



HAL
open science

Maximum likelihood trajectory for the velocity of an artillery shell in free-flight

R Séailles, Nicolas Petit

► **To cite this version:**

R Séailles, Nicolas Petit. Maximum likelihood trajectory for the velocity of an artillery shell in free-flight. [Research Report] Mines ParisTech - Ecole des mines de Paris. 2021. hal-03216640

HAL Id: hal-03216640

<https://minesparis-psl.hal.science/hal-03216640>

Submitted on 4 May 2021

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

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

Maximum likelihood trajectory for the velocity of an artillery shell in free-flight

R. Séailles^a, N. Petit^b

a. P19, romain.seailles@mines-paristech.fr

b. CAS, MINES ParisTech, PSL University, 60 bd Saint-Michel, 75272, Paris, France,
nicolas.petit@mines-paristech.fr

Key words : velocity estimation; onboard sensors; maximum likelihood estimation; causal estimator;

Abstract : *The paper considers the problem of determining the translational velocity of an artillery shell from noisy information. To optimize the smoothness of the estimate which is to be used in a attitude observer through an analysis of its slope, the unknown is sought after under the form of an element of a family of possible curves, solely parameterized by their initial condition. The nonlinear dynamics of the velocity being a contraction, it is shown that a measurement bias is observable. The estimator statistical properties are studied when the number of measurement is increased.*

Contents

1	Introduction and context	2
1.1	Wahba's problem, definition and application to a shell in free-flight	2
1.2	Finding the pitch angle using an estimate of the translational velocity	3
2	Proposed estimator : find the unknown among a set of pre-calculated curves	4
2.1	Curves generation	4
2.2	About causality	5
2.3	Stochastic approach	6
2.4	Maximum likelihood Estimator	6
2.5	First results	6
3	Effects of a constant bias	7
3.1	Contracting behavior	8
3.2	Impact of a constant bias on the estimator	8
3.3	Correction of our estimator	9
3.4	A new estimator	9
4	Convergence analysis	9
4.1	Main result	9
4.2	Proof	10
4.2.1	A Lemma from the literature	10
4.2.2	Analysis of the solutions of the differential equation	11

4.2.3	Rewriting the likelihood function	11
4.2.4	Standard deviation	11
4.2.5	Determining the limit function of L_n	13
4.2.6	Continuity and global minimum	14
4.2.7	Lemma of uniform continuity	14
4.2.8	Analysis of the error $L_n(u_0) - L_n(\tilde{u}_0)$	15
4.2.9	$\frac{1}{n+1} \sum_{i=0}^n N_i - \bar{N} $ is bounded in probability	16
4.2.10	Conclusion	17
5	Results, conclusion and perspectives	17
6	Acknowledgments	18

1 Introduction and context

In this paper, we propose a technique to improve the method proposed in [4, 5, 6] to estimate the attitude of an artillery shell in free flight. Determining the attitude of a shell means finding its 3D-orientation with respect to fixed reference frame. This information is instrumental to redirect the shell while flying, to improve its accuracy, and to increase its range and its overall efficiency. It is an essential step in the development of smart shells or guided-ammunition, which is a very promising technology [4, 2].

1.1 Wahba's problem, definition and application to a shell in free-flight

Attitude Estimation of a shell in free flight is a well-known problem that can be mathematically described using *Wahba's problem* [10].

This problem is to find a rotation matrix (i.e. a matrix from $SO_3(\mathbb{R})$) between two coordinate systems, in our case between the shell frame and the (fixed) reference Earth frame, though the minimization of a cost function expressed using the measurements of direction vectors. The cost function that *Wahba's problem* seeks to minimize is defined as follows :

$$J(\mathbf{R}) = \sum_{k=1}^N \|\vec{u}_k - \mathbf{R}\vec{v}_k\|^2$$

where $\mathbf{R} \in SO_3$, \vec{u}_k is a vector measurement in the shell frame and \vec{v}_k is the same vector expressed in the reference frame.

What is important to consider while considering the *Wahba's problem* is that a solution can only be found if $\mathbf{N} \geq 2$: at least 2 linearly independent vector measurement expressed in both frames to be able are required to find a unique solution.

In most applications of attitude determination using Inertial Measurement Units (IMU), accelerometers and magnetometers are employed. Under the usual assumption that the acceleration is zero (on average), the accelerometer measures the Earth gravity which is a fixed direction. The magnetometer measures the Earth magnetic field. Both vector measurements can be readily normalized, as explained in [4] to provided two linearly independent directions.

However, as the shell is in free flight, it is impossible to read gravity from any on-board accelerometers. Therefore, the direct solution of Wahba's problem is impossible here. One has to find another solution to get more information.

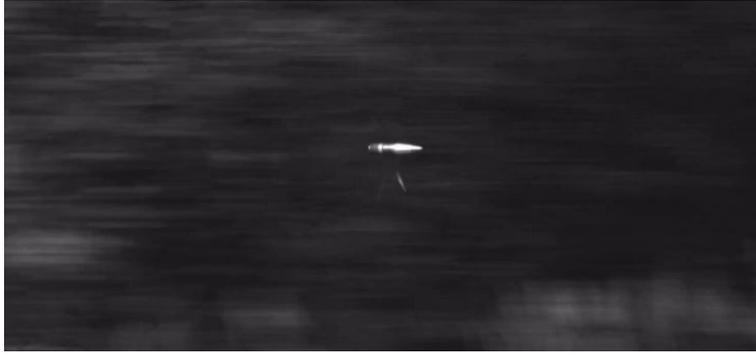


Figure 1: ISL Shell

Following the method proposed in [4]:

Instead of 2 vectors, the *Wahba's problem* can be solved using *1 vector* and *1 angle*.

This means that one needs to find 1 angle measurement in the shell and then, considering one already got Earth magnetic field direction using magnetometers, it would be possible to find the attitude of the shell. For this, a specific method has been developed. The convergence of the proposed attitude estimation methodology has been established, stressing the complementarity of pitch angle information and magnetic vector measurement.

1.2 Finding the pitch angle using an estimate of the translational velocity

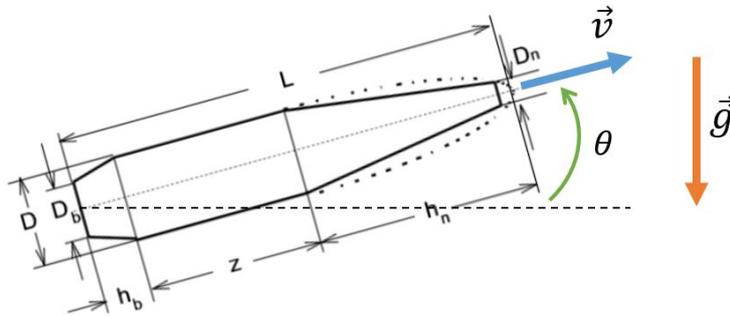


Figure 2: Pitch angle definition

The *pitch angle* is defined as the angle between the horizontal plane and the velocity vector \vec{v} of the shell. Using the dynamic equations projected along \vec{v} , one has

$$\dot{v} = -\frac{\rho S C_D}{2M} v^2 - g \sin(\theta)$$

where the quadratic term is the aerodynamic drag, and the term related to g comes from gravity. This equation gives a direct relationship between v and θ , which is almost equal to the *pitch angle* we are seeking. Loosely speaking, the time variations of the velocity give direct information on the value of θ . In turn, this stresses that, in view of determining the full attitude of the shell, any candidate method to estimate v should avoid to features bumps and other oscillations.

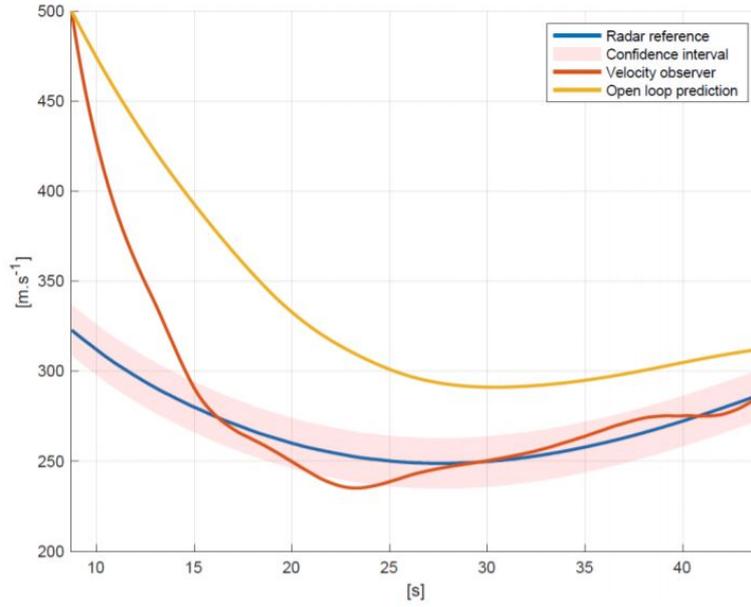


Figure 3: Estimation of v using a state observer

Figure 3 reports typically results obtained in [4] using a Luenberger-type state observer to estimate the velocity. As the pitch angle θ is directly linked to the slope of this curve, those irregularities lead to massive inaccuracy while trying to estimate θ .

It looks to be the main performance bottleneck. In this paper we propose *a new estimator of v which gives us a smoother curve.*

The paper is organized as follows. In Section 2, we present the new estimator. In Section 3, we extend it to account for a (constant) bias on the measurement. In Section 4, we conduct a convergence analysis. In Section 5 some numerical results are presented along with conclusions and perspectives.

2 Proposed estimator : find the unknown among a set of pre-calculated curves

2.1 Curves generation

The rotational and translational dynamics gives us the following set of equations, (see [9] for detailed notations). p, q and r are angular velocities of the shell among 3 axis. Those four equations are linked (because $\dot{\theta} = p$) and one need to solve them all together to get v among time.

$$\begin{cases} \dot{v} = -\frac{\rho S C_D}{2M} v^2 - g \sin(\theta) \\ \dot{p} = \frac{\rho(h) S D^2 C_{spin} v}{2I_t} p + \frac{1}{2} \rho(h) S D \delta_{fincant} C_{l\delta} v^2 \\ \dot{q} = \frac{1}{I_t} ([I_l - I_t] p r + \frac{1}{2} \rho S p D^2 C_{mag} v \beta + \frac{1}{2} \rho S D C_{M\alpha} v^2 \alpha + \frac{1}{2} \rho S D^2 C_{Mq} v q) \\ \dot{r} = \frac{1}{I_t} ([I_l - I_t] p q + \frac{1}{2} \rho S p D^2 C_{mag} v \alpha - \frac{1}{2} \rho S D C_{M\alpha} v^2 \beta + \frac{1}{2} \rho S D^2 C_{Mr} v r) \end{cases} \quad (1)$$

With those equations, and with enough computational power, one can integrate those laws over the (short) duration of the flight and obtain a set of simulated curves of v through time. Those curves are generated based on a uniformly distributed list of initial velocities in a compact \mathcal{K} , covering all the possible values for the initial translational velocity. As it is physically impossible to have a negative velocity, as of now, it will be acknowledging that $\mathcal{K} \in \{v_0 \in \mathbb{R} | \forall t, v(v_0, t) \geq 0\}$.

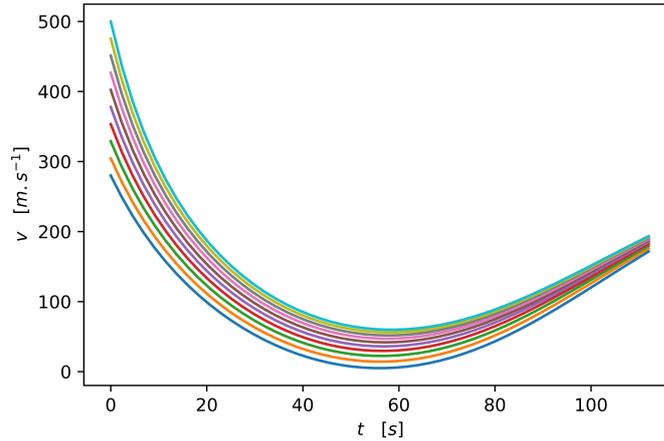


Figure 4: 10 velocity curves initiated with $v_0 \in \mathcal{K} = [280m.s^{-1}, 500m.s^{-1}]$

Straightforwardly, from the Cauchy–Lipschitz theorem states, for a given initial condition, there exists a unique solution to our differential equation. Thus, it is perfectly equivalent to know the value of v at every moment and to only know the value of v_0 that produce this velocity profile.

According to this consideration, this paper will now focus on building a good estimator of v_0 instead of v at any given time. If, by using all the data, one can find the best value of v_0 that generate a curve v that closely matches at every time step the data we got, then this value allows us to find the best value v at any moment.

2.2 About causality

What is important to know is that our method is designed to be used in real-time applications. Therefore, it is mandatory to talk about *causality*. Causality means that at a given time t of the flight, the only data available to estimate the shell velocity t are the measurements made before t . To formalize this concept, the current time of the experience is named t_f . Getting $n \in \mathbb{N}$, n measurements are available, respectively taken at $t_k = t_f \frac{k}{n}$.

Gathering the considerations above, the purpose of our estimator is to determine the approximation of v_0 that maximizes the likelihood of the velocity of our shell right now (i.e. at t_f) to be $v(v_0, t_f)$ (which is the value at t_f of the velocity profile generated by the initial value v_0).

2.3 Stochastic approach

Let $v_0 \in \mathcal{K}$, where \mathcal{K} is a compact set¹. Our objective is to build an estimator \hat{v}_0 that gives us a good estimation of v_0 , and therefore v at any given time t_f using our pre-calculated curves.

Our data is impacted by some random noise. Thus, we define :

$$Y_{t_i} = v(v_0, t_i) + N_i \quad (2)$$

where (N_i) are random variables called “noise”, as pictured in Figure 5.

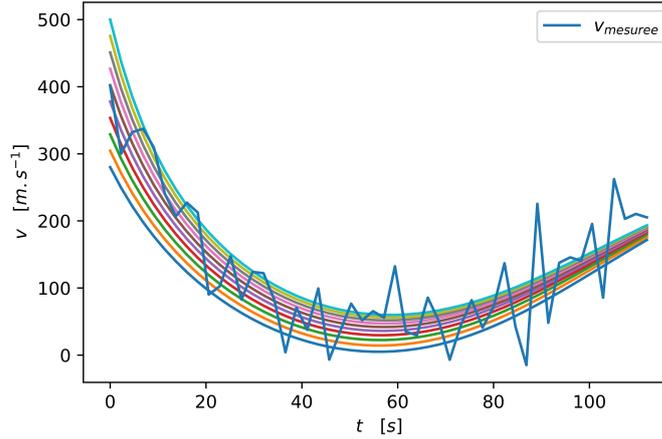


Figure 5: Examples of a list of measurements with a white Gaussian noise.

Our objective is to find the curve among those in the background that most fit our measured velocity (here in top in blue).

2.4 Maximum likelihood Estimator

For a set of measurements (y_0, \dots, y_{t_n}) , we define :

$$L(u_0 | y_{t_0}, \dots, y_{t_n}) : u_0 \rightarrow P_{u_0}((Y_{t_0} = y_{t_0}), (Y_{t_1} = y_{t_1}), \dots, (Y_{t_n} = y_{t_n}))$$

where P_{u_0} is a (joint) probability depending of one parameter u_0 , which corresponds to our initial velocity.

Maximum likelihood estimation (MLE) is a method of estimating this parameter. The value of u_0 that maximizes L is the parameter that defined the statistical model where the observed data is most probable. Therefore, one can imagine using as an estimator for v_0 :

$$\hat{v}_0 = \operatorname{argmax}_{\mathbf{u}_0} L(\mathbf{u}_0 | y_{t_0}, \dots, y_{t_n})$$

2.5 First results

Let us, for the sake of an example, imagine that our noise across our data (N_i) is independent, and identically distributed following a centered Gaussian law $\mathcal{N}(0, \sigma)$. Then, as $N_i = Y_{t_i} - v(v_0, t_i)$, one gets

$$P_{u_0}(Y_{t_i} = y_{t_i}) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(y_{t_i} - v(u_0, t_i))^2}{2\sigma^2}$$

¹in real life, \mathcal{K} can be relatively large : it only implies that we have an idea of the range around the value v_0 that we are looking for, which is always true in a real situation

As those random quantities are independent :

$$L(u_0|y_{t_0}, y_{t_1}, \dots, y_{t_n}) = \prod_{i=0}^n P_{u_0}(Y_{t_i} = y_{t_i})$$

As log is an increasing function, it is equivalent to find the max of L than to find the one of $\log(L)$ which is

$$\log L(u_0|y_{t_0}, y_{t_1}, \dots, y_{t_n}) = \sum_{i=0}^n \log(P_{u_0}(Y_{t_i} = y_{t_i})) = \frac{-1}{2\sigma} \sum_{i=0}^n (y_{t_i} - v(u_0, t_i))^2 + cst$$

Then, we can simply rewrite our estimator as

$$\hat{v}_0 = \operatorname{argmax}_{\mathbf{u}_0} L(\mathbf{u}_0|y_{t_0}, \dots, y_{t_n}) = \operatorname{argmin}_{\mathbf{u}_0} \sum_{i=0}^n (y_{t_i} - v(\mathbf{u}_0, t_i))^2$$

With this particular case, it appears that our MLE is equivalent to the method of *non-linear least squares*.

Indeed, using the signals reported in Figure 5, the following estimation (see Figure 6) are obtained.

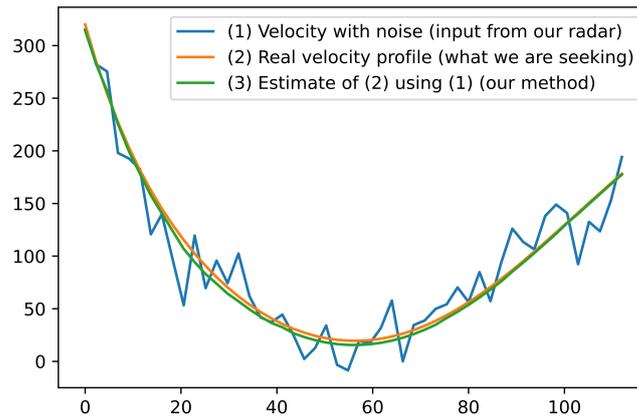


Figure 6: (Causal) Estimation from noisy measurements.

This method is robust and even with extreme value for σ , it converges really fast to the real solution. This comes from the fact that those measured are centered. So what happens when a measurement bias is introduced?

3 Effects of a constant bias

Let $N_i \rightsquigarrow \mathcal{N}(b, \sigma)$ with $b \in \mathbb{R}$. The notation \rightsquigarrow means that the random variable N_i is distributed according to the law $\mathcal{N}(b, \sigma)$, which in this case is a normal distribution of mean b and standard deviation σ). Figure 7 illustrates that, as the differential equation is not linear, a solution shifted up by a constant bias b is not anymore belonging to in the space of the solutions.

We are going to illustrate that with this bias, our previously developed method is exponentially diverging, and thus a new term shall be added to prevent this error.

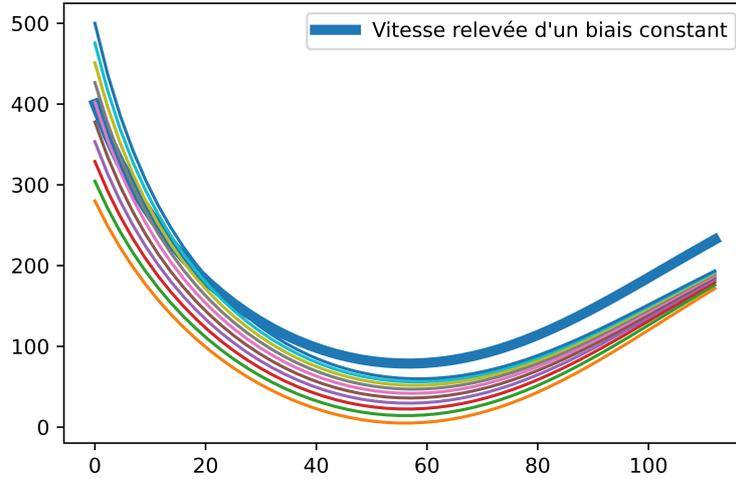


Figure 7: A biased velocity curve is not a trajectory $b = 10m.s^{-1}$

3.1 Contracting behavior

Note the error $\Delta(u_0, t) = v(u_0, t) - v(v_0, t)$. One has

$$\dot{v}(u_0, t) - \dot{v}(v_0, t) = -K(v(u_0, t)^2 - v(v_0, t)^2)$$

$$\dot{\Delta}(u_0, t) = -K\Delta(u_0, t)(v(u_0, t) + v(v_0, t))$$

This scalar differential equation can then be solved

$$\Delta(u_0, t) = (u_0 - v_0) \exp\left(-\int_{t_0}^t K(v(u_0, s) + v(v_0, s))ds\right)$$

$$|\Delta(u_0, t_{i+1})| \leq |\Delta(u_0, t_i)|$$

The solutions are contracting [7], which means that the velocity curves are getting closer and closer among time.

3.2 Impact of a constant bias on the estimator

On the other side :

$$u_0 - v_0 = \Delta(u_0, t) \exp\left(\int_0^t K(v(u_0, s) + v(v_0, s))ds\right)$$

For the sake of an example, let's take some speed around $v \sim 100m.s^{-1}$:

- With an error b at $t \sim 0$, which means that $\Delta(u_0, 0) = b$, then $u_0 - v_0 \sim b \times 1$
- But if there is the same error b at $t \sim 30s$, which means that $\Delta(u_0, 30s) = b$, then $u_0 - v_0 \sim b \exp\left(\int_0^{30} 2 * 100ds\right) \sim 2b$.

As it shows, as a result of the contracting behavior, the bias is exponentially more detrimental to the estimation of v_0 while one is looking at values of v further in time.

However, the nonlinear nature mentioned above of our equation makes it possible to observe this bias and therefore to avoid it. This is explained below.

3.3 Correction of our estimator

In order to take into account this flow, it is possible to add in our estimator this constant bias :

$$L(u_0|y_{t_0}, y_{t_1}, \dots, y_{t_n}) = \sum_{i=0}^n (y_{t_i} - [v(u_0, t_i) + b])^2$$

So how to choose b using our data $(y_{t_0}, y_{t_1}, \dots, y_{t_n})$? Simply, a look at the value of b that minimize the quantity

$$\frac{\partial L}{\partial b} = -2 \sum_{i=0}^n (y_{t_i} - v(u_0, t_i) - b) = 0$$

suggests the value for the estimator

$$\hat{b}(u_0) = \frac{1}{n+1} \sum_{i=0}^n (y_{t_i} - v(u_0, t_i))$$

3.4 A new estimator

All at once, we now use a joint estimator for v and b :

$$\hat{v}_0 = \operatorname{argmin}_{u_0} L_n(u_0)$$

with

$$L_n(u_0) = \frac{1}{n+1} \sum_{i=0}^n (y_{t_i} - [v(u_0, t_i) + \hat{b}(u_0)])^2 \quad (3)$$

and

$$\hat{b}(u_0) = \frac{1}{n+1} \sum_{k=0}^n (y_{t_k} - v(u_0, t_k)) \quad (4)$$

The convergence properties of this new estimator are studied next.

4 Convergence analysis

4.1 Main result

Theorem 1. Consider the differential equation $\dot{v} = -Kv^2 + g(t)$ (E) and its unique solution $v(v_0, t)$ for the initial value v_0 . Consider² $\mathcal{E} = \{v_0 \in \mathbb{R} | \forall t, v(v_0, t) \geq 0\}$. Let $\mathcal{K} \subset \mathcal{E}$ be a compact set. Choose $v_0 \in \mathcal{K}$, $n \in \mathbb{N}^*$, $t_0, t_f \in \mathbb{R}$, and $T_n = \frac{t_f - t_0}{n}$ a step size, $T_1 = t_f - t_0$ the duration of the flight and $t_i = t_0 + i \times T_n$ the instant of the i -th measure. Consider a set of $n+1$ noisy and biased measurements

$$(y_i = v(v_0, t_i) + b + n_i)_{0 \leq i \leq n}$$

where b is an unknown constant and where $(n_i)_{0 \leq i \leq n}$ are $(N_i)_{0 \leq i \leq n}$, where N_i are centered random variables of standard deviation σ and such as $\mathbb{E}(|N_i|)$ exists.

²In real application, any application always leads to an idea of the initial value of v_0 : one is never clueless on the real value of v_0 , and there is at least a wide range where we know that our initial velocity will be in.

With (3) and (4), consider the following estimator for v_0 :

$$\hat{v}_{0,n} = \underset{u_0 \in \mathcal{K}}{\operatorname{argmin}} L_n(u_0) \quad (5)$$

Then the following properties hold :

- $\hat{v}_{0,n}$ converges to v_0 when $n \rightarrow \infty$.
- $v(t_f, \hat{v}_{0,n})$ converges to $v(t_f, v_0)$ when $n \rightarrow \infty$.
- $\hat{b}_n(\hat{v}_{0,n})$ converges to b when $n \rightarrow \infty$.

4.2 Proof

4.2.1 A Lemma from the literature

This proof is based on a Lemma which is stated and proven in [3].

Using our notations, this Lemma can be reformulated as

Main Lemma ([3]). *Consider \mathcal{K} a compact set, and $v_0 \in \mathcal{K}$. Assume that $L_n(u_0)$ is continuous in u_0 and that there exists a function $L(u_0)$ such that :*

- $L_n(u_0)$ converges uniformly in probability to $L(u_0)$
- L is continuous
- L is uniquely minimized in v_0

Then \hat{v}_0 , defined as the value of \mathcal{K} that minimizes L_n for each $(Y_0, \dots, Y_{t_f}) = (y_0, \dots, y_{t_f})$ is converging in probability to v_0 .

In order to introduce the notion of *uniform convergence in probability*, let us first recall the notion of simple convergence in probability :

Definition (Simple convergence in probability).

Let $X_n(u)$ be a sequence of real random variables depending on one parameter u , and $l \in \mathbb{R}$.

It is said that the sequence X_n converges simply in probability to l when :

$$\forall u \in \mathcal{K}, \forall \epsilon, P(|X_n(u) - l| > \epsilon) \rightarrow 0 \quad \text{when } n \rightarrow \infty$$

This definition is analogue to the notion of *simple convergence* while facing a sequence of functions. Similarly, we have to consider the analogue of *uniform convergence*, because the simple convergence in probability will face the same cons as its analogue in \mathbb{R} .

Definition (Uniform convergence in probability).

Let $X_n(u)$ be a sequence of real random variables depending on one parameter u , and $l \in \mathbb{R}$.

It is said that the sequence X_n converges uniformly in probability to l when :

$$\forall \epsilon, P(\sup_{u \in \mathcal{K}} |X_n(u) - l| > \epsilon) \rightarrow 0 \quad \text{when } n \rightarrow \infty$$

Along the following proof, $\sup_{u \in \mathcal{K}} |F(u)|$ will be written $\|F\|_{\infty}^{\mathcal{K}}$.

4.2.2 Analysis of the solutions of the differential equation

Under a smoothness assumption on the coefficient K in the differential equation, any solution $v(u_0, \cdot)$ with $u_0 \in \mathcal{K}$ is \mathcal{C}^∞ . As previously, consider the error $\Delta(u_0, t) = v(u_0, t) - v(v_0, t)$.

It is straightforward to see that $v_0 \mapsto v(v_0, t)$ is *injective*. Indeed Lets choose u_1 and u_2 in \mathcal{K} such that $\exists \tau, v(u_1, \tau) = v(u_2, \tau)$. Using the Cauchy-Lipschitz theorem, there exists a unique solution for a set of initial condition $(\tau, v(u_1, \tau))$. That means $\forall t, v(u_1, t) = v(u_2, t)$, which leads to $u_1 = v(u_1, 0) = v(u_2, 0) = u_2$. Also, it is straightforward to see that $v_0 \mapsto v(v_0, t)$ is strictly monotonic. Indeed, choose $u_1 > u_2$, then $v(u_1, 0) > v(u_2, 0)$. Then $\forall t, v(u_1, t) > v(u_2, t)$. Indeed, assume that one can find τ such as $v(u_1, \tau) \leq v(u_2, \tau)$, then the Intermediate value theorem states that $\tau_0 \in [0, \tau]$ such that $v(u_1, \tau_0) = v(u_2, \tau_0)$, and so $u_1 = u_2$ which yields the conclusion. As a final remark, $v_0 \mapsto \Delta(v_0, t)$ is also strictly increasing.

4.2.3 Rewriting the likelihood function

Some calculations give

$$L_n(u_0) = \frac{1}{n+1} \sum_{i=0}^n (y_i - (v(u_0, t_i) + \hat{b}_n(u_0)))^2 = \frac{1}{n+1} \sum_{i=0}^n (\Gamma_i(u_0) - [\hat{b}_n(u_0) - b])^2$$

with $\Gamma_i(u) = v(v_0, t_i) - v(u, t_i) + N_i = \Delta_i(u) + N_i$. A closer look at $\hat{b}_n(u_0)$ reveals that

$$\hat{b}_n(u_0) = \frac{1}{n+1} \sum_{k=0}^n (y_k - v(u_0, t_k)) = \frac{1}{n+1} \sum_{k=0}^n (v(v_0, t_k) - v(u_0, t_k) + b + n_k) = b + \frac{1}{n+1} \sum_{k=0}^n \Gamma_k(u)$$

Which gives

$$\boxed{\hat{b}_n(u_0) - b = \frac{1}{n+1} \sum_{k=0}^n \Gamma_k(u) \triangleq \bar{\Gamma}_n(u)} \quad (6)$$

By injecting this result into the previous line, using *König-Huygens's* relation, one gets

$$\boxed{L_n(u) = \frac{1}{n+1} \sum_{i=0}^n (\Gamma_i(u) - \bar{\Gamma}_n(u))^2 = \frac{1}{n+1} \sum_{i=0}^n \Gamma_i(u)^2 - (\bar{\Gamma}_n(u))^2} \quad (7)$$

4.2.4 Standard deviation

Analysis of $\Gamma_i(u)$ As we defined it before, $\Gamma_i(u) = \Delta_i(u) + N_i$. The following results follow

- $\Gamma_i(u)$ is a random quantity, $\mathbb{E}(\Gamma_i(u)) \in \mathbb{R}$ and $\text{Var}(\Gamma_i(u)) \in \mathbb{R}$.
- Those random quantities $\Gamma_i(u)$ are independent.
- $\mathbb{E}(\Gamma_i(u)) = \Delta_i(u) + \mathbb{E}(n_i) = \Delta_i(u)$
- $\mathbb{E}(\Gamma_i(u)^2) = \mathbb{E}(\Delta_i(u)^2 + n_i^2 + 2\Delta_i(u)n_i) = \Delta_i(u)^2 + \mathbb{E}(n_i^2) + 0 = \Delta_i(u)^2 + \sigma^2$
 $\text{var}(\Gamma_i(u)) = \text{var}(\Delta_i(u) + n_i) = \text{var}(n_i) = \sigma^2$.

As for $\bar{\Gamma}(u)$:

- $\mathbb{E}(\bar{\Gamma}(u)) = \frac{1}{n+1} \sum_{i=0}^n \Delta_i(u) = \bar{\Delta}(u)$

- $\text{var}(\bar{\Gamma}(u)) = \frac{\sigma^2}{n+1}$.
- $\mathbb{E}(\bar{\Gamma}(u)^2) = \text{var}(\bar{\Gamma}(u)) + \mathbb{E}(\bar{\Gamma}(u))^2 = \frac{\sigma^2}{n+1} + \bar{\Delta}(u)^2$.

Expected value of $L_n(u_0)$

With the previous results in mind, it becomes easy to write the expected value of $L_n(u_0)$:

$$\begin{aligned} \mathbb{E}(L_n(u)) &= \frac{1}{n+1} \sum_{i=0}^n \mathbb{E}(\Gamma_i(u)^2) - \mathbb{E}(\bar{\Gamma}(u)^2) \\ &= \frac{1}{n+1} \sum_{i=0}^n (\Delta_i(u)^2 + \sigma^2) - \frac{\sigma^2}{n+1} - \bar{\Delta}(u)^2 = \frac{1}{n+1} \sum_{i=0}^n \Delta_i(u)^2 - \bar{\Delta}(u)^2 + \frac{n}{n+1} \sigma^2 \end{aligned}$$

$$\mathbb{E}(L_n(u)) = \frac{1}{n+1} \sum_{i=0}^n (\Delta_i(u) - \bar{\Delta}(u))^2 + \frac{n}{n+1} \sigma^2 \quad (8)$$

As a Riemann sum, this quantity converge when $n \rightarrow \infty$, and :

$$\mathbb{E}(L_n(u)) \rightarrow \frac{1}{T_1} \int_0^{t_f} (\Delta(s, u) - \bar{\Delta}(u))^2 ds + \sigma^2 = Q(u_0) \quad (9)$$

Standard deviation of $L_n(u_0)$

Calculating the exact standard deviation of $L_n(u_0)$ would be really time consuming and irrelevant. The only useful result is to prove that $\text{Var}(L_n(u_0)) \rightarrow 0$ when $n \rightarrow \infty$.

To prove this, let us first break down a set of useful proprieties :

1. If $X = O(n^\alpha)$ and $Y = O(n^\beta)$, then $\text{var}(X) = O(n^{2\alpha})$ and $\text{Cov}(X, Y) = O(n^{\alpha+\beta})$.
2. $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{Cov}(X, Y) = O(n^{\max(2\alpha, 2\beta)})$
3. Then, using the previous result, with n quantities which are $O(n^\alpha)$, $\text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i,j=1}^n \text{Cov}(X_i, X_j) = \frac{1}{n^2} \sum_{i,j=1}^n O(n^{2\alpha}) = O(n^{2\alpha})$
4. and, from the independence of the random quantities $O(n^\alpha)$, then $\text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{var}(X_i) = O(n^{2\alpha-1})$

It now becomes pretty easy to find an approximant of $\text{Var}(L_n(u_0))$. By assumption $\text{Var}(N_i^2)$ exists, then $\text{Var}(\Gamma_i^2)$ does exist as well. And as a sum of independent variables,

$$\text{Var}\left(\frac{1}{n+1} \sum_{i=0}^n \Gamma_i^2\right) = \frac{1}{n+1} \text{Var}(\Gamma_0^2) = O\left(\frac{1}{n}\right)$$

The second term rewrites :

$$\left(\frac{1}{n+1} \sum_{i=0}^n \Gamma_i\right)^2 = \frac{1}{(n+1)^2} \sum_{i,j=0}^n \Gamma_i \Gamma_j = \frac{1}{n+1} \left(\sum_{i=0}^n \Gamma_i \left[\frac{1}{n+1} \sum_{j=0, j \neq i}^n \Gamma_j \right] + \frac{1}{n+1} \Gamma_i^2 \right)$$

$$\text{Var}\left(\frac{1}{n+1} \sum_{j=0, j \neq i}^n \Gamma_j\right) = \frac{n}{(n+1)^2} \sigma^2 = O\left(\frac{1}{n}\right)$$

And because this quantity is independent of Γ_i :

$$\text{Var}\left(\Gamma_i \frac{1}{n+1} \sum_{j=0, j \neq i}^n \Gamma_j\right) = \text{Var}(\Gamma_i) \text{Var}\left(\frac{1}{n+1} \sum_{j=0, j \neq i}^n \Gamma_j\right) = \sigma^2 O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right)$$

$\text{Var}\left(\frac{1}{n+1} \Gamma_i^2\right) = O\left(\frac{1}{n^2}\right)$. Using the previous results :

$$\text{Var}\left(\Gamma_i \left[\frac{1}{n+1} \sum_{j=0, j \neq i}^n \Gamma_j\right] + \frac{1}{n+1} \Gamma_i^2\right) = O\left(\frac{1}{n}\right)$$

Finally, adding a sum :

$$\text{Var}\left(\frac{1}{n+1} \sum_{i=0}^n \Gamma_i^2\right) = \text{Var}\left(\frac{1}{n+1} \left(\sum_{i=0}^n \Gamma_i \left[\frac{1}{n+1} \sum_{j=0, j \neq i}^n \Gamma_j\right] + \frac{1}{n+1} \Gamma_i^2\right)\right) = O\left(\frac{1}{n}\right)$$

This leads to the desired result, as $L_n(u) = \frac{1}{n+1} \sum_{i=0}^n \Gamma_i(u)^2 - (\bar{\Gamma}_n(u))^2$:

$$\boxed{\text{var}(L_n(u)) = O\left(\frac{1}{n}\right)} \quad (10)$$

4.2.5 Determining the limit function of L_n

To use the referenced Lemma, it is first mandatory to prove the uniform convergence in probability of L_n to a known function L . As the uniform convergence in probability is stronger than the simple convergence in probability, then the only candidate function L is the simple limit in probability of L_n .

With the previous results, as the standard deviation of L_n is going down to 0 with $n \rightarrow \infty$, and because the sequence of functions $\mathbb{E}(L_n(u))$ converges to a function $Q(u_0)$, one can make the assumption that $\boxed{L(u_0) = Q(u_0)}$.

Let A_n, B_n and C_n be 3 positive random quantities such that $A_n \leq B_n + C_n$. Then $A_n \geq 2\epsilon \Rightarrow B_n + C_n \geq 2\epsilon$. Therefore, the event $(A_n \geq 2\epsilon)$ is included in the event $(B_n + C_n \geq 2\epsilon)$. Thus, $\mathbb{P}(A_n \geq 2\epsilon) \leq \mathbb{P}(B_n + C_n \geq 2\epsilon)$.

Further, if it is possible to decrease C_n to 0 when n goes to infinity, then let us fix a n_0 so than $\forall n \geq n_0, C_n \leq \epsilon$. Then, $\forall n \geq n_0, B_n + C_n \geq 2\epsilon \Rightarrow B_n \geq \epsilon$. Like before, the event $(B_n + C_n \geq 2\epsilon)$ is included in the event $(B_n \geq \epsilon)$, and so $\mathbb{P}(B_n + C_n \geq 2\epsilon) \leq \mathbb{P}(B_n \geq \epsilon)$.

With this, $\mathbb{P}(A_n \geq 2\epsilon) \leq \mathbb{P}(B_n \geq \epsilon)$ when $n \geq n_0$.

Finally, if $\mathbb{P}(B_n \geq \epsilon) \rightarrow 0$ when $n \rightarrow \infty$, then $\mathbb{P}(A_n \geq 2\epsilon) \rightarrow 0$ when $n \rightarrow \infty$

Let us use this previous proof technique with the elements of this paper. Let us fix $A_n = |L_n(u) - Q(u)|$, $B_n = |L_n(u) - \mathbb{E}(L_n(u))|$ and $C_n = |\mathbb{E}(L_n(u)) - Q(u)|$.

By the triangle inequality, one has : $|L_n(u) - Q(u)| \leq |L_n(u) - \mathbb{E}(L_n(u))| + |\mathbb{E}(L_n(u)) - Q(u)|$ which gives $A_n \leq B_n + C_n$.

Furthermore, as $C_n = |\mathbb{E}(L_n(u)) - Q(u)|$, because $\mathbb{E}(L_n(u)) \xrightarrow{n \rightarrow \infty} Q(u)$, then there exists a n_0 so that $\forall n \geq n_0, |\mathbb{E}(L_n(u)) - Q(u)| \leq \epsilon$. And, finally, because $\text{var}(L_n(u)) = O\left(\frac{1}{n^4}\right)$, using

Bienaymé-Tchebychev's inequality, one gets

$$\forall \epsilon > 0, \forall u, \mathbb{P}(|L_n(u) - \mathbb{E}(L_n(u))| > \epsilon) \leq \frac{\text{var}(L_n(u))}{\epsilon^2} \xrightarrow{n \rightarrow \infty} 0 \quad (11)$$

which means $P(B_n \geq \epsilon) \rightarrow 0$ when $n \rightarrow \infty$

As it is shown, this case perfectly fits the theory that was introduced before. As so, it is clear that :

$$\boxed{\forall \epsilon > 0, \forall u, \mathbb{P}(|L_n(u) - Q(u)| > 2\epsilon) \xrightarrow{n \rightarrow \infty} 0} \quad (12)$$

And so $L_n(u)$ simply converge in probability to $Q(u)$, $\forall u \in E$.

From now on, Q is renamed L , so

$$L(u) = \frac{1}{T_1} \int_0^{t_f} (\Delta(s, u) - \bar{\Delta}(u))^2 ds + \sigma^2 \quad (13)$$

4.2.6 Continuity and global minimum

From (13), the continuity of L continuity is pretty straightforward as it stems from the continuity of $u \mapsto \Delta(s, u)$. Further, because $(\Delta(s, u) - \bar{\Delta}(u))^2 \geq 0$ (as a squared real number), then :

$$L(u) = 0 \Rightarrow \forall t \in [0, t_f], \quad \Delta(t, u) - \bar{\Delta}(u) = 0 \quad (14)$$

Then,

$$\forall t \in [0, t_f], \Delta(t, u) = cst, \quad \forall t \in [0, t_f], \dot{\Delta}(t, u) = 0 = -K\Delta(u, t)(v(u, t) + v(v, t)) \quad (15)$$

So that

$$\forall t \in [0, t_f], \Delta(u_0, t) = 0, \quad \Delta(u, 0) = 0 = u - v_0, \quad u = v_0 \quad (16)$$

Then, L has a unique and global minimum in v_0 , which is the value we are seeking.

4.2.7 Lemma of uniform continuity

The last thing there is to show is that L_n uniformly converge to L in probability. At this point, we have only proved the simple convergence. The following Lemma is useful to establish uniform convergence in probability

Lemma 1 ([3]). *Assume that :*

1. L_n converges in probability to L .
2. L is continuous.
3. There exists $C(N_0, \dots, N_n)$ such that :

$$\forall (u_0, \tilde{u}_0) \in \mathcal{K}^2, |L_n(u_0) - L_n(\tilde{u}_0)| \leq C(N_0, \dots, N_n)|u_0 - \tilde{u}_0|$$

4. $C(N_0, \dots, N_n)$ is bounded in probability.

Then, L_n converges uniformly in probability to L .

The proof of this result can be found in [3].

The hypothesis formulated in this Theorem looks like a probabilistic version of *Lipschitz continuity*. Lipschitz continuity is a strong form of uniform continuity for functions. Therefore, it looks logic that this probabilistic version might imply a uniform convergence in probability.

Aside, the definition of “bounded in probability” is :

Definition (boundedness in probability). $C(N_0, \dots, N_n)$ is bounded in probability if for every $\epsilon > 0$ there exists N and η such that :

$$\forall n \geq N, P(|C(N_0, \dots, N_n)| > \eta) < \epsilon$$

4.2.8 Analysis of the error $L_n(u_0) - L_n(\tilde{u}_0)$

As a reminder :

$$L_n(u_0) = \frac{1}{n+1} \sum_{i=0}^n (\Gamma_i(u_0) - \bar{\Gamma}(u_0))^2$$

By subtracting $L_n(u_0)$ and $L_n(\tilde{u}_0)$ and by using the subtracted squares identity $a^2 - b^2 = (a - b)(a + b)$, we get a formula of the form

$$L_n(u_0) - L_n(\tilde{u}_0) = \frac{1}{n+1} \sum_{i=0}^n A_i B_i$$

with :

$$\begin{cases} A_i \triangleq (\Gamma_i(u_0) - \Gamma_i(\tilde{u}_0) - [\bar{\Gamma}(u_0) - \bar{\Gamma}(\tilde{u}_0)]) \\ B_i \triangleq (\Gamma_i(u_0) + \Gamma_i(\tilde{u}_0) - [\bar{\Gamma}(u_0) + \bar{\Gamma}(\tilde{u}_0)]) \end{cases}$$

Analysis of A_i As a reminder $\Gamma_i(u_0) = \Delta(u_0, t_i) + N_i$. Thus, $\Gamma_i(u_0) - \Gamma_i(\tilde{u}_0) = \Delta(u_0, t_i) - \Delta(\tilde{u}_0, t_i)$. This means that A_i is **not** a random variable, but only a function of (u_0, \tilde{u}_0, t) in \mathbb{R} . Yet, $\mathcal{K} \times [0, t_f]$ is a compact set (as a cross product of 2 compact sets). Knowing this, and by the fact that Δ is \mathcal{C}^1 , *Heine-Cantor theorem* [8] gives us :

$$\exists C, \forall t, |\Delta(u_0, t) - \Delta(\tilde{u}_0, t)| \leq C|u_0 - \tilde{u}_0|$$

This leads to :

$$|A_i| \leq 2C|u_0 - \tilde{u}_0|$$

Analysis of B_i This time, $\Gamma_i(u_0) + \Gamma_i(\tilde{u}_0) = 2N_i + \Delta(u_0, t_i) + \Delta(\tilde{u}_0, t_i)$. This leads to :

$$B_i = 2[N_i - \bar{N}] + \Delta(u_0, t_i) + \Delta(\tilde{u}_0, t_i) + \bar{\Delta}(u_0) + \bar{\Delta}(\tilde{u}_0)$$

Again, as $\mathcal{K} \times [0, t_f]$ is a compact set, one get that

$$\exists D, \forall t, \forall u, |\Delta(u, t)| \leq D$$

And so :

$$|B_i| \leq 2|N_i - \bar{N}| + 4D$$

Gathering the terms Using both A_i and B_i expressions in L_n expression, one has :

$$|L_n(u_0) - L_n(\tilde{u}_0)| \leq \frac{1}{n+1} \sum_{i=0}^n 2C|u_0 - \tilde{u}_0|(2|N_i - \bar{N}| + 4D)$$

This means :

$$|L_n(u_0) - L_n(\tilde{u}_0)| \leq |u_0 - \tilde{u}_0| \left[8CD + \frac{4C}{n+1} \sum_{i=0}^n |N_i - \bar{N}| \right]$$

4.2.9 $\frac{1}{n+1} \sum_{i=0}^n |N_i - \bar{N}|$ is bounded in probability

The most intuitive tool to prove this is *Markov's inequality* [1] :

$$P(|X| \geq \eta) \leq \frac{\mathbb{E}(|X|)}{\eta} \quad (17)$$

In our case, and using the linearity and the increasing propriety of the expected value :

$$P\left(\left|\frac{1}{n+1} \sum_{i=0}^n |N_i - \bar{N}|\right| \geq \eta\right) \leq \frac{\mathbb{E}\left(\left|\frac{1}{n+1} \sum_{i=0}^n |N_i - \bar{N}|\right|\right)}{\eta} \leq \frac{1}{\eta} \frac{1}{n+1} \sum_{i=0}^n \mathbb{E}(|N_i - \bar{N}|) \quad (18)$$

$$P\left(\frac{1}{n+1} \sum_{i=0}^n |N_i - \bar{N}| \geq \eta\right) \leq \frac{1}{\eta} \frac{1}{n+1} \sum_{i=0}^n [\mathbb{E}(|N_i|) + \mathbb{E}(|\bar{N}|)]$$

$$P\left(\frac{1}{n+1} \sum_{i=0}^n |N_i - \bar{N}| \geq \eta\right) \leq \frac{1}{\eta} \left[\frac{1}{n+1} \sum_{i=0}^n \mathbb{E}(|N_i|) + \mathbb{E}(|\bar{N}|) \right]$$

$$\mathbb{E}(|\bar{N}|) = \mathbb{E}\left(\left|\sum_{i=0}^n N_i\right|\right) \leq \sum_{i=0}^n \mathbb{E}(|N_i|). \text{ This means :}$$

$$P\left(\frac{1}{n+1} \sum_{i=0}^n |N_i - \bar{N}| \geq \eta\right) \leq \frac{2}{\eta} \frac{1}{n+1} \sum_{i=0}^n \mathbb{E}(|N_i|)$$

And, by assumption, $\mathbb{E}(|N_i|)$ exists :

$$P\left(\frac{1}{n+1} \sum_{i=0}^n |N_i - \bar{N}| \geq \eta\right) \leq \frac{2\mathbb{E}(|N_0|)}{\eta}$$

Some lines of calculus yields the relation we are seeking for :

$$P\left(\frac{4C}{n+1} \sum_{i=0}^n |N_i - \bar{N}| \geq \eta\right) \leq \frac{8C\mathbb{E}(|N_0|)}{\eta}$$

And adding to the constant $8CD$:

$$P\left(|8CD + \frac{4C}{n+1} \sum_{i=0}^n |N_i - \bar{N}|\right| \geq \eta\right) \leq P\left(|8CD + \frac{4C}{n+1} \sum_{i=0}^n |N_i - \bar{N}|\right| \geq \eta + 8CD\right) \leq \frac{8C\mathbb{E}(|N_0|)}{\eta}$$

We finally get :

$$P(|8CD + \frac{4C}{n+1} \sum_{i=0}^n |N_i - \bar{N}|| \geq \eta) \leq \frac{8CE(|N_0|)}{\eta}$$

which is the definition of boundedness in probability :

$$\forall \epsilon > 0, \exists \eta (= \frac{8CE(|N_0|)}{\epsilon}), \exists N(= 0), \forall n \geq N, P(|8CD + \frac{4C}{n+1} \sum_{i=0}^n |N_i - \bar{N}|| \geq \eta) \leq \epsilon$$

4.2.10 Conclusion

The last Lemma allows us to prove that L_n converges uniformly to L when $n \rightarrow \infty$. That was the last piece of the puzzle to use the main Lemma introduced at the beginning of this proof. And so, we have established that

$$\hat{v}_0 \text{ converges in probability to } v_0$$

As it was shown before, the contracting behavior of the solution space implies $|\Delta(t_f, u_0)| \leq |\Delta(0, u_0)|$. This means that $|\Delta(t_f, \hat{v}_0)| \leq |\Delta(0, \hat{u}_0)| \rightarrow 0$ when $n \rightarrow \infty$. Therefore :

$$\hat{v} = v(t_f, \hat{v}_0) \text{ converges in probability to } v(t_f, v_0). \quad (19)$$

5 Results, conclusion and perspectives

Using the proposed method on a simulation with identically distributed noises following a normal law of standard deviation $b = 20m.s^{-1}$ and $\sigma = 20m.s^{-1}$ gives us the curve on Figure 8.

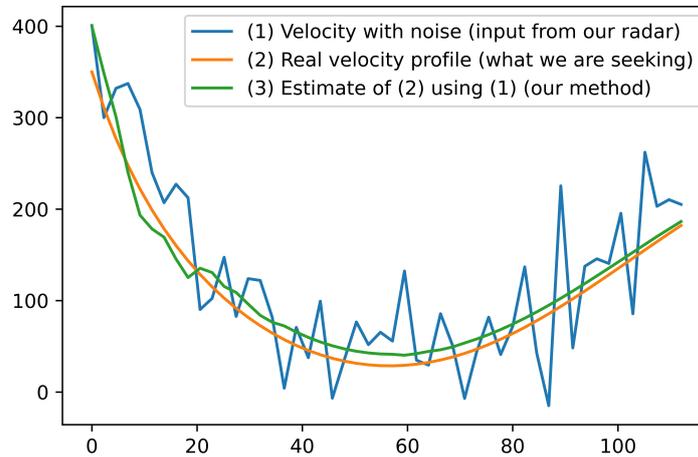


Figure 8: (N_i) i.i.d, et $N_i \rightsquigarrow \mathcal{N}(b = 20m.s^{-1}, \sigma = 20m.s^{-1})$

Even with pretty absurd value for b and σ as shown in Figure 9, our method looks really robust and consistent.

Unfortunately, I did not get the time to push further analysis of the consistency and robustness of this method. However, those first results look really promising and this method gives hope for the best.

Based on this first track of reflections, this work might be improved in the following ways :

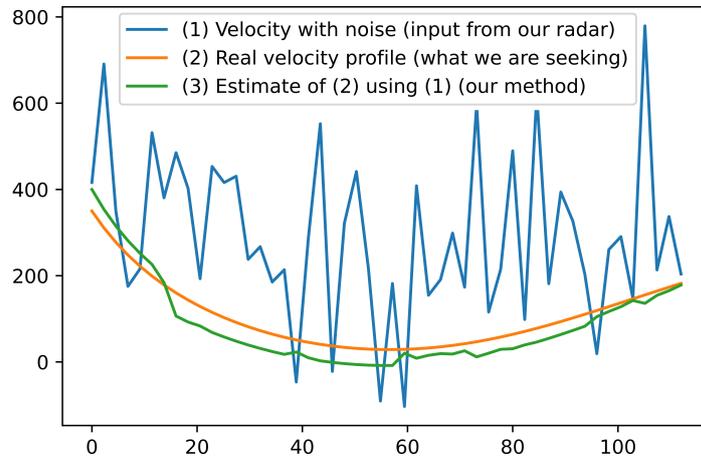


Figure 9: (N_i) i.i.d, et $N_i \rightsquigarrow \mathcal{N}(b = 200m.s^{-1}, \sigma = 150m.s^{-1})$

- An in deep study of the robustness of the detailed tools given
- Implement a real time implementation, probably using a recursive form.
- Developing a gradient descent implementation (instead of an comprehensive test on a set of a fixed number of velocity profile).

To conclude this study, the developed method is able to find using a set of biased measures the likeliest velocity profile on a set of pre-calculated curves. This method gives a smooth result which is a critical step to find the angle of pitch θ and to find the attitude of the shell.

This method appears to be relatively robust even with quite disproportionate noise, but further study of the robustness would be necessary to properly conclude.

6 Acknowledgments

This work would not have been possible without the help of my project supervisor Dr. Nicolas Petit, whom I warmly thank for his time and guidance during these 3 months of initiation to the world of research.

References

- [1] Ph. Barbe and M. Ledoux. *Probabilité*. EDP Sciences, 2007.
- [2] S. Changey, E. Pecheur, and T. Brunner. Attitude estimation of a projectile using magnetometers and accelerometers, experimental validation. In *Position, Location and Navigation Symposium - PLANS 2014, 2014 IEEE/ION*, pages 1168–1173, 2014.
- [3] H. Chernoff, D.B. Rubin, and D.O. Scharfstein. Asymptotic properties of the MLE. <http://www.biostat.jhsph.edu/bstcourse/bio771/>, 2007.

- [4] A. Fiot. *Attitude estimation of an artillery shell in free-flight from accelerometers and magnetometers*. PhD thesis, MINES ParisTech, PSL University, France, 2020.
- [5] A. Fiot, S. Changey, C. Combettes, and N. Petit. Estimation of air velocity for a high velocity spinning projectile using transverse accelerometers. In *Proc. of the AIAA Guidance, Navigation, and Control Conference, AIAA Science and Technology Forum and Exposition*, 2018.
- [6] A. Fiot, S. Changey, C. Combettes, and N. Petit. A gyroless adaptation of attitude complementary filtering. In *Proc. of the European Control Conference (ECC)*, 2019.
- [7] P. Hartman. *Ordinary Differential Equations*. Society for Industrial and Applied Mathematics, second edition, 2002.
- [8] E. Heine and G. Cantor. Heine–Cantor theorem. <https://planetmath.org/heinecantortheorem>, 1871.
- [9] R. L. McCoy. *Modern exterior ballistics*. Schiffer, 2nd edition, 1998.
- [10] G. Wahba. Problem 65-1: a least squares estimate of spacecraft attitude. *SIAM Review*, 7(3):409, 1965.