right\}$ is finite and continuous;*
- *There exists a neighborhood U of θ^* such that, for all θ in U,*
  - *$t \mapsto \mathbb{E}\left\{\nabla^2_\theta \log p_{(\theta,f^*)}(Y_t, t)\right\}$ is finite and continuous;*
  - *Lemma 1 applies to $\log p_{(\theta,f)}(Y_t, t)$, to $\left\|\nabla_\theta \log p_{(\theta,f^*)}(Y_t, t)\right\|^2$ and to all components of $\nabla^2_\theta \log p_{(\theta,f^*)}(Y_t, t)$.*

Introduce the parametric Fisher information matrix :

$$I(\theta) \stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left\{\frac{\nabla_\theta p_{(\theta,f^*)}}{p_{(\theta,f^*)}}(Y_t, t)\frac{\nabla_\theta p_{(\theta,f^*)}}{p_{(\theta,f^*)}}^{\mathrm{T}}(Y_t, t)\right\} \mathrm{d}t .$$

Theorem 5. *Under assumptions 2, 9 and 10, $\tilde{\theta}_n$ converges in probability to θ^* as n tends to infinity. Moreover, if $I(\theta^*)$ is non singular,*

$$\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big) = I^{-1}(\theta^*)\frac{1}{\sqrt{n}}\sum_{k=1}^n \frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}(Y_k, t_k) + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1) ,$$

and $\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big)$ converges in distribution, as $n \to \infty$, to $\mathcal{N}\left(0, I^{-1}(\theta^*)\right)$.

The proof follows the same lines as that of Theorems 1 and 2 and is left to the reader.
Notice that under the same assumptions, it is easy to prove that the parametric model is locally asymptotically normal in the sense of Le Cam (see [9]) so that if $I(\theta^*)$ is singular, there exists no regular estimator of θ which is $\sqrt{n}$-consistent. Thus if $I_R(\theta^*)$ is non singular and the assumptions in Theorem 2 hold, in which case the LSE is regular $\sqrt{n}$-consistent, then $I(\theta^*)$ is also non singular.

To investigate the optimality of possible estimators in the semiparametric situation, with f^* unknown but known to belong to $\mathcal{F}$, we use Le Cam's theory as developed for non i.i.d. observations by Mc Neney and Wellner [10]. Introduce the set $\mathcal{B}$ of integrable functions b on $\mathbb{R}^s$ such that :

- $\int_{\mathbb{R}^s} b(\zeta)\,\mathrm{d}\zeta = 0$ and there exists $\delta > 0$ such that $f^* + \delta b \geq 0$;
- for all t in $[0,1]$, for all θ in Θ, $\int_{\mathbb{R}^s} \Psi\left[S_\theta(t) + \zeta, t\right] b(\zeta)\,\mathrm{d}\zeta = 0$;
- $\int_0^1 \mathbb{E}\left\{\left(\frac{p_{(\theta^*,b)}}{p_{(\theta^*,f^*)}}(Y_t,t)\right)^2\right\}\mathrm{d}t < \infty$.

Let $\mathcal{H} = \mathbb{R}^d \times \mathcal{B}$ be endowed with the inner product :

$$\langle (a,b),(a',b')\rangle_{\mathcal{H}} \stackrel{\text{def}}{=} \int_0^1 \mathbb{E}\left\{\left(\frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}^{\mathrm{T}}(Y_t,t)\cdot a + \frac{p_{(\theta^*,b)}}{p_{(\theta^*,f^*)}}(Y_t,t)\right)\left(\frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}^{\mathrm{T}}(Y_t,t)\cdot a' + \frac{p_{(\theta^*,b')}}{p_{(\theta^*,f^*)}}(Y_t,t)\right)\right\}\mathrm{d}t\,.$$

We will need only local smoothness, so we introduce :

Assumption 11. *There exists a neighborhood U of θ^* such that for θ in U :*

- *for all (z,t) in $\mathbb{R}\times[0,1]$, the function $\theta \mapsto p_{(\theta,f^*)}(z,t)$ is twice continuously differentiable;*
- *The maps $t \mapsto \mathbb{E}\left\{\|\nabla_\theta \log p_{(\theta,f^*)}(Y_t,t)\|^2\right\}$ and $t \mapsto \mathbb{E}\left\{\nabla_\theta^2 \log p_{(\theta,f^*)}(Y_t,t)\right\}$ are finite and continuous;*
- *For any b in $\mathcal{B}$, for all (z,t) in $\mathbb{R}\times[0,1]$, the map $\theta \mapsto p_{(\theta,b)}(z,t)$ is continuously differentiable and $t \mapsto \mathbb{E}\left\{\left\|p_{(\theta^*,f^*)}[Y_t,t]^{-1}\cdot\nabla_\theta \int g\left\{Y_t - \Psi\left[S_\theta(t)+u,t\right]\right\}b(u)\,\mathrm{d}u\right\|\right\}$ is finite and continuous;*
- *Lemma 1 applies to $\left\|\nabla_\theta \log p_{(\theta,f^*)}[Y_t,t]\right\|^2$, all components of $\nabla_\theta^2 \log p_{(\theta,f^*)}[Y_t,t]$ and to $\left\|p_{(\theta^*,f^*)}(Y_t,t)^{-1}\cdot\nabla_\theta p_{(\theta,b)}(Y_t,t)\right\|$ for θ in U.*

Let $\mathbb{P}_{n,(\theta,f)}$ be the distribution of $Y_1,\ldots,Y_n$ when the parameter is θ and the density of the trajectory noise is f. For (θ,f) in $\Theta\times\mathcal{F}$, let

$$\Lambda_n(\theta,f) \stackrel{\text{def}}{=} \log\left[\frac{\mathrm{d}\mathbb{P}_{n,(\theta,f)}(Y_1,\ldots,Y_n)}{\mathrm{d}\mathbb{P}_{n,(\theta^*,f^*)}(Y_1,\ldots,Y_n)}\right] = n[J_n(\theta,f) - J_n(\theta^*,f^*)]\,.$$

Proposition 2. *Assume that Assumption 11 holds. Then the sequence of statistical models $\left(\mathbb{P}_{n,(\theta,f)}\right)_{(\theta,f)\in\Theta\times\mathcal{F}}$ is locally asymptotically normal with tangent space $\mathcal{H}$, that is, for (a,b) in $\mathcal{H}$,*

$$\Lambda_n\left(\theta^* + \frac{a}{\sqrt{n}}, f^* + \frac{b}{\sqrt{n}}\right) = W_n(a,b) - \frac{1}{2}\|(a,b)\|_{\mathcal{H}}^2 + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\,,$$

where

$$W_n(a,b) = \frac{1}{\sqrt{n}}\sum_{k=1}^n\left[\frac{\nabla_\theta p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}^{\mathrm{T}}(Y_k,t_k)\cdot a + \frac{p_{(\theta^*,b)}}{p_{(\theta^*,f^*)}}(Y_k,t_k)\right]\,,$$

and for any finite subset $h_1,\ldots,h_q\in\mathcal{H}$, the random vector $(W_n(h_1),\ldots,W_n(h_q))$ converges in distribution to the centered Gaussian vector with covariance $\langle h_i,h_j\rangle_{\mathcal{H}}$.

Proof.

$$\begin{aligned}\Lambda_n\left(\theta^* + \frac{a}{\sqrt{n}}, f^* + \frac{b}{\sqrt{n}}\right) &= \sum_{k=1}^n \log\left[1 + \frac{p_{(\theta^*+\frac{a}{\sqrt{n}},f^*)} - p_{(\theta^*,f^*)}}{p_{(\theta^*,f^*)}}(Y_k,t_k) + \frac{1}{\sqrt{n}}\frac{p_{(\theta^*+\frac{a}{\sqrt{n}},b)}}{p_{(\theta^*,f^*)}}(Y_k,t_k)\right]\\ &= W_n(a,b) - \frac{1}{2}\|(a,b)\|_{\mathcal{H}}^2 + \mathrm{o}_{\mathbb{P}_{\theta^*}}(1)\,,\end{aligned}$$

by using Taylor expansion till second order of $u \mapsto \log(1+u)$, Taylor expansion till second order of $\theta \mapsto p_{(\theta,f^*)}(z,t)$ and Taylor expansion till first order of $\theta \mapsto p_{(\theta,b)}(z,t)$, which gives the first order term $W_n(a,b)$, and then applying Lemma 1 to the second order terms to get $\frac{1}{2}\|(a,b)\|^2_{\mathcal{H}} + o_{\mathbb{P}_{\theta^*}}(1)$. The convergence of $(W_n(h))_{h\in\mathcal{H}}$ to the isonormal process on $\mathcal{H}$ comes from Lindeberg Theorem applied to finite dimensional marginals. □

The interest of Proposition 2 is that it gives indications on the limitations on the estimation of θ^* when f^* is unknown. Indeed, the *efficient Fisher information* I^* is given by

$$\inf_{b\in\mathcal{B}} \|(a,b)\|^2_{\mathcal{H}} = a^{\mathrm{T}} I^* a \ ,$$

and if I^* is non singular, any regular estimator $\widehat{\theta}_n$ that converges at speed $\sqrt{n}$ has asymptotic covariance Σ which is lower bounded (in the sense of positive definite matrices) by $(I^*)^{-1}$. In case $I_R(\theta^*)$ is non singular and the assumptions in Theorem 2 hold, one may deduce that I^* is non singular.

3.1. Application to BOT

As seen in Section 2.3, the set of isotropic densities with some compact support is a subset of $\mathcal{F}$. If g is twice differentiable, positive and upper bounded, if the trajectory model $\theta \mapsto S_\theta(t)$ is twice differentiable for all $t \in [0,1]$, then Assumptions 10 and 11 hold under almost any trajectory of the observer. Indeed, one may apply Lebesgue's Theorem to obtain derivatives of integrals, and use the fact that the function $z \mapsto \arctan z$ is infinitely differentiable, has vanishing derivatives at infinity, is bounded and has two bounded derivatives.

Moreover, as seen again in Section 2.3, if the trajectory model is (9) and satisfies Assumption 2, then $I_R(\theta^*)$ is non singular, so that the efficient Fisher information I^* is non singular, and all results of Section 3 apply.

4. Further considerations

It would be of great interest to have a more explicit general expression of I^* and of greater interest to exhibit an asymptotically regular and efficient estimator $\widehat{\theta}_n$. If one could approximate the profile likelihood $\sup_{f\in\mathcal{F}} J_n(\theta,f)$, one could hope that the maximizer $\widehat{\theta}_n$ of it be a good candidate.

Another possibility would be to use Bayesian estimators. Indeed, in the parametric context, the Bernstein-von Mises Theorem tells us that asymptotically, the posterior distribution of the parameter is gaussian, centered at the maximum likelihood estimator and with variance the inverse of Fisher information (see [14] for a nice presentation). Extensions to semiparametric situations are now available, see [4]. To obtain semiparametric Bernstein-von Mises Theorems, one has to verify assumptions relating the particular model and the choice of the non parametric prior. This could be the object of further work. Then, with an adequate choice of the prior on $\Theta \times \mathcal{F}$, taking advantage of MCMC computations, one could propose bayesian methods to estimate θ^* (mean posterior, maximum posterior, median posterior for example).

To extend the results of the preceding sections in the case where the trajectory noise is no longer a sequence of i.i.d. random variables, one needs to prove laws of large numbers and central limit theorems for empirical sums such as $n^{-1}\sum_{k=1}^n F(\varepsilon_k, t_k)$, we prove some below for stationary weakly dependent sequences $\{\zeta_k\}_{k\geq 0}$. In such a case, if $M(\theta)$ and $J(\theta,f^*)$ are still the limits of $M_n(\theta)$ and $J_n(\theta,f^*)$ respectively, then asymptotics for $\bar{\theta}_n$ and $\tilde{\theta}_n$ could be obtained. Here, $J_n(\theta,f^*)$ is no longer the normalized log-likelihood, rather the marginal normalized log-likelihood, but $J(\theta,f^*)$ is still a contrast function.

Since the convergence of the expectation relies on purely deterministic arguments (Rieman integrability), we focus on centered functions. We assume in this section that :

Assumption 12. *$\{\zeta_k\}_{k\geq 0}$ is a stationary sequence of random variables such that, for all t in $[0,1]$,*

$$\mathbb{E}[F(\zeta_1,t)] = 0 \ .$$

Denote by $\{\alpha_k\}_{k\geq 0}$ the strong mixing coefficients of the sequence $\{\zeta_k\}_{k\geq 0}$ defined as in [12], that is, for $k \geq 1$,

$$\alpha_k \stackrel{\text{def}}{=} 2 \sup_{\ell\in\mathbb{N},\, A\in\sigma(\zeta_i : i\leq \ell),\, B\in\sigma(\zeta_i : i\geq k+\ell)} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \ ,$$

and $\alpha_0 = 1/2$. Notice that they are also an upper bound for the strong mixing coefficients of the sequence $\{F(\zeta_k, t_k)\}_{k\geq 0}$ for any sequence $\{t_k\}_{k\geq 0}$ of real numbers in $[0,1]$.

Proposition 3. *Under Assumption 12, if α_k tends to 0 as $k \to \infty$, if $\sup_{t\in[0,1]} \mathbb{E}|F(\zeta_1, t)|$ is finite and $\sup_{t\in[0,1]} \mathbb{E}\{|F(\zeta_1, t)|\mathbb{1}_{|F(\zeta_1,t)|>M}\}$ tends to 0 as $M \to \infty$, then, as $n \to \infty$,*

$$n^{-1} \sum_{k=1}^{n} F(\zeta_k, t_k) = \mathrm{o}_{\mathbb{P}}(1) \ .$$

Proof. Using Ibragimov's inequality [6], for any $M > 0$,

$$\begin{aligned}
\mathbb{V}&\left(n^{-1}\sum_{k=1}^{n} F(\zeta_k, t_k)\mathbb{1}_{|F(\zeta_k,t_k)|\leq M}\right) \\
&= n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}\left[F(\zeta_i, t_i)\mathbb{1}_{|F(\zeta_i,t_i)|\leq M}, F(\zeta_j, t_j)\mathbb{1}_{|F(\zeta_j,t_j)|\leq M}\right] \ , \\
&\leq 2M^2 n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{|i-j|} \ , \\
&\leq 2M^2 n^{-1}\sum_{k=0}^{n-1} \alpha_k \ ,
\end{aligned}$$

which tends to 0 by Cesaro's lemma as $n \to \infty$. The end of the proof is similar to that of Lemma 1. □

Define now

$$\alpha^{-1}(u) \stackrel{\text{def}}{=} \inf\{k \in \mathbb{N} : \ \alpha_k \leq u\} = \sum_{i\geq 0} \mathbb{1}_{u<\alpha_i} \ .$$

Define also, for any t in $[0,1]$,

$$Q_t(u) \stackrel{\text{def}}{=} \inf\{x \in \mathbb{R} \ : \ \mathbb{P}\left[|F(\zeta_1, t)| > x\right] \leq u\} \ ,$$

and

$$Q(u) \stackrel{\text{def}}{=} \sup_{t\in[0,1]} Q_t(u) \ .$$

We shall assume that :

Assumption 13. $\displaystyle\int_0^1 \alpha^{-1}(u)\, Q^2(u)\, \mathrm{d}u < \infty$.

This condition is the same as the convergence of the series $\displaystyle\sum_{k\geq 0}\int_0^{\alpha_k} Q^2(u)\, \mathrm{d}u$.

Applying [12, Theorem 1.1], one gets, for any t in $[0,1]$ and $k \geq 0$,

$$|\mathrm{Cov}[F(\zeta_0, t), F(\zeta_k, t)]| \leq 2\int_0^{\alpha_k} Q^2(u)\, \mathrm{d}u \ ,$$

so that if Assumption 13 holds, one may define

$$\gamma^2 \stackrel{\text{def}}{=} \int_0^1 \mathbb{V}[F(\zeta_0,t)]\,\mathrm{d}t + 2\sum_{k=1}^{+\infty}\int_0^1 \mathrm{Cov}[F(\zeta_0,t),F(\zeta_k,t)]\,\mathrm{d}t\ . \tag{14}$$

Now :

Proposition 4. *Under Assumptions 12 and 13, if* $\sigma^2 > 0$ *and if for any integer* k*, the real function* $(t,u) \mapsto \mathrm{Cov}[F(\zeta_0,t),F(\zeta_k,u)]$ *is continuous on* $[0,1]^2$*, then*

$$n^{-1/2}\sum_{k=1}^{n} F(\varepsilon_k,t_k) \rightsquigarrow \mathcal{N}(0,\gamma^2)\ ,$$

as $n \to \infty$*.*

Proof. Let $S_n = \sum_{k=1}^n F(\zeta_k,t_k)$. First of all, let us prove that $n^{-1}\mathbb{V}(S_n)$ converges to σ^2 as $n \to \infty$.

$$\begin{aligned} n^{-1}\mathbb{V}(S_n) &= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}[F(\zeta_i,t_i),F(\zeta_j,t_j)]\ ,\\ &= n^{-1}\sum_{k=1-n}^{n-1}\sum_{i=1\vee(1-k)}^{n\wedge(n-k)}\mathrm{Cov}[F(\zeta_0,t_i),F(\zeta_k,t_{i+k}t)]\ . \end{aligned}$$

For any $K \geq 1$, using again [12, Theorem 1.1],

$$\left| n^{-1}\sum_{K\leq|k|\leq n-1}\sum_{i=1\vee(1-k)}^{n\wedge(n-k)}\mathrm{Cov}[F(\zeta_0,t_i),F(\zeta_k,t_{i+k})]\right| \leq 2\sum_{k\geq K}\int_0^{\alpha_k} Q^2(u)\,\mathrm{d}u\ ,$$

which is smaller than any positive ϵ for big enough K under Assumption 13. Now, for any fixed integer k,

$$\begin{aligned} &\left| n^{-1}\sum_{i=1\vee(1-k)}^{n\wedge(n-k)}\mathrm{Cov}[F(\zeta_0,t_i),F(\zeta_k,t_{i+k})] - \int_0^1 \mathrm{Cov}[F(\zeta_0,t),F(\zeta_k,t)]\,\mathrm{d}t\right| \\ &\qquad \leq \sup_{(t,u)\in[0,1]^2,\,|t-u|\leq k/n} |\mathrm{Cov}[F(\zeta_0,t),F(\zeta_k,t+u)] - \mathrm{Cov}[F(\zeta_0,t),F(\zeta_k,t)]| \\ &\qquad\quad + \frac{k}{n}\sup_{t\in[0,1]}|\mathrm{Cov}[F(\zeta_0,t),F(\zeta_k,t)]|\ , \end{aligned}$$

which goes to 0 as $n \to \infty$ under the continuity assumption. The convergence of $n^{-1}\mathbb{V}(S_n)$ to σ^2 follows. The end of the proof is a direct application of [11, Corollary 1]. □

5. Simulations

The simulations have been realized using `Matlab`. The minimisation is made with the function `searchmin` by setting to 2000 the options `MaxFunEvals` and `MaxIter`, so that the method reaches the minimum.

For all the simulations, the observation time is of 20 s. The trajectory of the observer has a speed with constant norm $\left\|\frac{\mathrm{d}O(t)}{\mathrm{d}t}\right\|$ equal to 0.25 km/s and makes maneuvers with norm of acceleration $\left\|\frac{\mathrm{d}^2O(t)}{\mathrm{d}t^2}\right\|$ of approximatively $50\,\mathrm{m/s}^2$. The trajectory is mainly composed of uniform linear motions and circular uniform motions. The different sequences of the trajectory of the platform are described in the following table. The null values of acceleration correspond to uniform linear motions and the others to uniform circular motion.

time interval (s)	$0-6$	$7-10$	$11-14$	$15-20$
norm of acceleration(m/s^2)	50	0	-55	0

The positive and negative values for norm of acceleration correspond respectively to anticlockwise and clockwise circular motion. The transition sequences between circular motion and linear motion which are the time intervals $[6,7]$, $[10,11]$, and $[14,15]$ are such that the whole trajectory is $\mathscr{C}^\infty$.

The assumed parametric model is a uniform linear motion with a speed of 0.27 km/s. The parameter θ is defined by

$$\theta = (x_0, y_0, v_x, v_y) \; .$$

where (x_0, y_0) denotes the initial position and (v_x, v_y) the speed vector. The parametric trajectory is then defined by

$$S_\theta(t) = (x_0 + v_x t,\; y_0 + v_y t) \; .$$

The observation noise is a sequence of i.i.d centered Gaussian variables with variance $\sigma = 10^{-3}$ rad. The platform receives 2000 observations.

For the first simulation, we consider a sequence $\{\zeta_k\}_{k\geq 0}$ of i.i.d Gaussian centered random variables with variance $\sigma_X^2 \times I_2$ and $\sigma_X = 10\,\mathrm{m}$[1]. The figure 3 shows the trajectory of the platform with a realization of a trajectory of the target and the parametric trajectory with parameter $\bar{\theta}_n$ and also the confidence area with level of 95% for the position at final time. The figure 6 presents the same for the maximum likelihood estimator (MLE) $\tilde{\theta}_n$.

By using Monte-Carlo methods with 1000 experiments, histograms of the coordinates of $\sqrt{n}\left(\bar{\theta}_n - \theta^*\right)$ are presented on figure 4 with the marginal probability densities of the asymptotic law $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$ in dotted line. The empirical cumulative distribution functions of the coordinates of $\sqrt{n}\left(\bar{\theta}_n - \theta^*\right)$ are presented on figure 5 juxtaposed to the marginal cumulative distributions of law $\mathcal{N}\left(0, I_M^{-1}(\theta^*)\right)$. These two figures illustrate the convergence in distribution given by Theorem 2, since the sequence $\{\zeta_k\}_{k\geq 0}$ is an i.i.d. sequence of isotropic random variables. The figure 7 present the histograms of the coordinates of $\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big)$ with the marginal probability densities of the asymptotic law $\mathcal{N}\left(0, I^{-1}(\theta^*)\right)$ in dotted line. Empirical cumulative distribution functions of the coordinates of $\sqrt{n}\big(\tilde{\theta}_n - \theta^*\big)$ and marginal cumulative distributions of law $\mathcal{N}\left(0, I^{-1}(\theta^*)\right)$ are presented on figure 8. These two figures illustrate the convergence in distribution given by Theorem 5. Confidence intervals for coordinates of θ^* with level of 95% are detailed in table 1 for $\bar{\theta}_n$ and in table 3 for $\tilde{\theta}_n$ and are respectively denoted by $\mathrm{IC}_1\left(\bar{\theta}_n\right)$ and $\mathrm{IC}_3\big(\tilde{\theta}_n\big)$. We also present in table 2 conservative confidence intervals denoted by $\mathrm{IC}_2\left(\bar{\theta}_n\right)$ built on the result provided by Theorem 3 with $R_{\min} = 6$ km. The choice of σ_X and $R_{\min}$ is a prior knowledge on the experiment and is made according to the knowledge of the tactical situation of BOT. Note that the majoration obtained in (10) shows that the accuracy of the conservative confidence intervals is proportional to the ratio $\mathbb{E}(\|\zeta_1\|^2)/R_{\min}^2$. This result is very interesting in practice since it shows that for high values of relative distance between target and observer and small values of state noise variance, conservative confidence intervals are of high accuracy.

For these simulations, one needs to calculate $I_\Psi\left(\bar{\theta}_n\right)$, $I_\Psi(\theta^*)$, $I\big(\tilde{\theta}_n\big)$ and $I(\theta^*)$ which involve expectations of functions of the r.v. ε_1 with law $\mathcal{N}(O, \sigma_X^2 \times I_2)$. All integrals of this type has been calculated using quadrature formula with 12 points. Abscissas and weight factors are given in [1]. Let us detail the numerical values of $I_\Psi\left(\bar{\theta}_n\right)$ and $\sigma^2 \times I_R\left(\bar{\theta}_n\right)$ for one experiment used to build the estimators $\bar{\theta}_n$ and $\tilde{\theta}_n$. These numerical values

[1]The variance is small enough to make assumption 7 valid, even if the noise is Gaussian.

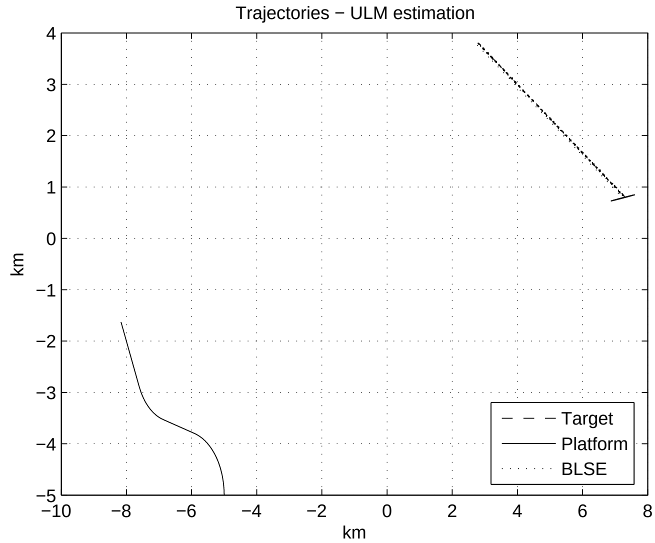

FIGURE 3. Trajectories with confidence area for BLSE at final position

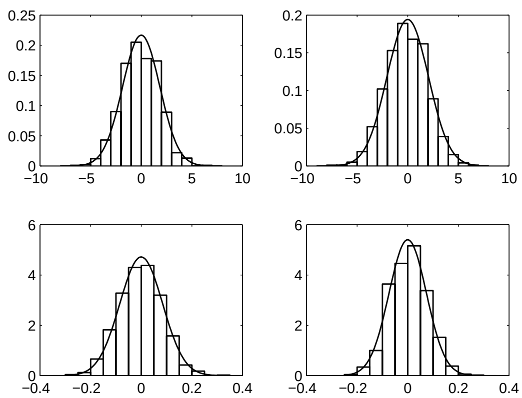

FIGURE 4. Histograms for BLSE with iid Gaussian isotropic sequence

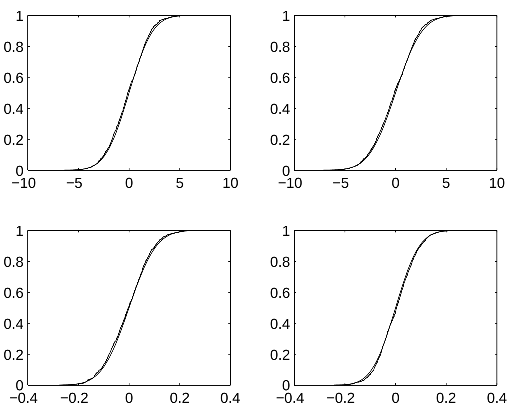

FIGURE 5. Cumulative distribution functions for BLSE with iid Gaussian isotropic sequence

illustrate that the contributions of state noise and observation noise are of the same level.

$$I_\Psi\left(\bar{\theta}_n\right) = 10^{-6} \times \begin{pmatrix} 0.0010 & -0.0014 & 0.0049 & -0.0094 \\ -0.0014 & 0.0024 & -0.0094 & 0.0220 \\ 0.0049 & -0.0094 & 0.0400 & -0.0950 \\ -0.0094 & 0.0220 & -0.0950 & 0.2709 \end{pmatrix},$$

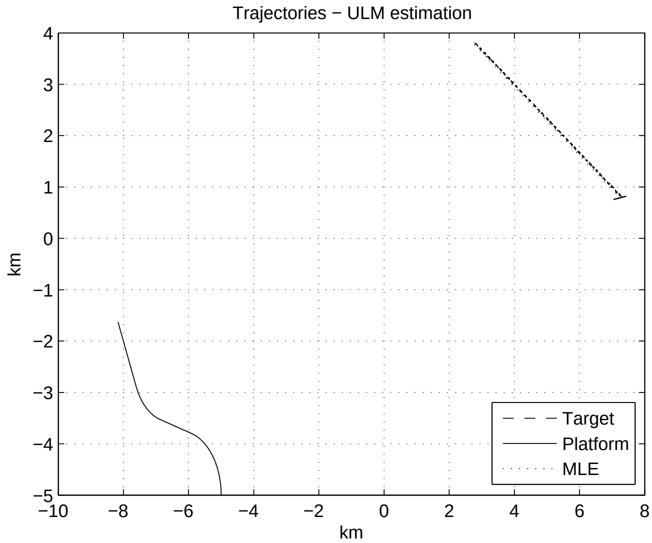

Figure 6. Trajectories with confidence area for MLE at final position

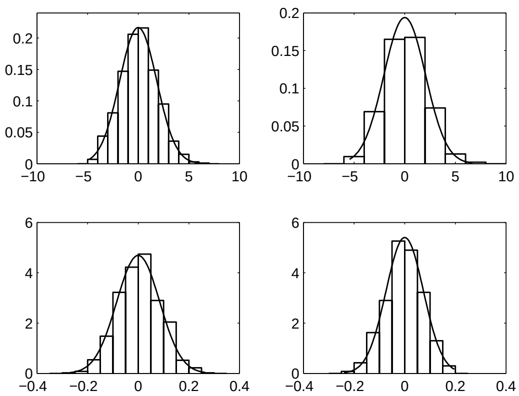

Figure 7. Histograms for MLE with iid Gaussian isotropic sequence

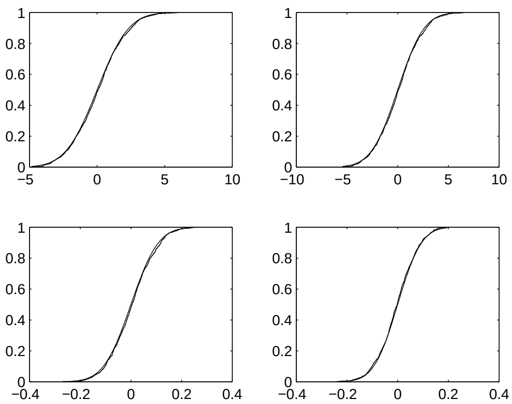

Figure 8. Cumulative distribution functions for MLE with iid Gaussian isotropic sequence

$$\sigma^2 \times I_R\left(\bar{\theta}_n\right) = 10^{-6} \times \begin{pmatrix} 0.0015 & -0.0023 & 0.0082 & -0.0169 \\ -0.0023 & 0.0043 & -0.0169 & 0.0428 \\ 0.0082 & -0.0169 & 0.0728 & -0.1853 \\ -0.0169 & 0.0428 & -0.1853 & 0.5639 \end{pmatrix}.$$

$\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)\right\rvert$
7.3128	7.5747	0.2619
0.8017	0.8456	0.0439
0.2253	0.2316	0.0063
-0.1558	-0.1503	0.0055

Table 1. Confidence intervals for BLSE at level 95%

$\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)\right\rvert$
6.0645	8.8230	2.7586
0.5917	1.0557	0.4640
0.1949	0.2619	0.0669
-0.1818	-0.1242	0.0576

Table 2. Conservative confidence intervals for BLSE at level 95%

$\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)$		$\left\lvert\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)\right\rvert$
7.1842	7.4430	0.2588
0.7815	0.8249	0.0434
0.2222	0.2285	0.0063
-0.1529	-0.1475	0.0054

Table 3. Confidence intervals for MLE at level 95%

Let us now precise the values of variance matrices. We have

$$I_M^{-1}\left(\bar{\theta}_n\right) = \begin{pmatrix} 3.4917 & 3.8949 & 0.1560 & -0.1399 \\ 3.8949 & 4.3496 & 0.1752 & -0.1561 \\ 0.1560 & 0.1752 & 0.0074 & -0.0062 \\ -0.1399 & -0.1561 & -0.0062 & 0.0056 \end{pmatrix},$$

and

$$I^{-1}\left(\tilde{\theta}_n\right) = \begin{pmatrix} 3.3918 & 3.7884 & 0.1526 & -0.1359 \\ 3.7884 & 4.2362 & 0.1715 & -0.1518 \\ 0.1526 & 0.1715 & 0.0072 & -0.0061 \\ -0.1359 & -0.1518 & -0.0061 & 0.0055 \end{pmatrix}.$$

The true parameter θ^* is

$$\theta^* = (2.8, 3.8, 0.225, -0.15)\ ,$$

and values of estimators $\bar{\theta}_n$ and $\tilde{\theta}_n$, used to calculate variance matrices, are

$$\bar{\theta}_n = (2.8753, 3.8841, 0.2284, -0.1530)\ , \qquad \tilde{\theta}_n = (2.8067, 3.8077, 0.2253, -0.1502)\ ,$$

with initial coordinates $x_0,\ y_0$ given in km and $v_x,\ v_y$ given in km/s and the position at final time is $(7.3, 0.8)$. It appears that the maximum likelihood estimator $\tilde{\theta}_n$ is a bit more accurate than $\bar{\theta}_n$. It is not surprising since the MLE is designed specifically for the model, and takes into account the state noise. Nevertheless, because of the high calculation cost for the MLE, the BLSE is in practice a very useful alternative.

For the second simulation, we consider the case of a sequence $\{\zeta_k\}_{k\geq 0}$ of i.i.d Gaussian centered random variables with variance $\sigma_X^2 \times \begin{pmatrix} 6^2 & 0 \\ 0 & 1 \end{pmatrix}$ and $\sigma_X = 10\,\mathrm{m}$. It seems that the results given by Theorems 2 and 5 still hold, even though the sequence $\{\zeta_k\}_{k\geq 0}$ does not have an isotropic distribution, see Figures 9, 10, 11 and 12. The estimators values are

$$\bar{\theta}_n = (2.8383, 3.8440, 0.2264, -0.1516)\ , \qquad \tilde{\theta}_n = (2.7984, 3.7999, 0.2253, -0.1499)\ .$$

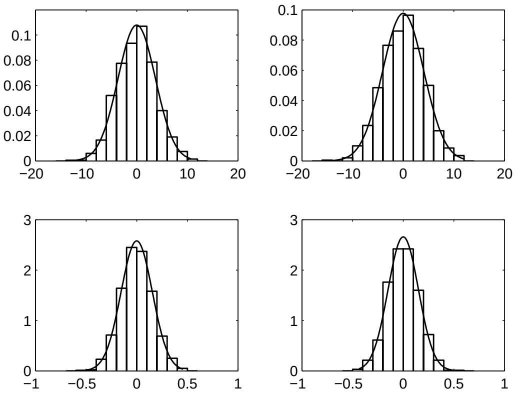

FIGURE 9. Histograms for BLSE with iid Gaussian non-isotropic sequence

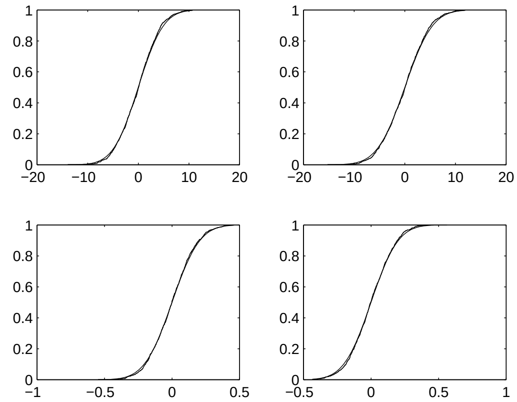

FIGURE 10. Cumulative distribution functions for BLSE with iid Gaussian non-isotropic sequence

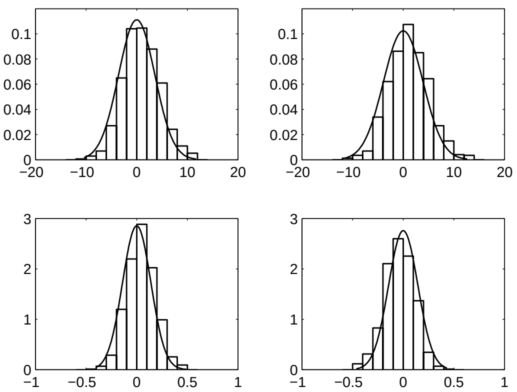

FIGURE 11. Histograms for MLE with iid Gaussian non-isotropic sequence

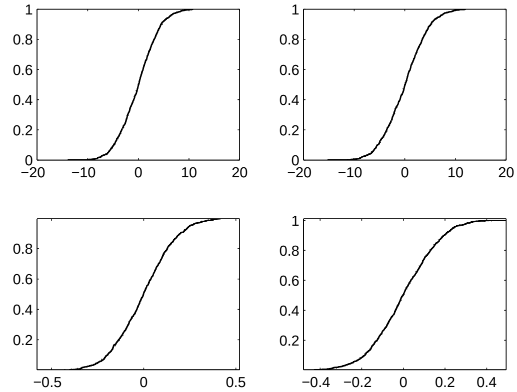

FIGURE 12. Cumulative distribution functions for MLE with iid Gaussian non-isotropic sequence

The values of variance matrices for the two estimators are

$$I_M^{-1}\left(\bar{\theta}_n\right) = \begin{pmatrix} 15.4505 & 17.0122 & 0.6174 & -0.6253 \\ 17.0122 & 18.7661 & 0.6863 & -0.6889 \\ 0.6174 & 0.6863 & 0.0263 & -0.0250 \\ -0.6253 & -0.6889 & -0.0250 & 0.0253 \end{pmatrix},$$

and

$$I^{-1}(\tilde{\theta}_n) = \begin{pmatrix} 12.9538 & 14.0399 & 0.4766 & -0.5214 \\ 14.0399 & 15.2720 & 0.5262 & -0.5661 \\ 0.4766 & 0.5262 & 0.0197 & -0.0192 \\ -0.5214 & -0.5661 & -0.0192 & 0.0210 \end{pmatrix}.$$

$\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)$		$\left\|\mathrm{IC}_1\left(\bar{\theta}_{n,i}\right)\right\|$
7.1040	7.6275	0.5235
0.7698	0.8552	0.0854
0.2204	0.2323	0.0119
-0.1574	-0.1457	0.0117

Table 4. Confidence intervals for BLSE at level 95%

$\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)$		$\left\|\mathrm{IC}_2\left(\bar{\theta}_{n,i}\right)\right\|$
1.5049	13.2266	11.7218
-0.1740	1.7990	1.9730
0.0842	0.3686	0.2844
-0.2740	-0.0291	0.2449

Table 5. Conservative confidence intervals for BLSE at level 95%

$\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)$		$\left\|\mathrm{IC}_3\left(\tilde{\theta}_{n,i}\right)\right\|$
7.0721	7.5366	0.4645
0.7643	0.8388	0.0746
0.2201	0.2305	0.0103
-0.1552	-0.1446	0.0107

Table 6. Confidence intervals for MLE at level 95%

The confidence intervals detailed in table 4 and table 6 show that the maximum likelihood estimator $\tilde{\theta}_n$ is significantly more accurate than the BLSE. Comparing to the first simulation where the difference is not so large, the higher accuracy of $\tilde{\theta}_n$ can be understood because of the higher level state noise in this simulation. Then, taking into account this state noise for estimating the parameter provides a significantly better result. The conservative intervals for $R_{\min} = 6\,\mathrm{km}$ described in table 5 are quite large compared to those obtained for the first simulation. This inaccuracy results directly from the large value of $\mathbb{E}\left(\|\zeta_1\|^2\right)$ chosen for the state noise.

For the third and last simulation, the sequence $\{\zeta_k\}_{k\geq 0}$ is an AR(1) series such that, for all k integer,

$$\zeta_{k+1} = \Phi\zeta_k + \eta_k\ ,$$

where $\Phi = 0.6$ and $\{\eta_k\}_{k\geq 0}$ is a sequence of i.i.d. random variables with law $\mathcal{N}(0, \sigma_\eta^2)$ and $\sigma_\eta = 8\,\mathrm{m}$. Thus, the sequence of state noise $\{\zeta_k\}_{k\geq 0}$ is a dependent stationary sequence such that the mixing coefficient α_k tends exponentially fast to zero as k tends to infinity. Then, we observe the predicted behavior described by Proposition 4. Indeed, by drawing the densities and cumulative distribution functions of the centered Gaussian law with the empirical variance, we observe a very good adequacy to the Gaussian behavior, see figures 13 and 14.

The authors want to thank Jerôme Dedecker for helpful discussions about dependent sequences of random variables.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. National Bureau of Standards Applied Mathematics Series. 1964.

[2] A. Averbuch, Y. Bar-Shalom, E. Mazor, and J. Dayan. Interacting multiple model methods in target tracking: A survey. *IEEE Trans. Aerospace Elect. Systems*, 34(1):103–123, January 1998.

[3] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. Wiley-Interscience, 2001.

[4] I. Castillo. A semi-parametric bernstein-von mises theorem. 2008. Submitted.

[5] A. Doucet, N. De Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer, New York, 2001.

[6] I.A. Ibragimov. Some limit theorems for stationary processes. *Theory of probability and its applications*, 7:349–382, 1962.

[7] B. Landelle. *Étude statistique du problme de la trajectographie passive*. PhD thesis, Universit Paris-Sud, Laboratoire de Mathmatiques, 91405 ORSAY Cedex, 2008.

[8] B. Landelle. Robustness considerations for bearings-only tracking. In *11th International Conference on Information Fusion*, 2008.

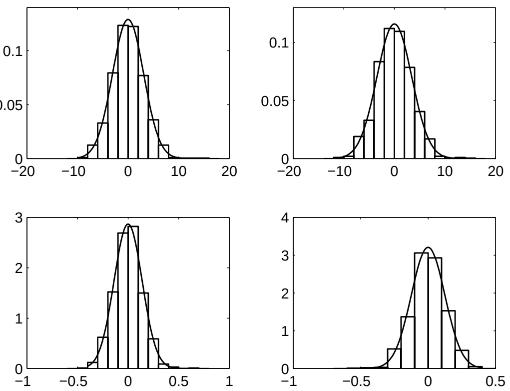

FIGURE 13. Histograms for AR(1) sequence, Gaussian adequacy

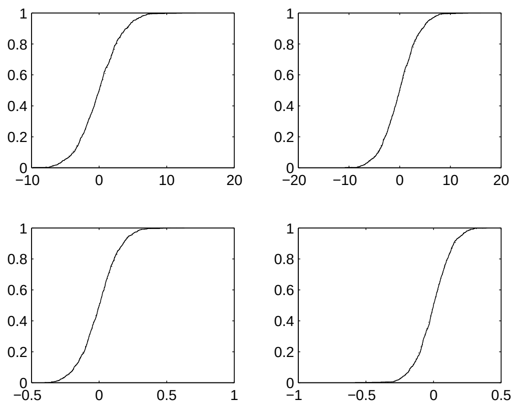

FIGURE 14. Cumulative distribution functions for AR(1) sequence, Gaussian adequacy

[9] L. Le Cam. *Asymptotic methods in statistical decision theory.* Springer Series in Statistics. New York etc.: Springer-Verlag. XXVI, 742 p. DM 138.00, 1986.
[10] Brad McNeney and Jon A. Wellner. Application of convolution theorems in semiparametric models with non-i.i.d. data. *J. Statist. Plann. Inference*, 91(2):441–480, 2000.
[11] E. Rio. About the lindeberg method for strongly mixing sequences. *ESAIM : Probability and Statistics*, 1:35–61, 1995.
[12] E. Rio. *Thorie asymptotique des processus alatoires faiblement dpendants*. Mathmatiques et Applications. Springer, 2000.
[13] B. Ristic, S. Arulampalam, and A. Gordon. *Beyond Kalman Filters: Particle Filters for Target Tracking.* Artech House, 2004.
[14] A. W. Van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.