

Huberian approach for reduced order ARMA modeling of neurodegenerative disorder signal

Christophe Corbier, Mohamed El Badaoui, Hector Manuel Romero Ugalde

► **To cite this version:**

Christophe Corbier, Mohamed El Badaoui, Hector Manuel Romero Ugalde. Huberian approach for reduced order ARMA modeling of neurodegenerative disorder signal. Signal Processing, Elsevier, 2015, 113, pp.273-284. <10.1016/j.sigpro.2015.02.010>. <inserm-01142514>

HAL Id: inserm-01142514

<http://www.hal.inserm.fr/inserm-01142514>

Submitted on 15 Apr 2015

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.

Huberian Approach for Reduced Order ARMA Modeling of Neurodegenerative Disorder Signal

Christophe Corbier^{a,*}, Mohamed El Badaoui^a, Hector Manuel Romero Ugalde^b

^aUniversite de Lyon, F-42023 Saint Etienne, France
 Universite de Saint Etienne, Jean Monnet, F-42000 Saint-Etienne, France
 LASPI, F-42334 IUT de Roanne, France

^bUniversite de Rennes 1, 35042, Rennes, France
 LTSI

Abstract

The purpose of this paper is to address the question of the existence of auto regressive moving average (ARMA) models with reduced order for neurodegenerative disorder signals by using Huberian approach. Since gait rhythm dynamics between Parkinson's disease (PD) or Huntington's disease (HD) and healthy control (CO) differ, and since the stride interval presents great variability, we propose a different ARMA modeling approach based on a Huberian function to assess parameters. Huberian function as a mixture of L_2 and L_1 norms, tuned with a threshold γ from a new curve, is chosen to deal with stride signal disorders. The choice of γ is crucial to ensure a good treatment of NO and allows to reduce the model order. The disorders induce disturbances in the classical estimation methods and increase of the number of parameters of the ARMA model. Here, the use of the Huberian function reduces the number of parameters of the estimated models leading to a *disease transfer function* with low order for PD and HD. Mathematical approach is discussed and experimental results based on a database containing 16 CO, 15 PD, and 19 HD are presented.

Keywords: Reduced order ARMA model, Gait signal, Huberian function, Tuning function, L_1 contribution, Neurodegenerative disease

1. Introduction

This paper introduces a new parametric approach for the estimation problem of the reduced order auto regressive moving average (ROARMA) model of human gait rhythm signal [13]. ARMA system identification is a well-defined problem in several science and engineering areas such as speech signal processing, adaptive filtering, radar Doppler processing or biomechanics. There exists different methods to deal with the ARMA estimation problem. Based on the fractional signal processing approach, Chaudhary et al [11] proposes a fractional least mean square (LMS) algorithm for parameter estimation of Hammerstein nonlinear ARMA system with exogenous noise.

*Corresponding author. Tel.: +(33)477448157
 E-mail address: christophe.corbier@univ-st-etienne.fr

19 This algorithm has still been used in other studies [2] [41] [10]. Another approach uses a two-stage fractional LMS
20 identification algorithm for parameter estimation of controlled ARMA (CARMA) systems [33]. The main idea is
21 to use fractional LMS identification (FLMSI) and two-stage FLMSI (TS-FLMSI) algorithms for CARMA models
22 which are decomposed into a system and noise models. Based on robust estimation, Chakhchoukh [9] introduces a
23 new robust method to estimate the parameters of a Gaussian ARMA model contaminated with outliers [18]. This
24 method makes use of a median and is termed ratio-of-medians estimator (RME). Among the problems of ARMA
25 identification, the model order estimation is crucial. Al-Qawasmi et al [4] propose a new technique for model
26 to estimate order in a general ARMA process based on a rounding approach. Most of the time, these estimation
27 procedures are performed by the implicit assumption that the processes are Gaussian [34]. However, most real
28 world signals are non-Gaussian and different methods such as higher order statistics are used [3] [40]. Moreover,
29 these methods are based on the assumption that the signal does not contain outliers or a low density of outliers
30 less than 1%. A reference paper in a robust estimation framework uses Huberian function for ARMA models [30].
31 This work shows that the Huberian-estimates are closely related to those based on a robust filter, but they have two
32 important advantages: they are consistent and the asymptotic theory is tractable. However, in this analysis, the
33 residuals are computed so the effect of one outlier is limited to the period where it occurs. Moreover, experimental
34 results only focus on the Monte Carlo simulations, not real measurements. A recent paper [45] developed a sys-
35 tematic procedure of statistical inference for the ARMA model with unspecified and heavy-tailed heteroscedastic
36 noises. The authors compare some estimators such that LSE, Huberian function and generalized Huberian func-
37 tion with outliers in a simulated ARMA process. In our framework, the measurements are real and contain natural
38 outliers (NO) due to the neurodegenerative disorders of each disease.

39 Neurodegenerative disorders have a direct consequence on the human behavior by introducing NO in biomechanic
40 time-signals. These points are crucial in the study of neurodegenerative diseases and provide information of the
41 degree of disorder. Here, the Parkinson's disease (PD) and Huntington disease (HD) are studied through the *stride*
42 *time-signal* (STS) of human gait rhythm, corresponding to the time from initial contact of when one foot to the
43 subsequent contact of the same foot [21]. Walking is one of the most fundamental and important activities of
44 human that is strongly related to human health [39]. This is a complex process which we have only recently begun
45 to understand through the study of the interval data in a complete gait cycle [35] [36]. Gait rhythm can also be
46 described in terms of swing and stance intervals corresponding to the time of one foot is in the air and the time
47 of bilateral foot contact, respectively (Fig.A.1). Human locomotion is regulated by the central nervous system
48 (CNS). In the CNS of the human body, motor neurons are the nerve cells that process sensory information and
49 control voluntary muscle movement [37]. Serving as a pivotal part of the human motor system, the basal ganglia
50 process motor impulses originating from the cerebral cortex and the brain stem, and also sends sensory informa-
51 tion through the projecting loops in the CNS [42]. Basal ganglia dysfunction affects motor function and may lead

52 to balance impairment or altered gait rhythm. PD is a chronic and progressive hypokinetic disorder of the CNS
53 induced by basal ganglia dysfunction. HD is a progressive neurodegenerative disorder with autosomal dominant
54 inheritance. Analysis of gait parameters is very useful for a better understanding of the mechanisms of movement
55 disorders, in particular for neurodegenerative diseases.

56 Different approaches exist to analyze gait rhythm time-signals, such as the kinematic aspect [29] [24], Gaussian
57 approach [43] [23], Huberian framework [13], and cyclostationary analysis [28] [44]. Wu and Krishnan [43] de-
58 veloped a framework through Gaussian statistical analysis applied to PD, amyotrophic lateral sclerosis, and gait
59 maturation in children. The main drawback of studies based on the Gaussian framework is the not well treatment
60 of the NO in the time-signal. Indeed, during the 5-min walking period, every time the subjects reached the end of
61 the hallway, they had to turn around, and finally they continued walking. The time-signal stride recorded during
62 these walking turns should be treated as NO. The authors replaced these points by the median value of the stride
63 interval time series, using the *three-sigma rule*, in order to avoid disturbance of the statistical moments. Unfortu-
64 nately, these authors neglected relevant information about the time-signal dynamics, since these NO give capital
65 information during the short phase of the walking turn. These subjects present difficulties to turn and it seems
66 fundamental to consider these points. Therefore, Gaussian-based estimation cannot be applied.

67 Here we propose a reduced order ARMA modeling approach based on a Huberian function to assess parameters
68 and experimental results are performed with STS real measurements of CO, PD and HD. Huberian function is a
69 mixture of L_2 and L_1 norms with a threshold γ . The choice of γ is crucial to ensure a good treatment of NO and
70 allows to reduce the model order. A large section in this paper discusses on the choice of γ using a new curve.
71 A relevant choice of γ in a new interval range ensures both convergence and consistency of the robust estimator.
72 Convergence is shown and a new method to assess the variance/covariance matrix of the estimator is proposed.
73 This paper is organized as follows: Section 2 gives the Huberian mathematical context of the ARMA estimator. Ex-
74 perimental results based on a database containing 16 CO, 15 PD, and 19 HD are shown in Section 3. Conclusions
75 and perspectives are drawn in Section 4.

76 2. Huberian mathematical framework

77 This section presents the Huberian framework mathematical basis. The choice of the threshold in Huber's
78 function is presented and discussed. Asymptotic convergence in law of the robust estimator is shown, consid-
79 ering the stochastic differentiability approach [31] and the m -dependence context. A new method to assess the
80 variance/covariance matrix of the estimator is proposed.

81 *2.1. Huberian function and estimation criterion*

82 Let (S, \mathbf{S}, P) be a probability space and $\{X_k\}_{k=1}^N$ a sequence of i.i.d.r.v's with values in S . Let Θ be a Borel
83 subset in \mathbb{R}^d and Γ a compact subset of \mathbb{R} . Let $\rho_\gamma^H : S \times \Theta \times \Gamma \rightarrow \mathbb{R}$ be a symmetric function such that $\rho_\gamma^H(\bullet, (\theta, \gamma))$
84 is measurable for each $\theta \in \Theta$ and $\gamma \in \Gamma$. The estimator $\hat{\theta}_N^H$ is defined by a minimum of the form

$$N^{-1} \sum_{k=1}^N \rho_\gamma^H(X_k(\hat{\theta}_N^H, \hat{\gamma})) = \inf_{\theta \in \Theta, \gamma \in \Gamma} N^{-1} \sum_{k=1}^N \rho_\gamma^H(X_k(\theta, \gamma)) \quad (1)$$

85 with

$$\rho_\gamma^H(X) = \begin{cases} \frac{X^2}{2} & \text{for } |X| \leq \gamma \\ \gamma|X| - \frac{\gamma^2}{2} & \text{for } |X| > \gamma \end{cases} \quad (2)$$

86 where γ is a threshold to be determined to improve efficiency, convergence, and stability of $\hat{\theta}_N^H$ [22] [12]. Let us
87 introduce two index sets in $\theta \in \mathbb{R}^d$ defined by $v_2(\theta, \gamma) = \{k : |\varepsilon_k(\theta, \gamma)| \leq \gamma\}$ and $v_1(\theta, \gamma) = \{k : |\varepsilon_k(\theta, \gamma)| > \gamma\}$ such
88 that $\text{card}[v_2(\theta, \gamma)] + \text{card}[v_1(\theta, \gamma)] = N \forall \theta \in \mathcal{D}_M, \gamma \in \mathcal{D}_\gamma$, where \mathcal{D}_M and \mathcal{D}_γ are compact subsets and M a
89 model structure. Let $M(\theta)$ be a particular model corresponding to the parameter vector value θ . Let us define
90 $\tilde{\theta} = [\theta \ \gamma]$. Let $W_N(\theta, \gamma)$ be the estimation criterion of the parameter vector θ for a threshold $\gamma > 0$. We denote
91 $s_k(\theta, \gamma)$, $k = 1, \dots, N$ the sign function such that $s_k(\theta, \gamma) = 1$ for $\varepsilon_k(\theta, \gamma) > \gamma$, $s_k(\theta, \gamma) = -1$ for $\varepsilon_k(\theta, \gamma) < -\gamma$ and
92 $s_k(\theta, \gamma) = 0$ for $|\varepsilon_k(\theta, \gamma)| < \gamma$. Let $\varepsilon_k(\theta, \gamma) = y_k - \hat{y}_{k|k-1}(\theta, \gamma) = y_k - \varphi_k^T(\theta, \gamma)\theta$ be the prediction error where y_k is
93 the process output, $\hat{y}_{k|k-1}(\theta, \gamma)$ the prediction model and $\varphi_k(\theta, \gamma) \in \mathbb{R}^d$ the regressor vector. This criterion contains
94 a L_2 part to treat small prediction errors and a L_1 part to deal with NO. Consider a batch of data from the system
95 $\tilde{Z}^N = [y_1 \dots y_N]$. Roughly speaking, we have to determine a mapping from the data \tilde{Z}^N to the set $\mathcal{D}_M \times \mathcal{D}_\gamma$

$$\tilde{Z}^N \longrightarrow \hat{\theta}_N^H = [\hat{\theta}_N^H \ \hat{\gamma}] \in \mathcal{D}_M \times \mathcal{D}_\gamma \quad (3)$$

96 The robust estimation criterion can be written as

$$W_N(\theta, \gamma) = \frac{1}{N} \sum_{k \in v_2(\theta, \gamma)} \frac{\varepsilon_k^2(\theta, \gamma)}{2} + \frac{\gamma}{N} \sum_{k \in v_1(\theta, \gamma)} \left(|\varepsilon_k(\theta, \gamma)| - \frac{\gamma s_k^2(\theta, \gamma)}{2} \right) \quad (4)$$

97 Let us denote $\|X\|^2 = \sum_i x_i^2$ and $|X| = \sum_i |x_i|$ where $X = [x_1 \dots x_N]^T$. We define the following rule: $x_{v_i, k} = x_k$ for all
98 $k \in v_i(\theta, \gamma)$ and $x_{v_i, k} = 0$ otherwise. We define the sparse matrix in $\mathbb{R}^{N \times d}$ over $v_i(\theta, \gamma)$ ($i = 1, 2$) respectively given
99 by

$$\Phi_{v_i}(\theta, \gamma) = \begin{bmatrix} \varphi_{v_i, 1}^T(\theta, \gamma) \\ \dots \\ \varphi_{v_i, N}^T(\theta, \gamma) \end{bmatrix}, \quad \varphi_{v_i, k}(\theta, \gamma) = \begin{cases} \varphi_k(\theta, \gamma) & \text{for } k \in v_i(\theta, \gamma) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

100 On the other hand, we define $Y_{v_i} = [y_{v_i,1} \dots y_{v_i,N}]^T$ the process output vector and $S_{v_1} = [s_{v_1,1} \dots s_{v_1,N}]^T$ the sign vector.
 101 The estimation criterion to be minimized is then given by

$$W_N(\theta, \gamma) = \frac{1}{2N} \|Y_{v_2} - \Phi_{v_2}(\theta, \gamma)\theta\|^2 + \frac{1}{N} \left[\gamma |Y_{v_1} - \Phi_{v_1}(\theta, \gamma)\theta| - \frac{\gamma^2}{2} \|S_{v_1}\|^2 \right] \quad (6)$$

102 This minimization algorithm is applied to yield a minimum corresponding to a given robust estimator for an
 103 appropriated choice of the threshold γ . In the sequel, we show this choice from two joint approaches. The first
 104 one comes from the maximum of the bias by defining a new function with properties to reduce the effect of NO
 105 in prediction errors. A new curve is presented and locates a new investigation interval of γ . From this, the second
 106 approach is to seek a local or global minimum of the robust estimation criterion with respect to θ and γ .

107 2.2. ARMA model in Huber's framework

108 The process output data are denoted as $\delta t_k, k = 1 \dots N$ corresponding to the STS of human gait rhythm. Figure
 109 A.1 shows an example of the left gait signal from heel toe force sensors underneath the left foot where appear the
 110 different phases. Now assuming that δt_k is generated according to

$$\delta t_k = H_0(q) e_k \quad (7)$$

111 where $H_0(q)$ is the noise filter and $e_k, k = 1 \dots N$ a random variables sequence with zero mean and variances λ . The
 112 ARMA model set is parametrized by a d -dimensional real-valued parameter vector θ , i.e.,

$$\delta t_k = H(q, \theta) e_k = \frac{C(q, \theta)}{\mathcal{A}(q, \theta)} e_k \quad (8)$$

113 with $\mathcal{A}(q, \theta) = 1 + \sum_{i=1}^{n_A} a_i q^{-i}$, $C(q, \theta) = 1 + \sum_{i=1}^{n_C} c_i q^{-i}$ and $\theta = [a_1 \dots a_{n_A} c_1 \dots c_{n_C}]^T$. Moreover, q^{-1} is the lag operator
 114 such that $q^{-l} \delta t_k = \delta t_{k-l}, l \in \mathbb{N}$.

115 In Huber's framework, the prediction errors depends on θ and γ . We write $\varepsilon_k(\theta, \gamma) = \delta t_k - \hat{\delta t}_k(\theta, \gamma)$ where
 116 $\hat{\delta t}_k(\theta, \gamma) = \varphi_k^T(\theta, \gamma)\theta$ is the prediction model. The regressor is $\varphi_k(\theta, \gamma) = [-\delta t_{k-1} \dots -\delta t_{k-n_A} \varepsilon_{k-1}(\theta, \gamma) \dots \varepsilon_{k-n_C}(\theta, \gamma)]^T$
 117 and $\psi_k(\theta, \gamma)$ is the gradient with respect to θ of $\hat{\delta t}_k(\theta, \gamma)$ given by $\psi_k(\theta, \gamma) = \frac{1}{C(q, \theta)} \varphi_k(\theta, \gamma)$, meaning that $\psi_k(\theta, \gamma)$
 118 is obtained by filtering the vector $\varphi_k(\theta, \gamma)$ through a stable linear filter.

119 2.3. Choice of γ

120 2.3.1. Location of γ

121 In the prediction error procedure, there appears an inner feedback loop to compute the pseudolinear prediction
 122 model $\hat{y}_{k|k-1}(\theta, \gamma)$. The estimated residuals are treated by a parametric adaptive algorithm which includes $W_N(\theta, \gamma)$

123 to be minimized. The presence of NO in the process output y_k induces large values in $\varepsilon_k(\theta, \gamma)$. A convenient choice
 124 of γ improves the robustness by reducing the effects of these large deviations. In the literature, γ is chosen in the
 125 interval range $[1, 2]$ for linear models. However, this choice does not ensure convergence, consistency nor stability
 126 of $\hat{\theta}_N^H$. Accordingly, the probability density function (pdf) of $\varepsilon_k(\theta, \gamma)$ is strongly disturbed and presents heavy tails.
 127 It is shown that *Huber's* estimators are not always robust and efficient when $\gamma \in [1, 2]$. In a recent paper [14]
 128 on piezoelectric-systems, the use of small values of γ in $[0.01, 0.5]$ led to derive relevant output error models. In
 129 this work, even though the prediction errors were disturbed by numerous NO, the choice of the small values of γ
 130 around 0.05 allowed to obtain interesting results in the frequency interval range for the vibration drilling control.
 131 In the sequel, we introduce a new curve ensuring a reduction of the bias and we show the choice of γ in low values.
 132 In [12] (chapter 6, p.130), we studied the quality of the robustness through influence function [19] of the robust
 133 estimator. We showed that the upper bound of the bias is proportional to the high NO, denoted \mathcal{L}^p and a new
 134 function named *tuning function*, denoted $f^\omega(\gamma)$. Figure A.2 shows this curve. It appears the *classical interval*,
 135 denoted C_γ where $\gamma \in [1, 1.5]$ and a new interval, named *extended interval*, denoted E_γ where $\gamma \in [0.001, 0.2]$. We
 136 showed that

$$\sup_{F_N \in \mathcal{P}_{\Phi_N}(\omega)} |\hat{\theta}_N^H - \theta^*| = b_N^\omega(k) \leq \hat{\kappa}^N f^\omega(\gamma) |\mathcal{L}^p| \quad (9)$$

137 where κ^N is independent of γ , θ^* is the true parameter, $\mathcal{P}_{\Phi_N}(\omega)$ is the corrupted distribution model and F_N the
 138 contaminated Gaussian. An approximation can be written as $f^\omega(\gamma) \approx 0.034\gamma^5 - 0.316\gamma^4 + 1.113\gamma^3 - 1.773\gamma^2 +$
 139 $1.088\gamma - 0.002$. From a linearization of $f^\omega(\gamma)$ in C_γ and E_γ , in absolute value, the slope in E_γ is six times as
 140 important as that of the slope in C_γ . Accordingly, the sensitivity to reduce the influence of high NO in E_γ is six
 141 times as important. Therefore, this new curve allows to locate a new investigation interval of γ in low values in
 142 order to get low values of $f^\omega(\gamma)$ to decrease the effects of NO.

143 2.3.2. Convergence domain of γ

144 Consider the differential of $W_N(\theta, \gamma)$ with respect to θ and γ given by

$$dW_N(\theta, \gamma) = \partial_\theta W_N(\theta, \gamma) d\theta + \partial_\gamma W_N(\theta, \gamma) d\gamma \quad (10)$$

145 where ∂_X is the derivative with respect to X . In detail

$$\partial_\theta W_N(\theta, \gamma) = \frac{-1}{N} \sum_{k \in \mathcal{V}_2(\theta, \gamma)} \psi_k(\theta, \gamma) \varepsilon_k(\theta, \gamma) - \frac{\gamma}{N} \sum_{k \in \mathcal{V}_1(\theta, \gamma)} \psi_k(\theta, \gamma) s_k(\theta, \gamma) \quad (11)$$

146 with $\psi_k(\theta, \gamma) = -\partial_\theta \varepsilon_k(\theta, \gamma)$ and

$$\partial_\gamma W_N(\theta, \gamma) = \frac{1}{N} \sum_{k \in \nu_2(\theta, \gamma)} \phi_k(\theta, \gamma) \varepsilon_k(\theta, \gamma) + \frac{1}{N} \sum_{k \in \nu_1(\theta, \gamma)} \left(|\varepsilon_k(\theta, \gamma)| - \gamma s_k^2(\theta, \gamma) + \gamma \phi_k(\theta, \gamma) s_k(\theta, \gamma) - \frac{\gamma^2}{2} \phi_k^*(\theta, \gamma) s_k(\theta, \gamma) \right) \quad (12)$$

147 with $\phi_k(\theta, \gamma) = \partial_\gamma \varepsilon_k(\theta, \gamma)$ and $\phi_k^*(\theta, \gamma) = \partial_\gamma s_k(\theta, \gamma)$. Let us define $\tilde{\Psi}(\theta, \gamma) = \frac{dW_N(\theta, \gamma)}{d\theta} = [\Psi(\theta, \gamma) \quad \partial_\gamma W_N(\theta, \gamma)]^T$,

148 where $\tilde{\Psi}(\theta, \gamma) \in \mathbb{R}^{d+1}$ and $\Psi(\theta, \gamma) = \partial_\theta W_N(\theta, \gamma)$ named Ψ -function.

149 We seek an optimal value of γ such that $W_N(\theta, \gamma)$ presents a global minimum with probability one (w.p.1) as N

150 tends to infinity, denoted $\overline{W}(\theta, \gamma) = \lim_{N \rightarrow \infty} E W_N(\theta, \gamma)$. This involves that the solution of $\tilde{\Psi}(\hat{\theta}_N^H, \hat{\gamma}) = 0$ is unique.

151 However, it may happen that $\overline{W}(\theta, \gamma)$ does not have a unique global minimum, then we define two compact subsets

152 \mathcal{D}_c^θ and \mathcal{D}_c^γ such that $\hat{\theta}_N^H \rightarrow \mathcal{D}_c^\theta$ w.p.1 as $N \rightarrow \infty$ and $\hat{\gamma} \rightarrow \mathcal{D}_c^\gamma$. We then have

$$\hat{\theta}_N^H = [\hat{\theta}_N^H \quad \hat{\gamma}] \rightarrow \mathcal{D}_c^\theta \times \mathcal{D}_c^\gamma \text{ w.p.1 as } N \rightarrow \infty \quad (13)$$

153 If we denote $\mathcal{D}_c^{\theta\gamma} = \mathcal{D}_c^\theta \times \mathcal{D}_c^\gamma$ then

$$\mathcal{D}_c^{\theta\gamma} = \operatorname{argmin}_{\theta \in \mathcal{D}_M, \gamma \in \mathcal{D}_\gamma} \overline{W}(\theta, \gamma) = \left\{ \theta \in \mathcal{D}_M, \gamma \in \mathcal{D}_\gamma \mid \overline{W}(\theta, \gamma) = \min_{\theta' \in \mathcal{D}_M, \gamma' \in \mathcal{D}_\gamma} \overline{W}(\theta', \gamma') \right\} \quad (14)$$

154 **theorem 1.** Consider a uniformly stable, linear model structure M . Assume that the data set $\tilde{Z}^\infty = \lim_{N \rightarrow \infty} \tilde{Z}^N$, then

$$\sup_{\theta \in \mathcal{D}_M, \gamma \in \mathcal{D}_\gamma} |W_N(\theta, \gamma) - \overline{W}(\theta, \gamma)| \rightarrow 0 \Rightarrow \inf_{\tilde{\theta}^* \in \mathcal{D}_c^{\theta\gamma}} |\hat{\theta}_N^H - \tilde{\theta}^*| \rightarrow 0 \text{ w.p.1 as } N \rightarrow \infty, \tilde{\theta}^* = [\theta^* \quad \gamma^*] \quad (15)$$

155 See proof in ([12], chap.4 p.69). In the case where the condition $\tilde{\Psi}(\hat{\theta}_N^H, \hat{\gamma}) = 0$ does not present a unique solution,

156 there exists a convergence domain of $\hat{\gamma}$ involving a local minimum of $\hat{\theta}_N^H$ such that $\hat{\gamma} \rightarrow \gamma^*$ and $\hat{\theta}_N^H \rightarrow \theta^*$ w.p.1 as

157 $N \rightarrow \infty$. Using theorem 1 and $\inf_{\tilde{\theta}^* \in \mathcal{D}_c^{\theta\gamma}} |\hat{\theta}_N^H - \tilde{\theta}^*| \rightarrow 0$ w.p.1 as $N \rightarrow \infty$, the consistency of the robust estimator is

158 proved.

159 Main properties of the robust estimator related to the covariance matrix and asymptotic normality of $\sqrt{N}(\hat{\theta}_N^H - \theta^*)$

160 are given. In the sequel we assume that $\hat{\gamma}$ converges to γ^* satisfying the conditions of theorem 1. Hence we suppose

161 that the set $\mathcal{D}_c^{\theta\gamma}$ consists only one point $\tilde{\theta}^* = [\theta^* \quad \gamma^*]$. We shall work with the expression $W_N(\theta, \gamma^*)$, $\theta \in \mathcal{D}_M$ and

162 the derivatives will be carried out with respect to θ and will be denoted $\partial_\theta W_N(\theta, \gamma^*)$ and $\partial_{\theta\theta}^2 W_N(\theta, \gamma^*)$ for the first

163 and second derivatives respectively.

164 2.4. ML robust estimator

165 The robust estimator $\hat{\theta}_N^H$ is a maximum likelihood estimator (MLE) satisfying $\rho_\gamma^H(X, \gamma) \sim -\log f_H(X, \gamma)$ where
 166 $f_H(X, \gamma)$ is the pdf defined by

$$f_H(X, \gamma) = \begin{cases} f_{L_2}(X, \gamma) = C(\gamma) e^{-\frac{X^2}{2\phi^2}} & \text{for } |X| \leq \gamma \\ f_{L_1}(X, \gamma) = C(\gamma) e^{-\frac{\gamma|X|}{\phi^2} + \frac{\gamma^2}{2\phi^2}} & \text{for } |X| > \gamma \end{cases} \quad (16)$$

167 $C(\gamma) = \frac{1}{2(K_1(\gamma) + K_2(\gamma))}$ with

$$\begin{cases} K_1(\gamma) = e^{\frac{\gamma^2}{2\phi^2}} \frac{\phi^2}{\gamma} \Gamma\left(1, \frac{\gamma^2}{\phi^2}\right) & \text{for } |X| > \gamma \\ K_2(\gamma) = \frac{\phi}{\sqrt{2}} \left[\Gamma\left(\frac{1}{2}\right) - \Gamma\left(\frac{1}{2}, \frac{\gamma^2}{2\phi^2}\right) \right] & \text{for } |X| \leq \gamma \end{cases} \quad (17)$$

168 $\Gamma(a)$ and $\Gamma(a, X)$ are respectively the complete and incomplete Euler's gamma functions. The parameter ϕ is the
 169 standard deviation of f_H and we can verify that $\forall X \in \mathbb{R}, f_H(X, \gamma) \geq 0$ and $\int_{\mathbb{R}} f_H(X, \gamma) dX = 1$, which ensure that
 170 f_H is a pdf.

171 2.5. Asymptotic covariance matrix of $\hat{\theta}_N^H$ in ARMA model

172 Since $\hat{\theta}_N^H$ minimizes $W_N(\theta, \gamma^*)$ then $\partial_\theta W_N(\hat{\theta}_N^H, \gamma^*) = 0$. Expanding this expression into Taylor's series around
 173 θ^* gives

$$\hat{\theta}_N^H - \theta^* = - \left[\overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*) \right]^{-1} \partial_\theta W_N(\theta^*, \gamma^*) \quad (18)$$

174 where $\partial_\theta W_N(\theta^*, \gamma^*)$ is given by (11) and $\overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*) = \lim_{N \rightarrow \infty} E \partial_{\theta\theta}^2 W_N(\hat{\theta}_N^H, \gamma^*)$ is the symmetric non-negative definite
 175 $d \times d$ limit Hessian matrix with

$$\partial_{\theta\theta}^2 W_N(\theta, \gamma^*) = \frac{-1}{N} \sum_{k \in \nu_2(\theta, \gamma^*)} (\partial_\theta \psi_k^T(\theta, \gamma^*) \varepsilon_k(\theta, \gamma^*) - \psi_k(\theta, \gamma^*) \psi_k^T(\theta, \gamma^*)) - \frac{\gamma}{N} \sum_{k \in \nu_1(\theta, \gamma^*)} \partial_\theta \psi_k^T(\theta, \gamma^*) s_k(\theta, \gamma^*) \quad (19)$$

176 See proof in ([12], chap.4 p.63). From (18) and for N sufficiently large, the asymptotic covariance matrix of the
 177 robust estimator is given by

$$\text{cov}(\hat{\theta}_N^H) \sim \frac{\left[\overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*) \right]^{-1} Q(\theta^*, \gamma^*) \left[\overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*) \right]^{-1}}{N} = \frac{\mathcal{P}(\theta^*, \gamma^*)}{N} \quad (20)$$

178 where $Q(\theta^*, \gamma^*) = \lim_{N \rightarrow \infty} NE \partial_\theta W_N(\theta^*, \gamma^*) \partial_\theta W_N(\theta^*, \gamma^*)^T$ is named Q-matrix.

179 **Remark**

180 For the user, having processed N data points and determined $\hat{\theta}_N^H$ and γ^* , we may use

$$\text{cov}(\hat{\theta}_N^H) = \frac{\left[\partial_{\theta\theta}^2 W_N(\hat{\theta}_N^H, \gamma^*) \right]^{-1} Q(\hat{\theta}_N^H, \gamma^*) \left[\partial_{\theta\theta}^2 W_N(\hat{\theta}_N^H, \gamma^*) \right]^{-1}}{N} \quad (21)$$

181 as an estimate of $\frac{\mathcal{P}(\theta^*, \gamma^*)}{N}$.

182 ARMA models involve a pseudolinear prediction model in $\hat{\delta}_k(\theta, \gamma)$. On the other hand $\psi_k(\theta, \gamma) = \frac{1}{C(q, \theta)} \varphi_k(\theta, \gamma)$
 183 meaning that the matrix $\partial_\theta \psi_k^T(\theta, \gamma^*)$ in (19) is not equal to zero. The main drawback is the infinite sum of Taylor's
 184 expansion of $\psi_k(\theta, \gamma^*)$ and $\partial_\theta \psi_k^T(\theta, \gamma^*)$, increasing the computational cost of the estimated covariance matrix (21).
 185 Here, we show the main results of our method to limit Taylor's expansion with a large order. For more details see
 186 ([12], chap.5 p.74) . After straightforward calculations, we have

$$\psi_k(\hat{\theta}_N^H, \gamma^*) = \sum_{m=0}^{\infty} A_m^N \varphi_{k-m}(\hat{\theta}_N^H, \gamma^*), \quad A_m^N \leq 1 \quad (22)$$

187 with $A_m^N \approx -2 \sum_{k=1}^{\mathcal{F}(n_C/2)} \tilde{\mu}_k \rho_k^{m-1} \cos(\Omega_k^m)$, where $\Omega_k^m = \tilde{\theta}_k + (m-1)\tilde{\varphi}_k$ if n_C is an even number and $\Omega_k^m = l\pi$, $l =$
 188 $\{m, m-1, 1, 0\}$ if n_C is an odd number. $\mathcal{F}(n)$ is the nearest integer less than or equal to n . The coefficients
 189 $\tilde{\mu}_k$, ρ_k , $\tilde{\theta}_k$, $\tilde{\varphi}_k$ are given by the n_C -poles $\{\pi_k\}_{k=1}^{n_C} = \rho_k e^{j\tilde{\varphi}_k}$, where $\rho_k < 1$ for $k = 1 \dots n_C$ and k -th residue
 190 $\text{Res}(\tilde{\Phi}; \pi_k) = \tilde{\mu}_k e^{j\tilde{\theta}_k}$ of the transfer function

$$\tilde{\Phi}(e^{j\omega}, \theta) = 1 - \frac{1}{C(e^{j\omega}, \theta)} = \frac{c_1 e^{j\omega(n_C-1)} + \dots + c_{n_C}}{e^{j\omega n_C} + c_1 e^{j\omega(n_C-1)} + \dots + c_{n_C}} \quad (23)$$

191 We show that A_m^N decrease like $\xi_2(m) = \frac{\beta_1}{m^2} + \frac{\beta_2}{m^4}$ for $m \geq 1$ where β_1, β_2 are determined with well chosen values of
 192 m . We define the *large order* \mathcal{L} to limit the development of (22) by the condition $\xi_2(\mathcal{L}) = \tau$ where τ is a threshold
 193 corresponding to 1% of $\max(A_m^N)$. The large order is then given by

$$\mathcal{L} = \mathcal{F} \left[\sqrt{\frac{1}{2\tau} \left(\sqrt{(\beta_1^N)^2 + 4\beta_2^N \tau} + \beta_1^N \right)} \right] \quad (24)$$

194 Moreover we show that $\sup_k \left\| \psi_k(\hat{\theta}_N^H, \gamma^*) - \psi_k^{\mathcal{L}}(\hat{\theta}_N^H, \gamma^*) \right\| \leq \frac{C}{(\mathcal{L})^2}$ meaning that the bias decreases like $\frac{1}{\mathcal{L}^2}$, ensuring a
 195 good convergence of ψ_k . The limited expression of $\psi_k(\hat{\theta}_N^H, \gamma^*)$ is then yielded by

$$\psi_k^{\mathcal{L}}(\hat{\theta}_N^H, \gamma^*) = \sum_{m=0}^{\mathcal{L}} A_m^N \varphi_{k-m}(\hat{\theta}_N^H, \gamma^*) \quad (25)$$

196 Analogous approach can be made for $\partial_\theta \psi_k^T(\hat{\theta}_N^H, \gamma^*)$. Indeed, its limited Taylor's development has the same large
 197 order \mathcal{L} and we show that $\sup_k \left\| \partial_\theta \psi_k(\hat{\theta}_N^H, \gamma^*)^T - \partial_\theta \psi_k^{\mathcal{L}}(\hat{\theta}_N^H, \gamma^*)^T \right\|_{\infty} \leq \frac{C}{\mathcal{L}^2}$. We then get

$$\partial_\theta \psi_k^{\mathcal{L}}(\hat{\theta}_N^H, \gamma^*)^T = C_k(\hat{\theta}_N^H, \gamma^*) + C_k^T(\hat{\theta}_N^H, \gamma^*) \quad (26)$$

198 where the matrix $C_k(\hat{\theta}_N^H, \gamma^*) \in \mathbb{R}^{d \times d}$ is

$$C_k(\hat{\theta}_N^H, \gamma^*) = \begin{pmatrix} O_{n_A \times d} \\ \hline - \sum_{m=0}^{\mathcal{L}} \sum_{l=0}^{\mathcal{L}} A_m^N A_l^N \varphi_{k-1-m-l}^T(\hat{\theta}_N^H, \gamma^*) \\ \dots \\ \dots \\ - \sum_{m=0}^{\mathcal{L}} \sum_{l=0}^{\mathcal{L}} A_m^N A_l^N \varphi_{k-n_C-m-l}^T(\hat{\theta}_N^H, \gamma^*) \end{pmatrix} \quad (27)$$

199 In the following section, proof of the asymptotic convergence in law of $\sqrt{N}(\hat{\theta}_N^H - \theta^*)$ is considered. This requires
200 the stochastic differentiability and m -dependence approaches.

201 2.6. Asymptotic convergence in law

202 For the asymptotic convergence in law of $\sqrt{N}(\hat{\theta}_N^H - \theta^*)$, let us consider the following technical points related
203 to the signal models of $\varepsilon_k(\theta^*, \gamma^*)$ and $\psi_k(\theta^*, \gamma^*)$.

204 2.6.1. Signal models

205 Assume $\tilde{Z}^\infty = \lim_{N \rightarrow \infty} \tilde{Z}^N$ the data set and consider $(\Omega_j(\theta^*, \gamma^*))_{j \in \nu_1(\theta^*, \gamma^*)}$, $(\phi_j(\theta^*, \gamma^*))_{j \in \nu_1(\theta^*, \gamma^*)}$ the NO in $\varepsilon_k(\theta^*, \gamma^*)$
206 and $\psi_k(\theta^*, \gamma^*)$ respectively. We can write

$$\varepsilon_k(\theta^*, \gamma^*) = \underbrace{\sum_{m \geq 0} \beta_{k,m}(\theta^*, \gamma^*) e_{k-m}}_{k \in \nu_2(\theta^*, \gamma^*)} + \underbrace{\sum_j \Omega_j(\theta^*, \gamma^*) \delta_{k,j}}_{k \in \nu_1(\theta^*, \gamma^*)} \quad (28)$$

$$\psi_k(\theta^*, \gamma^*) = \underbrace{\sum_{m \geq 0} \alpha_{k,m}(\theta^*, \gamma^*) e_{k-m}}_{k \in \nu_2(\theta^*, \gamma^*)} + \underbrace{\sum_j \phi_j(\theta^*, \gamma^*) \delta_{k,j}}_{k \in \nu_1(\theta^*, \gamma^*)} \quad (29)$$

208 for some filters

$$\{\alpha_{k,m}(\theta^*, \gamma^*), \beta_{k,m}(\theta^*, \gamma^*)\} = f_{k,m}(\theta^*, \gamma^*)$$

209 Here $\delta_{t,j}$ is the Kronecker function and

210 **H1:**

- 211 1. $\{e_k\}$ is a sequence of independent rv's with zero mean values and bounded moments of order $4 + \delta$, for $\delta > 0$.
- 212 2. The family of filters $f_{k,m}(\theta^*, \gamma^*)$, $k = 1, 2, \dots$ is uniformly stable for all k, θ^*, γ^* with $f_{k,m}(\theta^*, \gamma^*) < \mu_m$ and

$$213 \sum_{m \geq 0} \mu_m < \infty.$$

214 3. Natural outliers $\Omega_j(\theta^*, \gamma^*)$ and $\phi_j(\theta^*, \gamma^*)$ are bounded for all θ^* , γ^* and j , $\sup_{j, \theta^*, \gamma^*} |\Omega_j(\theta^*, \gamma^*)| = \hat{\Omega}$ and
 215 $\sup_{j, \theta^*, \gamma^*} |\phi_j(\theta^*, \gamma^*)| = \hat{\phi}$.

216 2.6.2. Stochastic differentiability

217 In the literature, the standard asymptotic normality results for MLE requires that (4) be twice continuously
 218 differentiable, which is not the case here by the presence of the sign function. There exists, however, asymptotic
 219 normality results for non-smooth functions and we will hereafter use the one proposed by Newey and McFadden
 220 [31] and Andrews [6]. The basic insight of their approaches is that the smoothness condition of (4), $W_N(\theta, \gamma)$
 221 can be replaced by a smoothness of its limit, which in the standard maximum likelihood case corresponds to the
 222 expectation $-\overline{E} \ln f_H(\varepsilon_k(\theta, \gamma)) = \overline{W}(\theta, \gamma)$, with the requirement that certain remainder terms are small. Hence,
 223 the standard differentiability assumption is replaced by a *stochastic differentiability* condition, which can then be
 224 used to show that the MLE $\hat{\theta}_N^H$ is asymptotically normal. Recall that the derivative w.r.to θ of ρ_γ^H is $\Psi_k(\theta, \gamma)$. If
 225 this function is differentiable in θ , one can establish the asymptotic normality of $\hat{\theta}_N^H$ by expanding $\sqrt{N}(\hat{\theta}_N^H - \theta^*)$
 226 about θ^* using element by element mean value expansions. This is the standard way of establishing asymptotic
 227 normality of the estimator. In a variety of applications, however, $\Psi_k(\theta, \gamma)$ is not differentiable in θ , or not even
 228 continuous, due to the appearance of a sign function. In such a case, one can still establish asymptotic normality
 229 of the estimator provided $\overline{E}\Psi_k(\theta, \gamma)$ is differentiable in θ . Since the expectation operator is a smoothing operator,
 230 $\overline{E}\Psi_k(\theta, \gamma)$ is often differentiable in θ , even though $\Psi_k(\theta, \gamma)$ is not.

231 2.6.3. m -dependence

232 Let us consider m a non-negative interger, then a sequence X_ν of random variables is m -dependent if X_1, X_2, \dots, X_s
 233 is independent of X_k, X_{k+1}, \dots provided $k - s > m$ [32] [38]. Here, this approach is applied since the terms in
 234 $\partial_\theta W_N(\theta, \gamma)$ are not independent. The purpose is to split the sum in (11) into one part that satisfies a certain in-
 235 dependence condition (m -dependence) among its terms and one part that is small. With assumptions H1, the
 236 dependence between distant terms will decrease. Thus, let us consider two following lemmas

237 Lemma 1

238 Consider the sum of doubly indexed rv's $\{x_{k,N}\}$ such that $S_N = \sum_{k=1}^N x_{k,N}$, where $E x_{k,N} = 0$ and $\{x_{1,N}, \dots, x_{s,N}\}$,
 239 $\{x_{k,N}, x_{k+1,N}, \dots, x_{n,N}\}$ are independent for $k - s > m$. If

$$\limsup_{N \rightarrow \infty} \sum_{k=1}^N E x_{k,N}^2 < \infty \quad (30)$$

240 and

$$\lim_{N \rightarrow \infty} \sum_{k=1}^N E |x_{k,N}|^{2+\delta} = 0, \quad \delta > 0, \text{ Lyapunov's condition} \quad (31)$$

241 , then S_N is asymptotically normal distributed with zero mean and covariance matrix $Q = \lim_{N \rightarrow \infty} ES_N S_N^T$. See [32]
242 and [38].

243 Lemma 2

244 Let $S_N = Z_{m,N} + X_{m,N}$, $m, N = 1, 2, \dots$ such that

- 245 • $EX_{m,N}^2 \leq C_m$, $\lim_{m \rightarrow \infty} C_m = 0$.
- 246 • $P(Z_{m,N} \leq z) = F_{m,N}(z)$.

247 Then $\lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} P(Z_{m,N} \leq z) = F(z)$. See [16] and [5].

248 To prove the asymptotic normality of $\sqrt{N}(\hat{\theta}_N^H - \theta^*)$, signal models, stochastic differentiability and m -dependence
249 are required. Let us consider the following theorem

250 **theorem 2.** Let $\varepsilon_1(\theta^*, \gamma^*), \dots, \varepsilon_N(\theta^*, \gamma^*)$ be iid rv's from the pdf f_H with an unknown parameter θ^* , $\theta^* \in \mathcal{D}_c^\theta$ with
251 \mathcal{D}_c^θ a compactness and \mathcal{D}_c^θ interior of \mathcal{D}_c^θ . Then the MLE $\hat{\theta}_N^H$ of θ^* is asymptotically normal

$$\sqrt{N}(\hat{\theta}_N^H - \theta^*) \xrightarrow{d} \mathcal{N}(0, \mathcal{P}(\theta^*, \gamma^*)) \quad (32)$$

252 where $\mathcal{P}(\theta^*, \gamma^*)$ is the asymptotic covariance matrix given by (21).

253 In order to do so, all the following assumptions hold. Suppose $W_N(\hat{\theta}_N^H, \gamma^*) \geq \sup_{\theta \in D_M, \gamma^* \in \mathcal{D}_c^\gamma} W_N(\theta, \gamma^*) - o_p(N^{-1})$,

254 $\hat{\theta}_N^H \xrightarrow{prob} \theta^*$, and

255 (i) $W(\theta, \gamma^*)$ is maximized on D_M at θ^*

256 (ii) θ^* is an interior point of D_M

257 (iii) $W(\theta, \gamma)$ is twice differentiable at (θ^*, γ^*) with nonsingular second derivative $\overline{\partial_{\theta\theta}^2 W_N}(\theta, \gamma)$

258 (iv) $\sqrt{N}(E\partial_\theta W_N(\theta, \gamma^*))_{\hat{\theta}_N^H} \xrightarrow{d} \mathcal{N}(0, Q(\theta^*, \gamma^*))$

259 (v) For any $\delta_N \rightarrow 0$, $\sup_{\|\hat{\theta}_N^H - \theta^*\| \leq \delta_N, \gamma^* \in \mathcal{D}_c^\gamma} \left| \frac{\hat{R}_N(\hat{\theta}_N^H, \gamma^*)}{1 + \sqrt{N}\|\hat{\theta}_N^H - \theta^*\|} \right| \xrightarrow{prob} 0$ with the remainder

$$\hat{R}_N(\theta, \gamma^*) = \sqrt{N} \frac{W_N(\theta, \gamma^*) - W_N(\theta^*, \gamma^*) - (\partial_\theta W_N(\theta, \gamma^*))_{\theta^*} (\theta - \theta^*) - W(\theta, \gamma^*) + W(\theta^*, \gamma^*)}{\|\theta - \theta^*\|} \quad (33)$$

260 then $\sqrt{N}(\hat{\theta}_N^H - \theta^*) \xrightarrow{d} \mathcal{N}(0, \mathcal{P}(\theta^*, \gamma^*))$. The proof is given in Appendix A.

261 3. Experimental results

262 Experimental results are presented over 16 CO, 15 PD, and 19 HD, left and right feet for different estimation
263 norms. The L_2 norm corresponds to the LSE (least square estimation), L_1 norm to the least sum absolute deviation
264 (LSAD) and L_∞ norm to the supremum norm given by $(\hat{\theta}_N^\infty = \min_\theta \max_t |\varepsilon_t(\theta, \gamma^*)|)$. In the Huberian context, a
265 campaign of estimations is carried out in C_γ with $\gamma^* = 1.5$ ([22]) and E_γ with $0.001 \leq \gamma^* \leq 0.2$. For each
266 estimator, comparisons between CO vs PD and HD for left and right feet are given. Table. A.1 shows the means

267 of γ^* , $RMSE$, $FIT(\%)$, $L_2C(\%)$, $L_1C(\%)$ and the total number of parameters $n = n_A + n_C$. The RMSE is the
 268 root mean square error between process output and prediction model output. The FIT is given by $100 \left(1 - \frac{y - \hat{y}}{y - \langle y \rangle}\right)$
 269 where y , \hat{y} and $\langle y \rangle$ are the process output, the prediction model output and the mean of the process output,
 270 respectively. L_2C and L_1C are the L_2 and L_1 contributions respectively given by $L_iC = \frac{card[y_i(\hat{\theta}_N^H, \gamma^*)]}{N}$. These are
 271 indicators of the density of NO in the prediction errors. If $L_2C = 40\%$ this means that 40% of prediction errors
 272 belong to the interval $[-\gamma^*, \gamma^*]$ and deal with the L_2 norm in the Huberian function. Here, the threshold γ in E_γ
 273 was varied among the range $[0.001; 0.2]$ with an incremental step of 0.001 for CO, PD and HD. We focus on the
 274 main results in Table. A.1. First, the L_2 , L_1 and L_∞ norms give bad results with large RMSE, low FIT and large
 275 number of parameters between 40 and 70. The lacks of robustness and degree of freedom (DOF) in these norms
 276 lead to an overestimation of the number of parameters n . On the other hand, each FIT presents a low value. In C_γ
 277 for $\gamma^* = 1.5$, the number of parameters is reduced with $25 \leq n \leq 32$ but not sufficient for a reduced order ARMA
 278 modeling. We can notice a great L_2 contribution, meaning a too large contribution of the L_2 norm, very sensitive
 279 to the large NO in the prediction errors.

280 The Huberian approach in E_γ leads to relevant results. Indeed, this remains in agreement with the formal point of
 281 view related to the bias and the new curve in section 2.3: low values of γ involve reduced bias and improve the FIT
 282 of the reduced order model. In Corbier and Carmona [15] we showed that the Huberian model order denoted d_M^H
 283 is such that $d_M^H < d_M^{L_1} < d_M^{L_2}$ since the Huberian function has one DOF and can be tuned from γ , by improving the
 284 estimation and reducing the number of parameters for pseudolinear models.

285 First we notice that $\langle \gamma_{control}^* \rangle \approx 2 \langle \gamma_{disease}^* \rangle$, meaning that there are twice more NO in STS-PD and STS-HD
 286 than STS-CO. Indeed, for PD and HD, the estimation requires a low value of γ^* involving a large value of the L_1
 287 contribution close to 70%. For CO, $\gamma^* \approx 0.19$ and $L_1C \approx 58\%$. Table. A.2 shows the parameters and variance
 288 of each parameter for CO and PD left with $\gamma^* = 0.05$ and $\gamma^* = 0.003$ respectively. For the variance/covariance
 289 matrix of these models, the large order \mathcal{L} is equal to 10 ensuring a low computational cost of $C_k(\hat{\theta}_N^H, \gamma^*)$. Table.
 290 A.3 yields the coefficients A_m^N for $m = 0..10$. Figure. A.4 and A.5 show two ARMA models for left CO ($\gamma^* = 0.05$)
 291 and left PD ($\gamma^* = 0.003$) respectively with a FIT close to 83%. In Figure. A.5 NO clearly appear in index-times
 292 $k = 52, k = 113, k = 190$ and $k = 247$ with high levels corresponding to the turn around during the walking period.
 293 In this phase, the classical estimators are highly disturbed and achieve sometimes the leverage point [22]. We can
 294 notice the good behavior of the Huberian reduced order ARMA model during this phase. Equation (34) shows the
 295 reduced order ARMA model of left PD for $\gamma^* = 0.003$.

$$\delta t_k = 0,712\delta t_{k-1} + 0,022\delta t_{k-2} + 0,018\delta t_{k-3} + 0,181\delta t_{k-4} + 0,060\delta t_{k-5} + e_k - 0,236e_{k-1} - 0,065e_{k-2} + 0.141e_{k-3}$$

$$-0,098e_{k-4} \quad (34)$$

297 The limited number of ARMA parameters contradicts conclusions in [20] and recently in [1]. These studies showed
298 a stride intervals of normal human walking which exhibit long-range temporal correlations. They presented a
299 highly simplified walking model by reproducing the long-range correlations observed in stride intervals without
300 complex peripheral dynamics. Based on fractal approach they showed an important point of view related to the
301 long-range *memory effect* of human walking. Our new approach shows a *short-range memory effect* for normal
302 and disease human walking. It remains to investigate this *memory effect* and try to interpret in physiological terms
303 the correlations with the CNS.

304 4. Conclusion

305 The main purpose of this paper has been to present a reduced order ARMA estimation method based on a
306 robust approach using Huberian function for the neurodegenerative disorder signal modeling. A new approach
307 has been presented to choose the threshold in Huberian function, allowing a best treatment of the natural outliers
308 contained in the signals. The reduced number of parameters is due to a relevant choice of this threshold in a
309 new interval range. Convergence and consistency properties of the robust estimator have been shown including
310 stochastic differentiability and m -dependence approaches. An estimations campaign has been conducted from STS
311 real measurements and it has been shown the relevance to use a Huberian function with DOF to tune its threshold in
312 order to assess a reduced order ARMA model. However, it remains to characterize more appreciably the diseases
313 to differentiate the neurodegenerative disorders. Accordingly, future work will focus on mixed L_p estimator [15]
314 to reduce the number of parameters providing new indicators and will investigate the *memory effect* of human
315 walking.

316 **Appendix A. Proof of the theorem 2**

 317 (i): From $E \ln f_H(\varepsilon_k(\theta, \gamma^*))$, we can deduce that

$$\theta^* = \underset{\theta \in D_M, \gamma^* \in \mathcal{D}_c^y}{\operatorname{argmax}} \left(\frac{-1}{N} \sum_{k=1}^N \rho_\gamma^H(\varepsilon_k(\theta, \gamma^*)) \right), \text{ as } N \rightarrow \infty \quad (\text{A.1})$$

318 which is equivalent to

$$\underset{\theta \in D_M, \gamma^* \in \mathcal{D}_c^y}{\operatorname{argmin}} \left(\frac{1}{N} \sum_{k=1}^N \rho_\gamma^H(\varepsilon_k(\theta, \gamma^*)) \right), \text{ as } N \rightarrow \infty \quad (\text{A.2})$$

 319 Since $W(\theta, \gamma^*) = E(\rho_\gamma^H(\varepsilon_k(\theta, \gamma^*)))$, then $W(\theta, \gamma^*)$ is maximized on D_M at θ^* .

 320 (ii): The interior condition is equivalent to the assumption $\theta^* \in \mathcal{D}_c^\theta$ where \mathcal{D}_c^θ is the interior of D_M .

 321 (iii): Using the stochastic differentiability condition, $E \partial_{\xi\xi}^2 \rho_\gamma^H(\varepsilon_k(\theta^*, \gamma^*)) = \overline{\partial_{\xi\xi}^2 W}(\theta^*, \gamma^*)$ is invertible as $N \rightarrow \infty$.

322 (iv): Using the mean value theorem, we get

$$\left(E \partial_{\xi\xi}^2 W_N(\xi, \gamma^*) \right)_{\hat{\theta}_N^H} (\hat{\theta}_N^H - \theta^*) = (E \partial_\theta W_N(\theta, \gamma^*))_{\hat{\theta}_N^H} - (E \partial_\theta W_N(\theta, \gamma^*))_{\theta^*} \quad (\text{A.3})$$

 323 with $\hat{\theta}_N^H \leq \tilde{\theta}_N \leq \theta^*$. For $N \rightarrow \infty$, $\tilde{\theta}_N \rightarrow \theta^*$, $(E \partial_\theta W_N(\theta, \gamma^*))_{\theta^*} = 0$ and $\lim_{N \rightarrow \infty} \left(E \partial_{\xi\xi}^2 W_N(\xi, \gamma^*) \right)_{\tilde{\theta}_N} \rightarrow \overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*)$. One

324 has

$$\sqrt{N} (\hat{\theta}_N^H - \theta^*) = \left(\overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*) \right)^{-1} \sqrt{N} (E \partial_\theta W_N(\theta, \gamma^*))_{\hat{\theta}_N^H} \quad (\text{A.4})$$

 325 The asymptotic normality of $\sqrt{N} (\hat{\theta}_N^H - \theta^*)$ only depends on the asymptotic normality of $\sqrt{N} (E \partial_\theta W_N(\theta, \gamma^*))_{\hat{\theta}_N^H}$.

 326 Let us denote $\partial_\theta W_N(\theta, \gamma) = \frac{1}{N} \sum_{k \in \mathcal{V}_2(\theta, \gamma)} \psi_k(\theta, \gamma) \varepsilon_k(\theta, \gamma) - \frac{\gamma}{N} \sum_{k \in \mathcal{V}_1(\theta, \gamma)} \psi_k(\theta, \gamma) s_k(\theta, \gamma) = \frac{1}{N} \sum_{k=1}^N \check{\Psi}_k(\theta, \gamma)$. Therefore,

$$-\sqrt{N} (E \partial_\theta W_N(\theta, \gamma^*))_{\hat{\theta}_N^H} = \sqrt{N} \left\{ \frac{1}{N} \sum_{k=1}^N [\check{\Psi}_k(\hat{\theta}_N^H, \gamma^*) - E \check{\Psi}_k(\hat{\theta}_N^H, \gamma^*)] \right\} - \sqrt{N} \frac{1}{N} \sum_{k=1}^N \check{\Psi}_k(\hat{\theta}_N^H, \gamma^*)$$

327 that is

$$= \frac{1}{\sqrt{N}} \sum_{k=1}^N (\check{\Psi}_k(\hat{\theta}_N^H, \gamma^*) - E \check{\Psi}_k(\hat{\theta}_N^H, \gamma^*)) - \frac{1}{\sqrt{N}} \sum_{k=1}^N \check{\Psi}_k(\hat{\theta}_N^H, \gamma^*) \quad (\text{A.5})$$

 328 Let us denote $S_N(\theta, \gamma^*) = \frac{1}{\sqrt{N}} \sum_{k=1}^N (\check{\Psi}_k(\theta, \gamma^*) - E \check{\Psi}_k(\theta, \gamma^*))$, then

$$-\sqrt{N} (E \partial_\theta W_N(\theta, \gamma^*))_{\hat{\theta}_N^H} = (S_N(\hat{\theta}_N^H, \gamma^*) - S_N(\theta^*, \gamma^*)) + S_N(\theta^*, \gamma^*) - \frac{1}{\sqrt{N}} \sum_{k=1}^N \check{\Psi}_k(\hat{\theta}_N^H, \gamma^*) \quad (\text{A.6})$$

 329 Since $\frac{1}{N} \sum_{k=1}^N \check{\Psi}_k(\hat{\theta}_N^H, \gamma^*) = 0$, the third term on the right hand side of (A.6) is $o(1)$. Its first term is $o(1)$ provided

 330 $\{S_N(\bullet, \gamma^*), N \geq 1\}$ is stochastically equicontinuous and $\hat{\theta}_N^H \xrightarrow{\text{prob}} \theta^*$. This follows because given any $\alpha > 0$ and

331 $\beta > 0$, there exists a $\delta > 0$ such that for $\Delta S(\hat{\theta}_N^H, \theta^*, \gamma^*) = S_N(\hat{\theta}_N^H, \gamma^*) - S_N(\theta^*, \gamma^*)$

$$\overline{\lim}_{N \rightarrow \infty} P\left(\left|\Delta S(\hat{\theta}_N^H, \theta^*, \gamma^*)\right| > \alpha\right) \leq$$

332
$$\overline{\lim}_{N \rightarrow \infty} P\left(\left|\Delta S(\hat{\theta}_N^H, \theta^*, \gamma^*)\right|, \left\|\rho_\gamma^H(\varepsilon_k(\hat{\theta}_N^H), \gamma^*) - \rho_\gamma^H(\varepsilon_k(\theta^*), \gamma^*)\right\| \leq \delta\right) + \overline{\lim}_{N \rightarrow \infty} P\left(\left\|\rho_\gamma^H(\varepsilon_k(\hat{\theta}_N^H), \gamma^*) - \rho_\gamma^H(\varepsilon_k(\theta^*), \gamma^*)\right\| > \delta\right)$$

333 (A.7)

$$\leq \overline{\lim}_{N \rightarrow \infty} P\left(\sup_{\theta \in D_M, \gamma^* \in \mathcal{D}_c^*} |S_N(\theta, \gamma^*) - S_N(\theta^*, \gamma^*)| > \alpha\right) < \beta$$

(A.8)

334 where the second inequality uses $\hat{\theta}_N^H \xrightarrow{prob} \theta^*$ and the third uses the stochastic equicontinuity. Accordingly, for a
 335 given threshold γ^* , this shows that for N tends to infinity, we have in law

$$\mathcal{L}\left(\sqrt{N}(\hat{\theta}_N^H - \theta^*)\right) \underset{N \rightarrow \infty}{\sim} \mathcal{L}(S_N(\theta^*, \gamma^*))$$

(A.9)

336 with

$$S_N(\theta^*, \gamma^*) = \frac{1}{\sqrt{N}} \sum_{k=1}^N \left(\frac{d}{d\theta} \rho_\gamma^H(\varepsilon_k(\theta, \gamma^*)) - E \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_k(\theta, \gamma^*)) \right)_{\theta^*}$$

(A.10)

337 The purpose is to prove that $S_N(\theta^*, \gamma^*)$ is a normal asymptotic distribution. For this, we show that the terms of
 338 $S_N(\theta^*, \gamma^*)$ are independent. As described above, we use the m -dependence approach to show the asymptotic normal
 339 behavior of $S_N(\theta^*, \gamma^*)$. Let us consider the following short expressions: $\varepsilon_{v_i, k}(\theta^*, \gamma^*) = \varepsilon_{i, k}^*$, $f_{t, k}(\theta^*, \gamma^*) = f_{t, k}^*$. We
 340 split $\varepsilon_{2, t}^*$ and $\psi_{2, t}^*$ into one part that satisfies m -dependence conditions among its terms and one part that is small.

341 We then have

$$\varepsilon_{2, t}^* = \varepsilon_{2, t}^{*, m} + \tilde{\varepsilon}_{2, t}^{*, m} + \varepsilon_{1, t}^* = \sum_{k=0}^m \beta_{t, k}^* e_{t-k} + \sum_{k=m+1}^{\infty} \beta_{t, k}^* e_{t-k} + \sum_j \Omega_j^* \delta_{t, j}^K$$

(A.11)

342 where m is an integer with $\Omega_j^* = \Omega_j(\theta^*, \gamma^*)$ and $\delta_{t, j}^K$ is the Kronecker's function. Analogously, we have

$$\psi_{2, t}^* = \psi_{2, t}^{*, m} + \tilde{\psi}_{2, t}^{*, m} + \psi_{1, t}^* = \sum_{k=0}^m \alpha_{t, k}^* e_{t-k} + \sum_{k=m+1}^{\infty} \alpha_{t, k}^* e_{t-k} + \sum_j \phi_j^* \delta_{t, j}^K$$

(A.12)

343 $S_N(\theta^*, \gamma^*)$ can be written as $S_N(\theta^*, \gamma^*) = Z_{m, N}(\theta^*, \gamma^*) + X_{m, N}(\theta^*, \gamma^*)$ with

$$Z_{m, N}(\theta^*, \gamma^*) = \frac{1}{\sqrt{N}} \sum_{t=1}^N \left(\frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) - E \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right)_{\theta^*}$$

(A.13)

344
$$X_{m, N}(\theta^*, \gamma^*) = \frac{1}{\sqrt{N}} \sum_{t=1}^N \frac{d}{d\theta} \left[\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*)) - \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right]_{\theta^*} - E \frac{d}{d\theta} \left[\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*)) - \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right]_{\theta^*}$$
 (A.14)

345 **Part1:**

346 From (A.13) in $Z_{m,N}(\theta^*, \gamma^*)$ and using the Lyapunov's condition, we obtain

$$E \left| \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) - E \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right|^{\delta+2} \leq 2^{\delta+1} E \left(\left| \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right|^{\delta+2} + E \left| \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right|^{\delta+2} \right) \quad (\text{A.15})$$

347

$$\leq 2^{\delta+2} E \left| \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right|^{\delta+2} \quad (\text{A.16})$$

348 with

$$\left| \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right|^{\delta+2} \leq (|\psi_{2,t}^{*,m}| |\varepsilon_{2,t}^{*,m}| + \gamma^* |\psi_{1,t}^{*,m}|)^{\delta+2} \leq 2^{\delta+1} (|\psi_{2,t}^{*,m}|^{\delta+2} |\varepsilon_{2,t}^{*,m}|^{\delta+2} + (\gamma^*)^{\delta+2} |\psi_{1,t}^{*,m}|^{\delta+2}) \quad (\text{A.17})$$

349 We deduce

$$2^{\delta+2} E \left| \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right|^{\delta+2} \leq 2^{2\delta+3} E |\psi_{2,t}^{*,m}|^{\delta+2} |\varepsilon_{2,t}^{*,m}|^{\delta+2} + 2^{2\delta+3} (\gamma^*)^{\delta+2} E |\psi_{1,t}^{*,m}|^{\delta+2} \quad (\text{A.18})$$

350 Using Schwarz's inequality

$$2^{\delta+2} E \left| \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right|^{\delta+2} \leq 2^{2\delta+3} \left(E |\psi_{2,t}^{*,m}|^{2\delta+4} E |\varepsilon_{2,t}^{*,m}|^{(2\delta+4)} \right)^{\frac{1}{2}} + 2^{2\delta+3} (\gamma^*)^{\delta+2} \left(E |\psi_{1,t}^{*,m}|^{2\delta+4} \right)^{\frac{1}{2}} \quad (\text{A.19})$$

351 The first and second terms on the right hand side of (A.19) are respectively denoted A^* and B^* .

352 • For A^* : in v_2 , for all t and θ^* , $|\varepsilon_{2,t}^{*,m}| \leq \gamma^*$. Therefore

$$E |\psi_{2,t}^{*,m}|^{2\delta+4} \leq 2^{2\delta+3} E |e_{t-k}|^{2\delta+4} \left(\sum_{k=0}^m \mu_k \right)^{2\delta+4} \quad (\text{A.20})$$

353 From **H1**, we have $E |\psi_{2,t}^{*,m}|^{2\delta+4} \leq C^*$ and $A^* \leq C^*$.

354 • For B^* : from **H1** we get $\sup_{t, \theta^*, \gamma^*} |\varepsilon_{1,t}^{*,m}| = \hat{\Omega}$ and $E |\psi_{1,t}^{*,m}|^{2\delta+4}$ are bounded. Accordingly, $B^* \leq C^*$.

355

356 Inserting $\left(\frac{1}{\sqrt{N}} \right)^{\delta+2}$, we finally obtain for all γ^*

$$E \left| \frac{1}{\sqrt{N}} \left[\frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) - E \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right] \right|_{\theta^*}^{\delta+2} \leq \frac{C}{N^{1+\frac{\delta}{2}}} \quad (\text{A.21})$$

357 Then

$$\lim_{N \rightarrow \infty} \sum_{t=1}^N E \left| \frac{1}{\sqrt{N}} \left[\frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) - E \frac{d}{d\theta} \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) \right] \right|_{\theta^*}^{\delta+2} \leq \lim_{N \rightarrow \infty} \frac{C}{N^{\frac{\delta}{2}}} \rightarrow 0 \quad (\text{A.22})$$

358 Expression (A.22) proves (30) and (31) in lemma 1 (section 2.6.3) with $Q_m(\theta^*, \gamma) = \lim_{N \rightarrow \infty} EZ_{m,N}(\theta^*, \gamma^*) Z_{m,N}^T(\theta^*, \gamma^*)$.

359 **Part2:**

360 In $X_{m,N}(\theta^*, \gamma^*)$, we can write

$$\begin{aligned}
\frac{d\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{d\theta} &= \frac{\partial\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{\partial\varepsilon_t(\theta, \gamma^*)} \frac{\partial\varepsilon_t(\theta, \gamma^*)}{\partial\theta} - \frac{\partial\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{\partial\varepsilon_t(\theta, \gamma^*)} \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} \\
&+ \frac{\partial\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{\partial\varepsilon_t(\theta, \gamma^*)} \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} - \frac{\partial\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{\partial\varepsilon_t^m(\theta, \gamma^*)} \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta}
\end{aligned} \tag{A.23}$$

362 Therefore

$$\begin{aligned}
\frac{d\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{d\theta} &= \frac{\partial\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{\partial\varepsilon_t(\theta, \gamma^*)} \left(\frac{\partial\varepsilon_t(\theta, \gamma^*)}{\partial\theta} - \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} \right) \\
&+ \left(\frac{\partial\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{\partial\varepsilon_t(\theta, \gamma^*)} - \frac{\partial\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{\partial\varepsilon_t^m(\theta, \gamma^*)} \right) \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta}
\end{aligned} \tag{A.24}$$

364 Using mean value theorem, we get

$$\frac{d\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{d\theta} = \frac{\partial^2\rho_\gamma^H(\xi_t(\theta, \gamma^*))}{\partial\xi_t(\theta, \gamma^*)^2} (\varepsilon_t(\theta, \gamma^*) - \varepsilon_t^m(\theta, \gamma^*)) \tag{A.25}$$

365 Hence

$$\begin{aligned}
\left| \frac{d\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{d\theta} \right| &\leq \left| \frac{\partial\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{\partial\varepsilon_t(\theta, \gamma^*)} \right| \left| \frac{\partial\varepsilon_t(\theta, \gamma^*)}{\partial\theta} - \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} \right| \\
&+ \left| \frac{\partial^2\rho_\gamma^H(\xi_t(\theta, \gamma^*))}{\partial\xi_t(\theta, \gamma^*)^2} \right| |\varepsilon_t(\theta, \gamma^*) - \varepsilon_t^m(\theta, \gamma^*)| \left| \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} \right|
\end{aligned} \tag{A.26}$$

367 From regularity conditions C1 in (see [25]) given by

- 368 • $\left\| \frac{\partial\rho(\varepsilon)}{\partial\varepsilon} \right\| \leq C|\varepsilon|, \theta \in D_M, \text{ all } t.$
- 369 • $\left\| \frac{\partial\rho(\varepsilon)}{\partial\theta} \right\| \leq C|\varepsilon|^2, \theta \in D_M, \text{ all } t.$
- 370 • $\left\| \frac{\partial^2\rho(\varepsilon)}{\partial\varepsilon^2} \right\| \leq C.$

372 We then have

$$\begin{aligned}
\left| \frac{d\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{d\theta} \right| &\leq C|\varepsilon_t(\theta, \gamma^*)| \left| \frac{\partial\varepsilon_t(\theta, \gamma^*)}{\partial\theta} - \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} \right| \\
&+ C|\varepsilon_t(\theta, \gamma^*) - \varepsilon_t^m(\theta, \gamma^*)| \left| \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} \right|
\end{aligned} \tag{A.27}$$

374 In detail

$$\left| \frac{\partial\varepsilon_t(\theta, \gamma^*)}{\partial\theta} - \frac{\partial\varepsilon_t^m(\theta, \gamma^*)}{\partial\theta} \right| = |\tilde{\psi}_t^m(\theta, \gamma^*)| \leq \sum_{k=m+1}^{\infty} |\alpha_{t,k}^*| |e_{t-k}| \leq \sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \tag{A.28}$$

375 and

$$|\varepsilon_t(\theta, \gamma^*) - \varepsilon_t^m(\theta, \gamma^*)| = |\tilde{\varepsilon}_t^m(\theta, \gamma^*)| \leq \sum_{k=m+1}^{\infty} |\beta_{t,k}^*| |e_{t-k}| \leq \sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \quad (\text{A.29})$$

376 Expression (A.27) becomes

$$\left| \frac{d\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{d\theta} \right| \leq C \left(\sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \right) (|\varepsilon_t(\theta, \gamma^*)| + |\psi_t^m(\theta, \gamma^*)|) \quad (\text{A.30})$$

377 Moreover

$$|\varepsilon_t(\theta, \gamma^*)| + |\psi_t^m(\theta, \gamma^*)| \leq 2 \sum_{k=0}^m \mu_k |e_{t-k}| + \sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \quad (\text{A.31})$$

378 Therefore

$$\left| \frac{d\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))}{d\theta} \right| \leq \alpha_t + \beta_t \quad (\text{A.32})$$

379 with

$$\alpha_t = 2C \left(\sum_{k=0}^m \mu_k \right) \left(\sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \right) \quad (\text{A.33})$$

380 and

$$\beta_t = C \left(\sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \right)^2 \quad (\text{A.34})$$

381 Therefore

$$X_{m,N}(\theta^*, \gamma^*) \leq \underbrace{\frac{1}{\sqrt{N}} \sum_{t=1}^N (\alpha_t - E\alpha_t)}_{X_{m,N}^\alpha(\theta^*, \gamma^*)} + \underbrace{\frac{1}{\sqrt{N}} \sum_{t=1}^N (\beta_t - E\beta_t)}_{X_{m,N}^\beta(\theta^*, \gamma^*)} \quad (\text{A.35})$$

382 Each term on the right hand side of (A.35) verifies the corollary of the lemma 2B.1 in [27](p.57). Hence, as $m \rightarrow \infty$

$$E \left(X_{m,N}^\alpha(\theta^*, \gamma^*) \right)^2 \leq K \left(\sum_{k=0}^m \mu_k \right) \left(\sum_{k=m+1}^{\infty} \mu_k \right) \rightarrow 0 \quad (\text{A.36})$$

383

$$E \left(X_{m,N}^\beta(\theta^*, \gamma^*) \right)^2 \leq K \left(\sum_{k=m+1}^{\infty} \mu_k \right)^2 \rightarrow 0 \quad (\text{A.37})$$

384 Hence, $Z_{m,N}(\theta^*, \gamma^*) \in \mathcal{AN}(0, Q_m(\theta^*, \gamma^*))$ and $S_N(\theta^*, \gamma^*) \in \mathcal{AN}(0, Q(\theta^*, \gamma^*))$ with $Q(\theta^*, \gamma^*) = \lim_{m \rightarrow \infty} Q_m(\theta^*, \gamma^*)$.

385 Which proves the point (iv) of the Theorem 2.

386 (v): Expanding $W(\theta, \gamma^*)$ into Taylor series around θ^* , we get

$$W(\hat{\theta}_N^H, \gamma^*) = W(\theta^*, \gamma^*) + \frac{1}{2} (\hat{\theta}_N^H - \theta^*)^T \overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*) (\hat{\theta}_N^H - \theta^*) + o\left(\|\hat{\theta}_N^H - \theta^*\|^2\right) \quad (\text{A.38})$$

387 Since $\overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*)$ is positive definite and nonsingular, there exists $C > 0$ and a neighborhood of θ^* such that

$$\frac{1}{2} (\hat{\theta}_N^H - \theta^*)^T \overline{\partial_{\theta\theta}^2 W}(\theta^*, \gamma^*) (\hat{\theta}_N^H - \theta^*) + o\left(\|\hat{\theta}_N^H - \theta^*\|^2\right) \leq C \|\hat{\theta}_N^H - \theta^*\|^2 \quad (\text{A.39})$$

388 we obtain $W(\hat{\theta}_N^H, \gamma^*) \leq W(\theta^*, \gamma^*) + C \|\hat{\theta}_N^H - \theta^*\|^2$. Moreover

$$W_N(\hat{\theta}_N^H, \gamma^*) - W_N(\theta^*, \gamma^*) + o\left(\frac{1}{N}\right) = W(\hat{\theta}_N^H, \gamma^*) - W(\theta^*, \gamma^*) + (\partial_{\xi} W_N(\xi, \gamma^*))_{\theta^*}^T (\hat{\theta}_N^H - \theta^*) + \|\hat{\theta}_N^H - \theta^*\| \hat{R}_N(\hat{\theta}_N^H, \gamma^*) + o\left(\frac{1}{N}\right) \quad (\text{A.40})$$

389 Therefore

$$\begin{aligned} W_N(\hat{\theta}_N^H, \gamma^*) - W_N(\theta^*, \gamma^*) &\leq C \|\hat{\theta}_N^H - \theta^*\|^2 + \|\partial_{\xi} W_N(\xi, \gamma^*)\|_{\theta^*} \|\hat{\theta}_N^H - \theta^*\| \\ &+ \|\hat{\theta}_N^H - \theta^*\| \left(1 + \sqrt{N} \|\hat{\theta}_N^H - \theta^*\|\right) o\left(\frac{1}{\sqrt{N}}\right) + o\left(\frac{1}{N}\right) \end{aligned} \quad (\text{A.41})$$

391 Since $\|\partial_{\xi} W_N(\xi, \gamma^*)\|_{\theta^*} \rightarrow 0$ as $N \rightarrow \infty$, then

$$W_N(\hat{\theta}_N^H, \gamma^*) - W_N(\theta^*, \gamma^*) \leq (C + o(1)) \|\hat{\theta}_N^H - \theta^*\|^2 + o\left(\frac{1}{\sqrt{N}}\right) \|\hat{\theta}_N^H - \theta^*\| + o\left(\frac{1}{N}\right) \quad (\text{A.42})$$

392 The remainder $\hat{R}_N(\hat{\theta}_N^H, \gamma^*)$ can be written as

$$\hat{R}_N(\hat{\theta}_N^H, \gamma^*) \leq \sqrt{N} \|\hat{\theta}_N^H - \theta^*\| (K + o(1)) \quad (\text{A.43})$$

393 then

$$\frac{\hat{R}_N(\hat{\theta}_N^H, \gamma^*)}{1 + \sqrt{N} \|\hat{\theta}_N^H - \theta^*\|} \leq \frac{\sqrt{N} \|\hat{\theta}_N^H - \theta^*\| (K + o(1))}{1 + \sqrt{N} \|\hat{\theta}_N^H - \theta^*\|} \quad (\text{A.44})$$

394 Since $\sqrt{N} \|\hat{\theta}_N^H - \theta^*\| \xrightarrow{prob} 0$ then

$$\sup_{\|\hat{\theta}_N^H - \theta^*\| \leq \delta_N, \gamma^* \rightarrow \mathcal{D}_c^{\gamma}} \left| \frac{\hat{R}_N(\hat{\theta}_N^H, \gamma^*)}{1 + \sqrt{N} \|\hat{\theta}_N^H - \theta^*\|} \right| \xrightarrow{prob} 0 \quad (\text{A.45})$$

395 which prove the point (v) and finally the theorem 2.

- 396 [1] J. Ahn, N. Hogan, Long-range correlations in stride intervals may emerge from non-chaotic walking dynamics, PLOS ONE, 8(9), 2013,
397 pp. 2-10.
- 398 [2] M.S. Aslam, M.A.Z. Raja, A new adaptive strategy to improve online secondary path modeling in active noise control system using
399 fractional signal processing approach, Signal Processing 107, 2015, pp. 433-443.
- 400 [3] A. Al-Smadi, A least-squares based algorithm for identification of non-gaussian ARMA models, Circuits Systems Signal Processing
401 26(5), 2007, pp. 715-731.
- 402 [4] K.E. Al-Qawasmī, A.M. Al-Smadi, A. Al-Hamami, A robust ARX and ARMA model order estimation via pivot-neighbors comparisons,

- 403 Recent Patents on Computing Science. 3, 2010, pp. 33-38.
- 404 [5] T.W. Anderson, On asymptotic distributions of estimated parameters of stochastic difference equation, *Ann. Math. Stat.*, 30, 1959, pp.
405 676-687.
- 406 [6] D.W.K. Andrews, Large sample estimation and hypothesis testing, *Handbook of Econometric*, Elsevier Science 4, 1994, pp. 2248-2292.
- 407 [7] V. Barnett, T. Lewis, *Outliers in statistical data*, 3rd ed. Wiley, 1998, New York.
- 408 [8] P.M. Broersen, The quality of models for ARMA processes, *IEEE Trans. Signal Process.* 46(6), 1998, pp. 1749-1752.
- 409 [9] Y. Chakhchoukh, A new robust estimation method for ARMA Models, *IEEE Tran on Sig Pro.* 58(7), 2010, pp. 3512-3522.
- 410 [10] N.I. Chaudhary, M.A.Z. Raja, J.A. Khan, M.S. Aslam, Identification of input nonlinear control autoregressive systems using fractional
411 signal processing approach, Hindawi Publishing Corporation, 2013. Article ID 467276. DOI:10.1155/2013/467276.
- 412 [11] N.I. Chaudhary, M.A.Z. Raja, Identification of Hammerstein nonlinear ARMAX systems using nonlinear adaptive algorithms, *Nonlinear*
413 *Dynamics*, 2014. DOI:10.1007/s11071-014-1748-8.
- 414 [12] C. Corbier, Contribution a l'estimation robuste de modeles dynamiques: application a la commande de systemes dynamiques complexes,
415 *Arts et Metiers ParisTech*, Thesis, Nov, 29, 2012.
- 416 [13] C. Corbier, M. El Badaoui, ARMA Modeling and Nonparametric Probability Density Function of Gait Signal Using $h_2 - L_1$ estimator in
417 Patients with Neuro-Degenerative Disease, *Computer Methods in Biomechanics and Biomedical Engineering*, Taylor and Francis.... 39th
418 Congress of the Societe de Biomecanique, Valenciennes, France, 27-29 August, 2014, 17(1), pp. 178-179.
- 419 [14] C. Corbier, J-C. Carmona, Extension of the Tuning Constant in the Hubers Function for Robust Modeling of Piezoelectric Systems,
420 *International Journal of Adaptive Control and Signal Processing*, August 2014. DOI:10.1002/acs.2517.
- 421 [15] C. Corbier, J-C. Carmona Mixed L_p -estimators Variety for Model Order Reduction in Control Oriented System Identification, Hindawi
422 Publishing Corporation *Mathematical Problems in Engineering*, July 2014. ID 349070.
- 423 [16] P.H. Diananda, Some probability limit theorems with statistical applications, *Proc. Cambridge Philos. Soc.*, 49, 1953, pp. 239-245.
- 424 [17] A.J. Fox, Outliers in time series, *Journal of the Royal Society*, 34(3), 1972, 350-363.
- 425 [18] A.S. Hadi, A.H.M. Rahmatullah Imon, M. Werner, Detection of outliers, John Wiley and Sons, Inc. *Wires Comp Stat*, 1, 2009, 57-70.
- 426 [19] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, W.A. Stahel, *Robust statistics: the approach based on influence function*, John Wiley and
427 Sons. New York, 1985.
- 428 [20] J.M. Hausdorff, C.K. Peng, Z. Ladin, J.Y. Wei, A.L. Goldberger, Is walking a random walk? Evidence for long-range correlations in
429 stride interval of human gait, *Journal of Applied Physiology*, 78(1), 1995, pp. 349-358.
- 430 [21] J.M. Hausdorff, M.E. Cudkowicz, R. Firtion, J.Y. Wei, A.L. Goldberger, Gait variability and basal ganglia disorders: stride-to-stride
431 variations of gait cycle timing in Parkinson's disease and Huntington's disease, *Movement disorders*, 13(3), 1998, 428-437.
- 432 [22] P.J. Huber, E.M. Ronchetti, *Robust statistics*. 2th (Eds), New York, NY: John Wiley and Sons, 2009.
- 433 [23] T. Iluz, E. Gazit, T. Herman, E. Sprecher, M. Brozgol, N. Giladi, A. Mirelman, J.M. Hausdorff, Automated detection of missteps during
434 community ambulation in patients with Parkinson's disease: a new approach for quantifying fall risk in the community setting, *Journal of*
435 *NeuroEngineering and Rehabilitation*, 2014, 11-48.
- 436 [24] Y-C. Lin, M. Gfoehler, M.G. Pandy, Quantitative evaluation of the major determinants of human gait. *J. Biomech*, 47, 2014, 1324-1331.
- 437 [25] L. Ljung, Convergence analysis of parametric identification methods, *IEEE Trans on Automatic Control*, vol. AC-23, no.5, 1978, pp.
438 770-783.
- 439 [26] L. Ljung, P.E. Caines, Asymptotic Normality of prediction error estimators for approximate systems models, *Stochastics*, 3, 1979, 29-46.
- 440 [27] L. Ljung, *System identification: theory for the user*, Prentice Hall PTR. New York, 1999.
- 441 [28] S. Maiz, M. El Badaoui, F. Bonnardot, C. Serviere, New order cyclostationary analysis and application to the detection and characterization
442 of a runner's fatigue. *Signal Processing* 102, 2014, 188-200.
- 443 [29] A.E. Martin, J.P. Schmiedeler, Predicting human walking gaits with a simple planar model, *J. Biomech*, 47, 2014, 1416-1421.

- 444 [30] N. Muler, D. Pena, V.J. Yohai, Robust estimation for ARMA models, *The Annals of Statistics*, 37(2), 2009, pp. 816840.
- 445 [31] W.K. Newey and D.L. McFadden, Large sample estimation and hypothesis testing, *Handbook of Econometric*, Elsevier Science 4, 1994,
446 pp. 2113-2247.
- 447 [32] S. Orey, A central limit theorem for m-dependent random variables, *Duke Math.J.*, 25, 1958, pp. 543-546.
- 448 [33] M.A.Z. Raja, N.I. Chaudhary, Two-stage fractional least mean square identification algorithm for parameter estimation of CARMA
449 systems, *Signal Processing*, 107, 2015, pp. 327-339. DOI:10.1016/j.sigpro.2014.06.015.
- 450 [34] C. Ran, Z. Deng, Self-tuning distributed measurement fusion Kalman estimator for multi-channel ARMA signal, *Signal Processing*, 91,
451 2011, pp. 2028-2041.
- 452 [35] L. Ren, R.K. Jones, D. Howard, Predictive modeling of human walking over a complete gait cycle, *J.Neurol., Neurosurgery, Psychiatry*,
453 79, 2007, 874-880.
- 454 [36] R.T. Roemmich, P.R., Zeilman, M.S.O., Vaillancourt, C.J., Hass, Gait variability magnitude but not structure is altered in essential tremor,
455 *J. Biomech*, 46(15), 2013, 2682-2687.
- 456 [37] A.H. Ropper, R.H., Brown, Adams and Victor's principles of neurology. 8th (Eds), McGraw-Hill, New York, NY, 2005.
- 457 [38] B. Rosen, On a central limit theorem for sums of dependent random variables, *Z., Wahrsch verw. Geb.*, 7, 1967, pp. 48-82.
- 458 [39] N. Scafetta, D., Marchi, B.J., West, Understanding the complexity of human gait dynamics, *Chaos*, 19(2), 2009, 026108.
- 459 [40] S. Sen Roy, S., Guria, Estimation of regression parameters in the presence of outliers in the response, *Statistics*, 25(5), 2009, 1000-1009.
- 460 [41] S.M. Shah, R. Samar, M.A.Z. Raja, J.A. Chambers, Fractional normalized filtered-error least mean squares algorithm for application in
461 active noise control systems, *Electronics Letters* 50 (14), 973-975, 2014. DOI: 10.1049/el.2014.1275.
- 462 [42] J. Sian, M. Gerlach, M.B.H. Youdim, P., Riederer, Parkinson's disease: a major hypokinetic basal ganglia disorder, *J.Neural Transmission*,
463 106(5-6), 1999, 443-476.
- 464 [43] Y. Wu, S., Krishnan, Statistical analysis of gait maturation in children using nonparametric probability density function. *Entropy*, 15,
465 2013, 753-766.
- 466 [44] F. Zacharia, M. El Badaoui, S. Maiz, F. Guillet, M. Khalil, K. Khalil, M. Halimi, Walking analysis: Empirical relation between kurtosis
467 and degree of cyclostationary, 2nd International Conference on Advances in Biomedical Engineering IEEE, 11-13 Sept, 2013, pp. 93-96.
- 468 [45] K. Zhu and S. Ling, LADE-based inference for ARMA models with unspecified and heavy-tailed heteroscedastic noises, *Chinese*
469 *Academy of Science, Hong Kong University of Science and Technology*, MPRA Paper No. 59099, 2014, pp. 2-33.

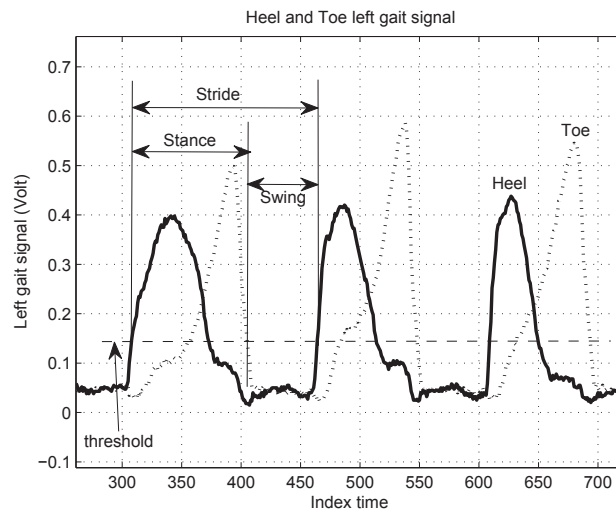


Figure A.1: Example of gait signals from heel and toe force sensors underneath the left foot. The threshold allows to compute the time-signals δt_k such as the stride, swing and stance.

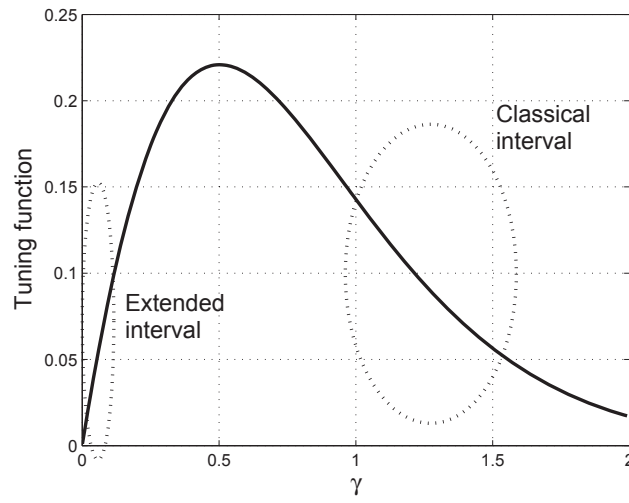


Figure A.2: Tuning function with two main intervals. The classical interval $\gamma \in [1, 1.5]$ and the extended interval $\gamma \in [0.001, 0.2]$.

Accepted manuscript

Table A.1: Means of γ^* , $RMSE$, $FIT(\%)$, $L_2C(\%)$, $L_1C(\%)$ and the total number of parameters $n = n_M + n_C$ over 16 CO 15 PD and 19 HD (left and right feet) for different estimation norms. L_2 is the LSE, L_1 is the LSAD, L_∞ is the supremum norm. C_γ is the classical interval in the Huber's context with $\langle \gamma^* \rangle = 1.5$. E_γ is the extended interval in the Huber's context with low values of γ^* .

			CO left							PD left			
Estimator	γ^*	$RMSE$	FIT	L_2C	L_1C	n	γ^*	$RMSE$	FIT	L_2C	L_1C	n	
L_2	–	11.2	10	100	0	70	–	13	9	100	0	70	
L_1	–	4.3	42	0	100	41	–	5.2	38	0	100	46	
L_∞	–	4.2	25	–	–	45	–	5.3	26	–	–	56	
Huber in C_γ	1.5	2.4	42	95	5	25	1.5	3.1	31	96	4	28	
Huber in E_γ	0.17	0.09	92	41	59	9	0.09	0.34	78	30	70	9	
			CO right							PD right			
L_2	–	10.2	9	100	0	70	–	13	9	100	0	70	
L_1	–	5.3	44	0	100	39	–	6.2	35	0	100	46	
L_∞	–	3.2	26	–	–	46	–	5.5	28	–	–	54	
Huber in C_γ	1.5	2.3	44	96	4	27	1.5	3.3	31	96	4	30	
Huber in E_γ	0.18	0.08	92	43	57	9	0.09	0.29	78	32	68	9	
			CO left							HD left			
L_2	–	11.2	10	100	0	70	–	8	17	100	0	70	
L_1	–	4.3	42	0	100	41	–	4.1	36	0	100	44	
L_∞	–	4.2	25	–	–	45	–	6.3	24	–	–	54	
Huber in C_γ	1.5	2.4	42	95	5	25	1.5	3.2	32	96	4	31	
Huber in E_γ	0.17	0.09	92	41	59	9	0.08	0.28	78	29	71	9	
			CO right							HD right			
L_2	–	10.2	9	100	0	70	–	13	9	100	0	70	
L_1	–	5.3	44	0	100	39	–	6.2	35	0	100	46	
L_∞	–	3.2	26	–	–	46	–	5.1	32	–	–	56	
Huber in C_γ	1.5	2.3	44	96	4	27	1.5	3.5	29	95	5	32	
Huber in E_γ	0.18	0.08	92	43	57	9	0.07	0.16	87	27	73	9	

Table A.2: Parameters of the CO ($\gamma^* = 0.05$) and PD ($\gamma^* = 0.003$) ARMA models and Huberian variance of each parameter λ^H .

	CO left				
i	1	2	3	4	5
a_i	-0,877	-0,152	0,173	-0,215	0,073
c_i	-0,236	-0,065	0,141	-0,098	-
$\lambda_{a_i}^H$	0.0021	0.0032	0.0015	0.0035	0.0026
$\lambda_{c_i}^H$	0.0012	0.0075	0.0056	0.0074	-
	PD left				
i	1	2	3	4	5
a_i	-0,712	-0,022	-0,018	-0,181	-0,060
c_i	-0,166	0,119	0,160	0,133	-
$\lambda_{a_i}^H$	0.0031	0.0022	0.0095	0.0015	0.0086
$\lambda_{c_i}^H$	0.0002	0.0005	0.0066	0.0024	-

Accepted manuscript

Table A.3: Coefficients A_m^N in the covariance matrix of the CO ($\gamma^* = 0.05$) and PD ($\gamma^* = 0.003$) ARMA models.

CO left											
m	0	1	2	3	4	5	6	7	8	9	10
A_m^N	1	0.91	0.86	0.74	0.62	0.45	0.33	0.22	0.19	0.11	0.09
PD left											
m	0	1	2	3	4	5	6	7	8	9	10
A_m^N	1	0.94	0.81	0.71	0.63	0.51	0.41	0.29	0.18	0.10	0.08

Accepted manuscript

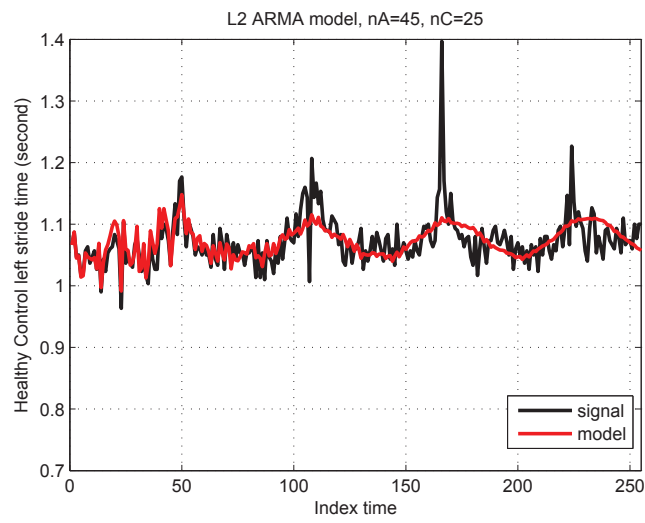


Figure A.3: Gaussian ARMA model of the left STS (red line) vs CO real signal (black line). $n_A = 45$, $n_C = 25$, $Fit = 9.5\%$.

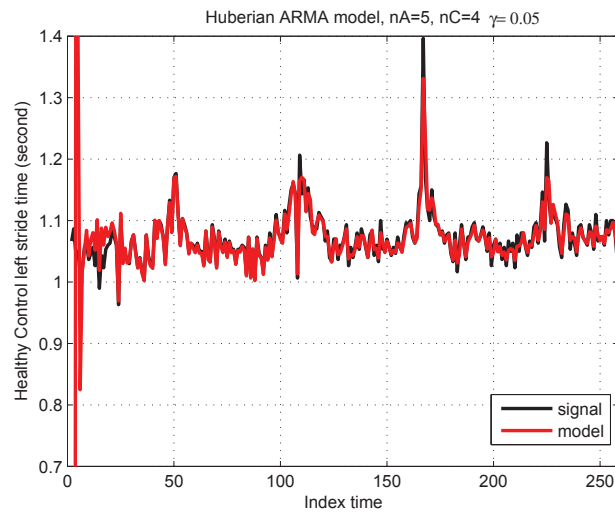


Figure A.4: Huberian ARMA model of the left STS (red line) vs CO real signal (black line). $n_A = 5$, $n_C = 4$, $Fit = 82.7\%$, $\gamma = 0.05$, $N = 253$.

Accepted manuscript

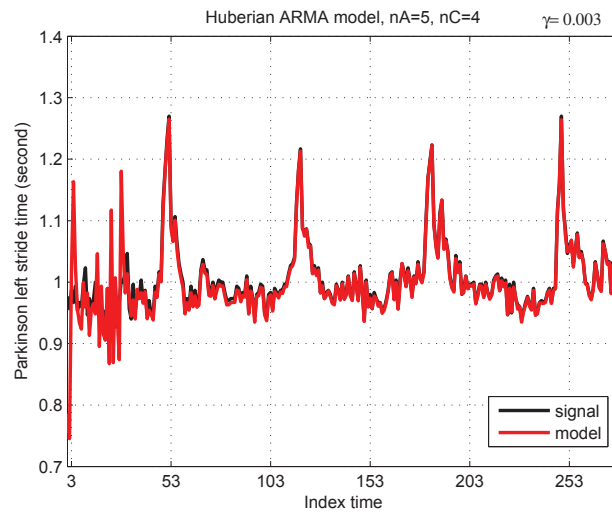


Figure A.5: Huberian ARMA model of the left STS (red line) vs PD real signal (black line). $n_A = 5$, $n_C = 4$, $Fit = 82.8\%$, $\gamma = 0.003$, $N = 288$.