CEC Dauphine | PSL

N°2025 – 06 • FEBRUARY 2025

WORKING PAPER

Forecasting Extreme Trajectories Using Seminorm Representations

Gilles de TRUCHIS [1], Sébastien FRIES [2*], Arthur THOMAS [3*4*]*

For (Xt) a two-sided α-stable moving average, this paper studies the conditional distribution of future paths given a piece of observed trajectory when the process is far from its central values. Under this framework, vectors of the form Xt = (Xt−m, ..., Xt, Xt+1, ..., Xt+h), m ≥ 0, h ≥ 1, are multivariate α-stable and the dependence between the past and future components is encoded in their spectral measures. A new representation of stable random vectors on unit cylinders {s ∈ R^(m+h+1) : ||s|| = 1} for ||·|| an adequate seminorm is proposed to describe the tail behaviour of vectors Xt when only the first m + 1 components are assumed to be observed and large in norm. Not all stable vectors admit such a representation, and (Xt) will have to be 'anticipative enough' for Xt to admit one. The conditional distribution of future paths can then be explicitly derived using the regularly varying tails property of stable vectors and has a natural interpretation in terms of pattern identification. Through Monte Carlo simulations, we develop procedures to forecast crash probabilities and crash dates and demonstrate their finite sample performances. As an empirical illustration, we estimate probabilities and reversal dates of El Niño and La Niña occurrences.

JEL CODES: C22, C53, C58, Q54

1*Orléans University, Laboratoire d'Economie d'Orléans (LEO),France, gilles.detruchis@univ-orleans.fr
2*Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, Amsterdam, Netherlands.
3* Paris Dauphine-PSL University, LEDa, CNRS, IRD, France.
4* Climate Economics Chair (CEC) – IEF/ILB, France.

Acknowledgements: We thank Anna Creti, Guillaume Chevillon, Elena Dumistrescu, Christian Francq, Christian Gourieroux, Alain Hecq, Joann Jasiak, Sebastien Laurent, Aryan Manafi Neyazi, Daniel Velasquez-Gavarira and Jean-Michel Zakoian for many useful suggestions and detailed comments, as well as Dauphine LEDa's Seminar, LEO Economic Seminar, Maastricht Econometric Seminars, York University Econometric Seminar and conference participants at the 39th Canadian Econometric Study Group Meeting (CESG 2024), 6th Quantitative Finance and Financial Econometrics (QFFE 2024) and 44th International Symposium on Forecasting (ISF), 18th International Joint Conference on Computational and Financial Econometrics (CFE) and Computational and Methodological Statistics (CMStatistics) for their insights.

KEYWORDS

Prediction | Stable random vectors | Spectral representation | Pattern identification | Extreme weather event

Executive summary

Stochastic processes that depend on the "future" values of an independent and identically distributed (i.i.d.) sequence, often referred to as anticipative, have experienced growing interest. This interest has been driven by their convenience in modeling exotic patterns in time series, such as explosive bubbles in financial prices. These anticipative processes are well known as solutions to the rational expectations model and as a method to address non-fundamentalness in macroeconometric literature. More recently, they have been considered an effective tool for modeling climate variables associated with extreme weather events, such as global sea level, greenhouse gas (GHG) emissions, global temperature, sea ice area, and various natural ocean oscillation indices (e.g., the Southern Oscillation Index, SOI).

Due to climate change, the frequency and strength of these extreme weather events are increasing. Predicting these shocks is of primary interest, as it provides numerous societal benefits, from extreme weather warnings to agricultural planning. However, while the estimation aspect of these anticipative processes is well-covered, the appealing flexibility of these processes cannot yet be fully leveraged. Their dynamics, particularly the conditional distribution of future paths given the observed past trajectory, remain largely unexplored. This limits the applicability for applied prediction purposes.

In this paper, we propose a new representation for general two-sided stable MA processes using a semi-norm. The conditional distribution of future paths can then be explicitly derived using the regularly varying tails property of stable vectors and has a natural interpretation in terms of pattern identification. This leads to some interesting results: under certain persistence conditions in the anticipative part, the crash date or reversal date could be determined with certainty. This is highly relevant for climate variables and is linked to the concept of tipping points.

Through Monte Carlo simulations, we develop procedures to forecast crash probabilities and crash dates and demonstrate their finite sample performance. As an empirical illustration, we estimate reversal probabilities and dates of El Niño and La Niña occurrences.

To replicate the numerical and empirical results presented in the paper and to demonstrate the generality of our approach, we have developed a web application (link to the web app). This application enables users to replicate our findings, explore examples of simulated trajectories, and apply our methods to other time series, particularly in the fields of macroeconomics, finance, and climate science.

.

1 Introduction

Stochastic processes depending on the "future" values of an independent and identically distributed (i.i.d.) sequence, often referred to as *anticipative*, have witnessed a recent surge of attention from the statistical and econometric literature. Historically, the growing interest in this process has been driven by its convenience for modelling exotic patterns in time series, such as explosive bubbles in financial prices (Hencic and Gouriéroux, 2015; Fries and Zakoian, 2019; Gouriéroux and Zakoian, 2017; Hecq et al., 2017a; Cavaliere et al., 2017; Gouriéroux and Jasiak, 2017; Hecq et al., 2020; Fries, 2021; Hecq and Voisin, 2021). This anticipative process-known as the solution to the rational expectations model (Gourieroux et al., 2020)-has also been explored as a method to address non-fundamentalness in the macroeconometric literature (Lanne and Luoto, 2013; Gouriéroux et al., 2019; Chahrour and Jurado, 2021; Hecq et al., 2024) and as an effective tool for modelling climate variables associated with extreme weather events, such as global sea level, greenhouse gas (GHG) emissions, global temperature, sea ice area, and various natural ocean oscillation indices (e.g., the Southern Oscillation Index, SOI), see e.g., Giancaterini et al. (2022). Despite the fact that the estimation aspect is well-covered in both univariate and multivariate settings (see Andrews et al., 2009; Lanne and Saikkonen, 2011; Behme, 2011; Behme et al., 2011; Lanne and Saikkonen, 2013; Hecq et al., 2016; Chen et al., 2017; Fries and Zakoian, 2019; Hecq et al., 2017b; Cavaliere et al., 2017; Velasco and Lobato, 2018; Hecq et al., 2020; Gouriéroux and Jasiak, 2016, 2017; Gourieroux and Jasiak, 2023), the appealing flexibility of anticipative processes cannot yet be fully leveraged. Their dynamics—particularly the conditional distribution of future paths given the observed past trajectory—remain largely unexplored. The absence of closed-form formulas has prompted the literature to propose numerical approaches (see Lanne et al., 2012; Lanne and Luoto, 2016; Gourieroux et al., 2021a,b). Unfortunately, these algorithms, while based on attractive heuristics, lack strong theoretical foundations and offer no guarantees of soundness or accuracy. Moreover, as noted by Hecq and Voisin (2021), they are computationally intensive for larger prediction horizons and fail to accurately capture the dynamics during explosive episodes.

An exception is the anticipative α-stable AR(1) process, for which partial results were obtained in Gouriéroux and Zakoian (2017) and later completed in Fries (2021).[1] Even in the simplest case within the family of anticipative processes, future realizations exhibit a complex dependence on the observed past. This is reflected in the functional forms of the conditional

[1] Gourieroux and Jasiak (2024) proposed a closed-form solution for the predictive density within a general semi-parametric anticipative framework. However, their approach ultimately relies on numerical approximation using a kernel estimator.

moments, which are generally nonlinear. Interestingly, the dynamics of the anticipative stable AR(1) process simplify during extreme events, appearing to follow an explosive exponential path with a determined killing probability. This naturally raises the question of whether, and in what form, such behavior might occur in more general stable processes. For X_t a two-sided moving average, defined by

$$X_t = \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k} \tag{1.1}$$

where (ε_t) is an i.i.d. α-stable sequence and (d_k) is a sequence of non-random coefficients, this paper examines the conditional distribution of future paths—specifically, $(X_{t+1},\dots,X_{t+h})$ given the observed trajectory $(X_{t-m},\dots,X_t)$ with $m \geq 0$ and $h \geq 1$ —when the process deviates significantly from its central values. Only mild summability conditions are assumed on the sequence (d_k), and in particular, we do not make any assumptions upfront about the anticipativeness or non-anticipativeness of (X_t). Under this general framework, any vector of the form $\boldsymbol{X_t} = (X_{t-m},\dots,X_{t+h})$ is multivariate α-stable and its distribution is characterised by a unique finite measure Γ on the Euclidean unit sphere $S_{m+h+1} = \{\boldsymbol{s} \in \mathbb{R}^{m+h+1} : \ \|\boldsymbol{s}\|_e = 1\}$, where $\|\cdot\|_e$ denotes the Euclidean norm (Theorem 2.3.1 in Samorodnitsky and Taqqu (1994)). The measure Γ completely describes the conditional distribution of the normalized paths $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|_e$, which represents the "shape" of the trajectory, when $\boldsymbol{X_t}$ is large according to the Euclidean norm and given some information about the observed first $m+1$ entries. A straightforward application of Theorem 4.4.8 by Samorodnitsky and Taqqu (1994) indeed shows that

$$\mathbb{P}\Big(\boldsymbol{X}_t/\|\boldsymbol{X}_t\|_e \in A \ \Big|\ \|\boldsymbol{X}_t\|_e > x \ \text{ and } \ \boldsymbol{X}_t/\|\boldsymbol{X}_t\|_e \in B\Big) \underset{x\longrightarrow\infty}{\longrightarrow} \frac{\Gamma(A\cap B)}{\Gamma(B)}, \tag{1.2}$$

for any appropriately chosen Borel sets $A, B \subset S_{m+h+1}$. As such, however, (1.2) is of limited value for prediction purposes, where only $X_{t-m},\dots,X_t$ are assumed to be observed. This is because the conditioning generally depends on the future realizations $X_{t+1},\dots,X_{t+h}$, mainly through the Euclidean norm of $\boldsymbol{X}_t$. The idea developed here is to obtain a version of (1.2) where the Euclidean norm is replaced by a seminorm $\|\cdot\|$ satisfying

$$\|(x_{-m},\dots,x_0,x_1,\dots,x_h)\| = \|(x_{-m},\dots,x_0,0,\dots,0)\|, \tag{1.3}$$

for any $(x_{-m},\dots,x_h) \in \mathbb{R}^{m+h+1}$. In this context, a new representation of stable random vectors on the "unit cylinder" $C_{m+h+1}^{\|\cdot\|} := \{\boldsymbol{s} \in \mathbb{R}^{m+h+1} : \ \|\boldsymbol{s}\| = 1\}$ is explored, where $\|\cdot\|$ is such a seminorm. Contrary to representations involving norms (see Theorem 2.3.8 in Samorodnitsky and Taqqu (1994)), not all stable random vectors can be characterized on unit cylinders, and a representation theorem is provided. It is shown that $\boldsymbol{X_t}$ admits a representation by a measure

$\Gamma^{|\cdot|}$ on $C^{|\cdot|}_{m+h+1}$ if and only if (X_t) is "anticipative enough". The equation (1.2) is then shown to hold with an adequate seminorm and with Γ (resp. S_{m+h+1}) replaced by $\Gamma^{\|\cdot\|}$ (resp. $C^{\|\cdot\|}_{m+h+1}$). The problem finally comes down to choosing the appropriate Borels B in (1.2) reflecting that only the past "shape" $(X_{t-m},\dots,X_t)/\|\boldsymbol{X_t}\|$ is observed.

The use of (1.2) to infer the future paths of (X_t) is related to the so-called *spectral process* introduced by Basrak and Segers (2009), which has opened a fruitful line of research (see, for instance, Basrak et al. (2016); Dombry et al. (2017); Janßen (2019); Janßen and Segers (2014); Meinguet and Segers (2010); Planinić and Soulier (2017)). This spectral process is defined as the limit in distribution of a vector of observations from a multivariate regularly varying time series, conditionally on the first observation being large. The approach followed here differs in that it operates at the representation level of α-stable vectors, establishing a link between the spectral representation and the tail conditional distribution of stable linear processes, while shedding light on the (un)predictability of their extremes. A natural interpretation of path prediction in terms of pattern identification emerges from (1.2) when applied to stable linear processes.

Section 2 characterizes the representation of general α-stable vectors on seminorm unit cylinders and shows that (1.2) can be restated under this new representation. Focusing on α-stable processes defined by 1.1, Section 3 studies under which condition the vector $(X_{t-m},\dots,X_{t+h})$ admits a representation on the unit cylinder $C^{\|\cdot\|}_{m+h+1}$. The anticipativeness of (X_t) interestingly emerges as a necessary condition for such a representation to exist. Section 4 leverages (1.2) to analyze the tail conditional distribution of specific processes: the anticipative AR(1), AR(2), and the anticipative fractionally integrated process. For the last two processes, it appears possible to determine, in theory with certainty, the dates of extreme events, provided the current pattern is properly identified. Section 5 presents a set of Monte Carlo simulations that illustrate how this pattern identification can be practically applied to accurately predict future trajectories. In particular, we propose simple procedures to forecast crash probabilities and determine crash dates. In Section 6, we demonstrate the empirical relevance of our theoretical results through a climate forecasting exercise. Specifically, we predict the occurrences of El Niño and La Niña, as well as their reversal dates, using the Southern Oscillation Index (SOI) data.[2] To replicate the numerical and empirical results presented in the paper and to demonstrate the generality of our approach, we have developed a web application. This application enables users to replicate our findings, explore examples of simulated trajectories, and apply our methods to

[2] Data and methodology are available at: https://www.ncei.noaa.gov/access/monitoring/enso/soi

other time series, particularly in the fields of macroeconomics, finance, and climate science.[3] Section 7 concludes the paper. Complementary empirical results are provided in Appendix B, and the proofs are collected in Appendix C.

2 Representation of stable random vectors on unit cylinders

We start with a review of the characterization of stable random vectors on the Euclidean unit sphere. First, we recall that a random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ is defined to be a stable random vector in $\mathbb{R}^d$ if and only if, for any positive numbers A and B there is a positive number C and a deterministic vector $\boldsymbol{D} \in \mathbb{R}^d$ such that

$$A\boldsymbol{X}^{(1)} + B\boldsymbol{X}^{(2)} \stackrel{d}{=} C\boldsymbol{X} + \boldsymbol{D},$$

where $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ are independent copies of $\boldsymbol{X}$. Furthermore, if $\boldsymbol{X}$ is stable, then there exists a real number $\alpha \in (0, 2]$ such that the above relation holds with $C = (A^\alpha + B^\alpha)^{1/\alpha}$, and $\boldsymbol{X}$ is referred to as α-stable. The Gaussian case ($\alpha = 2$) is henceforth excluded.

For, $\boldsymbol{X}$ an α-stable random vector with $0 < \alpha < 2$, one can define a unique pair $(\Gamma, \boldsymbol{\mu}^0)$, where Γ is a finite measure on S_d and $\boldsymbol{\mu}^0$ a non-random vector in $\mathbb{R}^d$, such that,

$$\mathbb{E}\left[e^{i\langle \boldsymbol{u}, \boldsymbol{X}\rangle}\right] = \exp\left\{ -\int_{S_d} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha \Big(1 - i\,\mathrm{sign}(\langle \boldsymbol{u}, \boldsymbol{s}\rangle) w(\alpha, \langle \boldsymbol{u}, \boldsymbol{s}\rangle)\Big)\Gamma(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \right\}, \quad \forall \boldsymbol{u} \in \mathbb{R}^d, \tag{2.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the canonical scalar product, $w(\alpha, s) = \mathrm{tg}\left(\frac{\pi\alpha}{2}\right)$, if $\alpha \neq 1$, and $w(1, s) = -\frac{2}{\pi}\ln|s|$ otherwise, for $s \in \mathbb{R}$. The pair $(\Gamma, \boldsymbol{\mu}^0)$ is called the *spectral representation* of the stable vector $\boldsymbol{X}$, Γ is its *spectral measure* and $\boldsymbol{\mu}^0$ its *shift vector*. In particular, $\boldsymbol{X}$ is symmetric if and only if $\boldsymbol{\mu}^0 = 0$ and $\Gamma(A) = \Gamma(-A)$ for any Borel set A in S_d (Theorem 2.4.3 in Samorodnitsky and Taqqu (1994)), and in that case

$$\mathbb{E}\left[e^{i\langle \boldsymbol{u}, \boldsymbol{X}\rangle}\right] = \exp\left\{ -\int_{S_d} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha \Gamma(d\boldsymbol{s}) \right\}, \quad \forall \boldsymbol{u} \in \mathbb{R}^d. \tag{2.2}$$

The representations (2.1) and (2.2) of a stable random vector involves integration over all directions of $\mathbb{R}^d$,[4] parameterized here by the unit sphere with respect to the Euclidean norm. Proposition 2.3.8 in Samorodnitsky and Taqqu (1994) shows that the unit sphere with respect to any norm can be used instead, provided a change of spectral measure and shift vector.

[3] The web application is available at the following link: https://marforecast.streamlit.app/?utm_medium=oembed

[4] By *direction* of $\mathbb{R}^d$, we mean the equivalence classes of the relation "$\equiv$" defined by: $\boldsymbol{u} \equiv \boldsymbol{v}$ if and only if there exists $\lambda > 0$ such that $\boldsymbol{u} = \lambda\boldsymbol{v}$, for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$.

We explore alternative representations where integration is carried out over a unit cylinder relative to a seminorm. However, for a given seminorm, not all stable vectors admit such a representation, motivating the following definition.

Definition 2.1 *Let $\|\cdot\|$ be a seminorm on $\mathbb{R}^d$, $C_d^{\|\cdot\|} := \{\boldsymbol{s} \in \mathbb{R}^d : \ \|\boldsymbol{s}\| = 1\}$ be the corresponding unit cylinder, and let $\boldsymbol{X} = (X_1, \ldots, X_d)$ be an α-stable random vector.*

(ι) *Asymmetric case:*

In the case where $\boldsymbol{X}$ is not symmetric, we say that $\boldsymbol{X}$ is representable on $C_d^{\|\cdot\|}$ if there exists a non-random vector $\boldsymbol{\mu}^0_{\|\cdot\|} \in \mathbb{R}^d$ and a Borel measure $\Gamma^{\|\cdot\|}$ on $C_d^{\|\cdot\|}$ satisfying for all $\boldsymbol{u} \in \mathbb{R}^d$

$$\int_{C_d^{\|\cdot\|}} |\langle \boldsymbol{u}, \boldsymbol{s} \rangle|^\alpha \Gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty, \tag{2.3}$$

if $\alpha \neq 1$, and if $\alpha = 1$,

$$\int_{C_d^{\|\cdot\|}} |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \left| \ln |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \right| \Gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty, \tag{2.4}$$

such that the joint characteristic function of $\boldsymbol{X}$ can be written as in (2.1) with $(S_d, \Gamma, \boldsymbol{\mu}^0)$ replaced by $(C_d^{\|\cdot\|}, \Gamma^{\|\cdot\|}, \boldsymbol{\mu}^0_{\|\cdot\|})$.

($\iota\iota$) *Symmetric case:*

In the case where $\boldsymbol{X}$ is symmetric α-stable ($S\alpha S$), $0 < \alpha < 2$, we say that $\boldsymbol{X}$ is representable on $C_d^{\|\cdot\|}$ if there exists a symmetric Borel measure $\Gamma^{\|\cdot\|}$ on $C_d^{\|\cdot\|}$ satisfying (2.3) such that the joint characteristic function of $\boldsymbol{X}$ can be written as in (2.2) with (S_d, Γ) replaced by $(C_d^{\|\cdot\|}, \Gamma^{\|\cdot\|})$.

The integrability conditions (2.3)-(2.4) ensure the validity of the above definition, as they account for the unbounded nature of unit cylinders. The following proposition begins by characterizing stable random vectors that can be represented on a unit cylinder defined by a given seminorm.

Proposition 2.1 *Let $\|\cdot\|$ be a seminorm on $\mathbb{R}^d$ and $C_d^{\|\cdot\|}$ be the corresponding unit cylinder. Denote $K^{\|\cdot\|} = \{x \in S_d : \|x\| = 0\}$. Let also $\boldsymbol{X}$ be an α-stable random vector on $\mathbb{R}^d$ with spectral representation $(\Gamma, \boldsymbol{\mu}^0)$ on the Euclidean unit sphere (with $\boldsymbol{\mu}^0 = 0$ if $\boldsymbol{X}$ is $S\alpha S$). If $\alpha \neq 1$ or if $\boldsymbol{X}$ is S1S, then*

$$\boldsymbol{X} \text{ is representable on } C_d^{\|\cdot\|} \iff \Gamma(K^{\|\cdot\|}) = 0.$$

If $\alpha = 1$ and $\boldsymbol{X}$ is not symmetric, then

$$\boldsymbol{X} \text{ is representable on } C_d^{\|\cdot\|} \iff \int_{S_d} \left| \ln \|\boldsymbol{s}\| \right| \Gamma(d\boldsymbol{s}) < +\infty.$$

Moreover, if $\boldsymbol{X}$ *is representable on* $C_d^{\|\cdot\|}$*, its spectral representation is then given by* $(\Gamma^{\|\cdot\|}, \boldsymbol{\mu}_{\|\cdot\|}^0)$ *where*

$$\Gamma^{\|\cdot\|}(d\boldsymbol{s}) = \|\boldsymbol{s}\|_e^{-\alpha}\, \Gamma \circ T_{\|\cdot\|}^{-1}(d\boldsymbol{s})$$

with $T_{\|\cdot\|} : S_d \setminus K^{\|\cdot\|} \longrightarrow C_d^{\|\cdot\|}$ *defined by* $T_{\|\cdot\|}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|$*, and*

$$\boldsymbol{\mu}_{\|\cdot\|}^0 = \begin{cases} \boldsymbol{\mu}^0, & \text{if} \quad \alpha \neq 1 \quad \text{or} \quad \text{if} \quad \boldsymbol{X} \quad \text{is} \quad S1S, \\ \boldsymbol{\mu}^0 + \tilde{\boldsymbol{\mu}}, & \text{if} \quad \alpha = 1 \quad \text{and} \quad \boldsymbol{X} \quad \text{is not symmetric}, \end{cases}$$

$$\tilde{\boldsymbol{\mu}} = (\tilde{\mu}_j), \qquad \text{and} \qquad \tilde{\mu}_j = -\frac{2}{\pi} \int_{S_d \setminus K^{\|\cdot\|}} s_j \ln \|\boldsymbol{s}\| \Gamma(d\boldsymbol{s}), \quad j = 1, \ldots, d.$$

It can be noticed from Proposition 2.1 that the representability condition in the case $\alpha = 1$ and $\boldsymbol{X}$ is not symmetric, is slightly stronger than that in the other cases. Indeed, $\int_{K^{\|\cdot\|}} \left| \ln \|\boldsymbol{s}\| \right| \Gamma(d\boldsymbol{s}) \leq \int_{S_d} \left| \ln \|\boldsymbol{s}\| \right| \Gamma(d\boldsymbol{s}) < +\infty$ necessarily implies that $\Gamma(K^{\|\cdot\|}) = 0$ since $\left| \ln \|\boldsymbol{s}\| \right| = +\infty$ for $\boldsymbol{s} \in K^{\|\cdot\|}$.

Remark 2.1 The case $d = 2$ is particularly insightful. In light of (1.2), the spectral measure of the α-stable vector (X_1, X_2) characterizes its likelihood of being in any particular direction of $\mathbb{R}^2$ when it is large in norm. Since unit spheres relative to norms encompass all directions in $\mathbb{R}^2$, spectral measures on such spheres can capture any potential tail dependence of (X_1, X_2). However, unit cylinders do not span all directions in $\mathbb{R}^2$ and therefore, spectral measures on such cylinders encode less information. Consider for instance the unit cylinder $C_2^{\|\cdot\|} = \{(s_1, s_2) \in \mathbb{R}^2 : \ |s_1| = 1\}$ associated to the seminorm such that $\|(x_1, x_2)\| = |x_1|$ for all $(x_1, x_2) \in \mathbb{R}^2$. It is straightforward to observe that $C_2^{\|\cdot\|}$ spans all directions of $\mathbb{R}^2$ except for those of $(0, -1)$ and $(0, +1)$. A stable vector (X_1, X_2) will admit a representation on $C_2^{\|\cdot\|}$ if these directions are irrelevant for characterizing its distribution, that is, if $\Gamma\Big(\{(0,-1),(0,+1)\}\Big) = 0$. In terms of tail dependence, the latter condition intuitively implies that realizations (X_1, X_2) where X_2 is extreme and X_1 is not, almost never occur (i.e., they occur with probability zero). The conditions $\Gamma(\{(0,-1),(0,+1)\}) = 0$ and $\int_{S_2} |\ln \|\boldsymbol{s}\||\Gamma(d\boldsymbol{s}) < +\infty$ can also be related to the stronger condition ensuring the existence of conditional moments of X_2 given X_1 as discussed in Cioczek-Georges and Taqqu (1994, 1998) - see also Theorem 5.1.3 in Samorodnitsky and Taqqu (1994). This stronger condition requires that Γ is not too concentrated around the points $(0, \pm 1)$. Specifically, assuming $\int_{S_2} |s_1|^{-\nu} \Gamma(ds) < +\infty$ for some $\nu \geq 0$, then $\mathbb{E}[|X_2|^\gamma | X_1] < +\infty$ for $\gamma < \min(\alpha + \nu, 2\alpha + 1)$, even though $\mathbb{E}[|X_2|^\alpha] = +\infty$. If this condition holds for some $\nu > 0$, then both of the aforementioned conditions must necessarily be satisfied.

Provided that the appropriate representation exists, (1.2) holds with seminorms instead of norms. This forms the foundation for studying the tail conditional distribution of stable processes and leads to the following proposition:

Proposition 2.2 *Let $\boldsymbol{X} = (X_1, \dots, X_d)$ be an α-stable random vector and let $\|\cdot\|$ be a seminorm on $\mathbb{R}^d$. If $\boldsymbol{X}$ is representable on $C_d^{\|\cdot\|}$, then for every Borel sets $A, B \subset C_d^{\|\cdot\|}$ with $\Gamma^{\|\cdot\|}\Big(\partial(A \cap B)\Big) = \Gamma^{\|\cdot\|}(\partial B) = 0$, and $\Gamma^{\|\cdot\|}(B) > 0$,*

$$\mathbb{P}_x^{\|\cdot\|}(\boldsymbol{X}, A|B) \underset{x\to+\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}(A \cap B)}{\Gamma^{\|\cdot\|}(B)}, \tag{2.5}$$

where ∂B (resp. $\partial(A \cap B)$) denotes the boundary of B (resp. $A \cap B$), and

$$\mathbb{P}_x^{\|\cdot\|}(\boldsymbol{X}, A|B) := \mathbb{P}\Big(\frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A \Big| \|\boldsymbol{X}\| > x, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in B\Big).$$

3 Unit cylinder representation for the paths of stable linear processes

Given a seminorm, Proposition 2.2 is only applicable to stable vectors that are representable on the corresponding unit cylinder. This section investigates under which condition on an stable moving average (X_t) vectors of the form $(X_{t-m}, \dots, X_t, X_{t+1}, \dots, X_{t+h})$ admit such representations. A characterization is proposed and extended to linear combinations of stable moving averages. Any seminorm satisfying (1.3) could be relevant for the prediction framework mentioned in the introduction. However, to simplify the discussion and avoid considering numerous cases for all possible kernels, we restrict attention to seminorms such that

$$\|(x_{-m}, \dots, x_0, x_1, \dots, x_h)\| = 0 \iff x_{-m} = \dots = x_0 = 0, \tag{3.1}$$

for any $(x_{-m}, \dots, x_h) \in \mathbb{R}^{m+h+1}$, which in particular satisfy (1.3). As an example, seminorms on $\mathbb{R}^{m+h+1}$ satisfying (3.1) can be naturally derived from norms on the $m+1$ first components of vectors. For any $p \in [1, +\infty]$, one can, for instance, consider seminorms of the form $\|\cdot\|$ defined by

$$\|(x_{-m}, \dots, x_0, x_1, \dots, x_h)\| = \Big(\sum_{i=-m}^{0} |x_i|^p\Big)^{1/p},$$

for any $(x_{-m}, \dots, x_0, x_1, \dots, x_h) \in \mathbb{R}^{m+h+1}$ with by convention $\big(\sum_{i=-m}^{0} |x_i|^p\big)^{1/p} = \sup_{-m\le i\le 0} |x_i|$ for $p = +\infty$.

Now, assuming any seminorm satisfying (1.3) and (3.1), we consider (X_t) the α-stable moving average defined by

$$X_t = \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k}, \qquad \varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta, \sigma, 0) \tag{3.2}$$

with (d_k) a real deterministic sequence such that

$$\text{if} \quad \alpha \neq 1 \quad \text{or} \quad (\alpha, \beta) = (1, 0), \qquad 0 < \sum_{k\in\mathbb{Z}} |d_k|^s < +\infty, \qquad \text{for some} \quad s \in (0, \alpha) \cap [0, 1], \tag{3.3}$$

and

$$\text{if} \quad \alpha = 1 \quad \text{and} \quad \beta \neq 0, \qquad 0 < \sum_{k\in\mathbb{Z}} |d_k| \Big| \ln |d_k| \Big| < +\infty. \tag{3.4}$$

Letting for $m \geq 0$, $h \geq 1$,

$$\boldsymbol{X_t} = (X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h}), \tag{3.5}$$

it follows from Proposition 13.3.1 in Brockwell and Davis (1991) that the infinite series converge almost surely and both (X_t) and $\boldsymbol{X}_t$ are well-defined and the random vector $\boldsymbol{X_t}$ is multivariate α-stable. Denoting $\boldsymbol{d}_k := (d_{k+m}, \ldots, d_k, d_{k-1}, \ldots, d_{k-h})$ for $k \in \mathbb{Z}$, the spectral representation of $\boldsymbol{X_t}$ on the Euclidean sphere is given by $(\Gamma, \boldsymbol{\mu}^0)$ with

$$\begin{aligned} \Gamma &= \sigma^\alpha \sum_{\vartheta\in S_1} \sum_{k\in\mathbb{Z}} w_\vartheta \|\boldsymbol{d}_k\|_e^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|_e}\right\}}, \\ \boldsymbol{\mu}^0 &= -\mathbb{1}_{\{\alpha=1\}} \frac{2}{\pi} \beta\sigma \sum_{k\in\mathbb{Z}} \boldsymbol{d}_k \ln \|\boldsymbol{d}_k\|_e, \end{aligned} \tag{3.6}$$

where $w_\vartheta = (1+\vartheta\beta)/2$, $S_1 = \{-1, +1\}$, δ is the Dirac mass and by convention, if for some $k \in \mathbb{Z}$, $\boldsymbol{d}_k = \boldsymbol{0}$, i.e. $\|\boldsymbol{d}_k\|_e = 0$, then the kth term vanishes from the sums. Notice in particular that for $\beta = 0$, it holds that $w_{-1} = w_{+1} = 1/2$, $\boldsymbol{\mu}^0 = \boldsymbol{0}$, and both the measure Γ and the random vector $\boldsymbol{X}_t$ are symmetric. The following representation theorem characterizes the representability of $\boldsymbol{X_t}$ on a unit cylinder for fixed m and h.

Theorem 3.1 *Let $\boldsymbol{X_t}$ satisfy (3.2)-(3.5) and let $\|\cdot\|$ be a seminorm on $\mathbb{R}^{m+h+1}$ satisfying (3.1). For $\alpha \neq 1$ or $(\alpha, \beta) = (1, 0)$, the vector $\boldsymbol{X_t}$ is representable on $C^{\|\cdot\|}_{m+h+1}$ if and only if*

$$\forall k \in \mathbb{Z}, \quad \Big[(d_{k+m}, \ldots, d_k) = \boldsymbol{0} \quad \Longrightarrow \quad \forall \ell \leq k-1, \quad d_\ell = 0\Big]. \tag{3.7}$$

For $\alpha = 1$ and $\beta \neq 0$, the vector $\boldsymbol{X_t}$ is representable on $C^{\|\cdot\|}_{m+h+1}$ if and only if in addition to (3.7), it holds that

$$\sum_{k\in\mathbb{Z}} \|\boldsymbol{d}_k\|_e \left| \ln\left(\|\boldsymbol{d}_k\|/\|\boldsymbol{d}_k\|_e\right)\right| < +\infty. \tag{3.8}$$

In the cases, $\alpha \neq 1$ and $(\alpha, \beta) = (1, 0)$, the representability of $\boldsymbol{X_t}$ on a seminorm unit cylinder depends on the number of observation $m + 1$ but not on the prediction horizon h.[5] This is particularly important for the applicability of Proposition 2.2 when studying the conditional dynamics of a given stable moving average. Indeed, this proposition relies on seminorms instead of norms, meaning that we use only a portion of size $m + 1$ of an observed path. This leads to the notion of *past-representability*, stated by the following definition:

[5] The case $\alpha = 1$, $\beta \neq 0$ is more intricate, as the roles of m and h in the validity of the additional requirement (3.8) are not as clear-cut.

Definition 3.1 Let (X_t) be an α-stable moving average satisfying (3.2)-(3.4). We say that the stable process (X_t) is *past-representable* if there exists at least one pair (m, h), $m \geq 0$, $h \geq 1$, such that $\boldsymbol{X_t} = (X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$ is representable on $C^{\|\cdot\|}_{m+h+1}$ for any seminorm satisfying (3.1). For any such pair (m, h), we will say that (X_t) is (m, h)-past-representable.

Definition 3.1, holds for any seminorm satisfying (3.1) because it ensures that all these seminorms have the same kernel. Moreover, it is straightforward to observe that this definition holds for any $m' \geq m$, because if (3.7) is true for some $m \geq 0$, it is true for any $m' \geq m$. Thus, the notion of past-representability can be defined independently of the particular choice of a seminorm but relies on the existence of at least one m, for which the considered process is (m, h)-past-representable.[6] The following proposition provides this characterization.

Proposition 3.1 *Let (X_t) be an α-stable moving average satisfying (3.2)-(3.4).*

(ι) *With the set $\mathcal{M} = \{m \geq 1 : \ \exists k \in \mathbb{Z}, \ d_{k+m} = \ldots = d_{k+1} = 0, \ d_k \neq 0\}$, define*

$$m_0 = \begin{cases} \sup \mathcal{M}, & \text{if} \quad \mathcal{M} \neq \emptyset, \\ 0, & \text{if} \quad \mathcal{M} = \emptyset. \end{cases} \tag{3.9}$$

(a) *For $\alpha \neq 1$ or $(\alpha, \beta) = (1, 0)$, the process (X_t) is past-representable if and only if*

$$m_0 < +\infty. \tag{3.10}$$

Moreover, letting $m \geq 0$, $h \geq 1$, the process (X_t) is (m, h)-past-representable if and only if (3.10) holds and $m \geq m_0$.

(b) *For $\alpha = 1$ and $\beta \neq 0$, the process (X_t) is past-representable if and only if in addition to (3.10), there exist an $m \geq m_0$ and an $h \geq 1$ such that (3.8) holds. If such a pair (m, h) exists, (X_t) is then (m, h)-past-representable.*

($\iota\iota$) *Let $\|\cdot\|$ a seminorm satisfying (3.1) and assume that (X_t) is (m, h)-past-representable for some $m \geq 0$, $h \geq 1$. The spectral representation $(\Gamma^{\|\cdot\|}, \boldsymbol{\mu}^{\|\cdot\|})$ of the vector $\boldsymbol{X_t} = (X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$ on $C^{\|\cdot\|}_{m+h+1}$ is then given by (3.6) with the Euclidean norm $\|\cdot\|_e$ replaced by the seminorm $\|\cdot\|$.*

It can be noted, that $m_0 = 0$ if and only if for some $k_0 \in \mathbb{Z} \cup \{-\infty\}$, $d_k \neq 0$ for all $k \geq k_0$ and $d_k = 0$ for all $k < k_0$.

Proposition 3.1 shows that for an α-stable moving average to be past-representable, sequences

[6] This will not be true in general under the weaker assumption (1.3) and different notions of representability of a process could emerge depending on the kernels of the seminorms.

of consecutive zero values in the coefficients (d_k) must be either of finite length or infinite to the right. This surprisingly makes the anticipativeness of a stable moving average a necessary condition, and sufficient for $\alpha \neq 1$ and $(\alpha, \beta) = (1, 0)$, for its past-representability. The less anticipative a moving average is (i.e., the larger the gaps of zeros on its forward-looking side), the higher m must be chosen to ensure the representability of $(X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$ on the appropriate unit cylinder. Purely non-anticipative moving averages are, in particular, immediately ruled out. This property is demonstrated in the following corollary.

Corollary 3.1 Let (X_t) an α-stable moving average satisfying (3.2)-(3.4). If (X_t) is purely non-anticipative, i.e., $d_k = 0$ for all $k \geq 1$, then (X_t) is not past-representable.

Corollary 3.1, sheds light on the predictability of extreme events in linear processes. For illustration, we consider the following two α-stable AR(1) processes, defined as the stationary solutions of:

$$X_t = \rho X_{t+1} + \varepsilon_t, \qquad \forall t \in \mathbb{Z}, \tag{3.11}$$

$$Y_t = \rho Y_{t-1} + \eta_t, \qquad \forall t \in \mathbb{Z}, \tag{3.12}$$

where $0 < |\rho| < 1$, and (ε_t), (η_t) are independent i.i.d. stable sequences. While (X_t) generates bubble-like trajectories –explosive exponential paths eventually followed by sharp returns to central values–, the trajectories of (Y_t) feature sudden jumps followed by exponential decays. In both processes, an extreme event results from a large realisation of an underlying error ε_τ or η_τ, at some time τ. On the one hand, for the non-anticipative AR(1) process (3.12), a jump does not manifest any early visible signs before its date of occurrence, as it is independent of the past trajectory. Jumps in the trajectory of (Y_t) are unpredictable, and one only has information about their unconditional likelihood of occurrence. On the other hand, for the anticipative AR(1) process (3.11), extremes do manifest early visible signs and are gradually reached as their occurrence dates approach. The past trajectory is informative about future extreme events, and, in particular, it is more informative than their plain unconditional likelihood of occurrence. Building on the "information encoding" interpretation of spectral measures given in Remark 2.1, the fact that (X_t) (resp. (Y_t)) is past-representable (resp. not past-representable) can be seen as a consequence of the dependence (resp. independence) of future extreme events on past ones. The following Corollary 3.2 states the condition for past-representability for ARMA.

Corollary 3.2 Let (X_t) be the strictly stationary solution of

$$\Psi(F)\Phi(B)X_t = \Theta(F)H(B)\varepsilon_t, \qquad \varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta, \sigma, 0),$$

where Ψ, Φ, Θ, H are polynomials of arbitrary finite degrees with roots located outside the unit disk and F (resp. B) is the forward (resp. backward) operator: $FX_t := X_{t+1}$ (resp. $BX_t := X_{t-1}$). We suppose furthermore that Ψ and Θ (resp. Φ and H) have no common roots. Then, for any $\alpha \in (0,2)$ and $\beta \in [-1,1]$, the following statements are equivalent:

(ι) (X_t) is past-representable,

$(\iota\iota)$ $\deg(\Psi) \geq 1$,

$(\iota\iota\iota)$ $m_0 < +\infty$,

with m_0 as in (3.9). Moreover, letting $m \geq 0$, $h \geq 1$, the process (X_t) is (m,h)-past-representable if and only if $m \geq m_0$ with $m_0 < +\infty$.

For the ARMA process, the past-representability condition simplifies and is equivalent to the autoregressive polynomial having at least one root inside the unit circle. Also, only the roots of the AR polynomial matter for past-representability, the MA part having no role. In the noncausal literature, if $\Theta = H = 1$, it is called a mixed-causal process $MAR(p,q)$ model, first introduced by Lanne and Saikkonen (2011), where p corresponds to $\deg(\Phi)$ and q corresponds to $\deg(\Psi)$.

4 Tail conditional distribution of stable anticipative processes

In this section, we will derive the tail conditional distribution of linear stable processes for which Proposition 2.2 will be applicable. The case of a general past-representable stable process is considered, along with particular examples.

To be relevant for the prediction framework, the Borel set B appearing in Proposition 2.2 must be chosen such that the conditioning event $\{\|\boldsymbol{X_t}\| > x\} \cap \{\boldsymbol{X_t}/\|\boldsymbol{X_t}\| \in B\}$ is independent of the future realizations $X_{t+1}, \ldots, X_{t+h}$. For $\|\cdot\|$ a seminorm on $\mathbb{R}^{m+h+1}$ satisfying (3.1), denote $S_{m+1}^{\|\cdot\|} = \{(s_{-m}, \ldots, s_0) \in \mathbb{R}^{m+1} : \|(s_{-m}, \ldots, s_0, 0, \ldots, 0)\| = 1\}$.[7] Then, for any Borel set $V \subset S_{m+1}^{\|\cdot\|}$, define the Borel set $B(V) \subset C_{m+h+1}^{\|\cdot\|}$ as

$$B(V) = V \times \mathbb{R}^h.$$

Notice in particular that for $V = S_{m+1}^{\|\cdot\|}$, we have $B(V) = C_{m+1}^{\|\cdot\|} = S_{m+1}^{\|\cdot\|}$. In the following, we will use Borel sets of the above form to condition the distribution of the entire vector $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ on the observed "shape" of the past trajectory. The latter information is contained in the Borel set V, which we will typically assume to be a small neighborhood on $S_{m+1}^{\|\cdot\|}$. It will be useful in

[7] The set $S_{m+1}^{\|\cdot\|}$ corresponds to the unit sphere of $\mathbb{R}^{m+1}$ relative to the restriction of $\|\cdot\|$ to the first $m+1$ dimensions.

the following to note that

$$V \times \mathbb{R}^h = \left\{ \boldsymbol{s} \in C_{m+h+1}^{\|\cdot\|} : \quad f(\boldsymbol{s}) \in V \right\},$$

where f is the function defined by

$$f : \begin{array}{ccc} \mathbb{R}^{m+h+1} & \longrightarrow & \mathbb{R}^{m+1} \\ (x_{-m}, \ldots, x_0, x_1, \ldots, x_h) & \longmapsto & (x_{-m}, \ldots, x_0) \end{array}. \tag{4.1}$$

4.1 Stable past-representable processes: general case

Let (X_t) an α-stable process satisfying Definition 3.1. This states that (X_t) is (m, h)-past-representable, for some $m \geq 0$, $h \geq 1$ and let $\boldsymbol{X_t}$ as in (3.5). Denoting $\Gamma^{\|\cdot\|}$ the spectral measure of $\boldsymbol{X}_t$ on the unit cylinder $C_{m+h+1}^{\|\cdot\|}$ for some seminorm satisfying (3.1), we know by Proposition 3.1 $(\iota\iota)$, that $\Gamma^{\|\cdot\|}$ is of the form

$$\Gamma^{\|\cdot\|} = \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_\vartheta \|\boldsymbol{d}_k\|^\alpha \delta_{\left\{ \frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \right\}}. \tag{4.2}$$

leading to the following proposition:

Proposition 4.1 *Under the above assumptions, we have*

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A \Big| B(V)\Big) \underset{x \to +\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}\left(\left\{ \frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \in A : \quad \frac{\vartheta f(\boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} \in V \right\} \right)}{\Gamma^{\|\cdot\|}\left(\left\{ \frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \in C_{m+h+1}^{\|\cdot\|} : \quad \frac{\vartheta f(\boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} \in V \right\} \right)}, \tag{4.3}$$

for any Borel sets $A \subset C_{m+h+1}^{\|\cdot\|}$, $V \subset S_{m+1}^{\|\cdot\|}$ *such that* $\left\{ \frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \in C_{m+h+1}^{\|\cdot\|} : \quad \frac{\vartheta f(\boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} \in V \right\} \neq \emptyset$, $\Gamma^{\|\cdot\|}\Big(\partial(A \cap B(V))\Big) = \Gamma^{\|\cdot\|}(\partial B(V)) = 0$, *where* $B(V) = V \times \mathbb{R}^h$ *and* f *is as in* (4.1).

From Proposition 4.1 and by setting $V = S_{m+1}^{\|\cdot\|}$, and A an arbitrarily small closed neighborhood of all the points $(\vartheta \boldsymbol{d}_k / \|\boldsymbol{d}_k\|)_{\vartheta,k}$, we can see that $\lim_{x \to +\infty} \mathbb{P}\Big(\boldsymbol{X_t}/\|\boldsymbol{X_t}\| \in A \Big| \|\boldsymbol{X_t}\| > x\Big) = 1$. In other words, when far from central values, the trajectory of the process (X_t) necessarily features patterns of the same shape as some $\vartheta \boldsymbol{d}_k / \|\boldsymbol{d}_k\|$, which is a finite piece of a moving average's coefficient sequence. The index k indicates to which piece $(d_{k+m}, \ldots, d_k, d_{k-1}, \ldots, d_{k-h})$ of this moving average it corresponds to, and, $\vartheta \in \{-1, +1\}$ indicates whether the pattern is flipped upside down (in case the extreme event is driven by a negative value of an error ε_τ). The likelihood of a pattern $\vartheta \boldsymbol{d}_k / \|\boldsymbol{d}_k\|$ can be evaluated by setting A to be a small neighborhood of that point. From this viewpoint, the observed path, $(X_{t-m}, \ldots, X_{t-1}, X_t)/\|\boldsymbol{X_t}\|$ will *a fortiori* be of the same shape as some $\vartheta(d_{k+m}, \ldots, d_{k+1}, d_k)/\|\boldsymbol{d}_k\|$ when an extreme event approaches in time. Observing the initial part of the pattern can provide information about the remaining

unobserved part: the conditional likelihood of the latter can be assessed by setting V to be a small neighborhood of the observed pattern.

Remark 4.1 The tail conditional distribution given in (4.3) highlights three types of uncertainty/approximation for prediction. We deal with two of them in the rest of the paper and leave the third for future research. [8]

(ι) In practice, events of the type

$$\{(X_{t-m},\dots,X_{t-1},X_t)/\|\boldsymbol{X_t}\| = \vartheta(d_{k+m},\dots,d_{k+1},d_k)/\|\boldsymbol{d}_k\|\}$$

have probability zero of occurring, and only noisy observations such as $(X_{t-m},\dots,X_{t-1},X_t)/\|\boldsymbol{X_t}\| \approx \vartheta(d_{k+m},\dots,d_{k+1},d_k)/\|\boldsymbol{d}_k\|$ are available on a realized trajectory. The choice of an adequate conditioning neighborhood V in (4.3) given a piece of the trajectory, will thus need to rely on a statistical approach. One could envision hypothesis tests to determine whether a piece of the realized (noisy) trajectory "is more similar" to a certain pattern 1 or to another pattern 2.

($\iota\iota$) Even if the observed path can be confidently identified with a particular pattern, uncertainty regarding the future trajectory may remain. It could indeed be that several patterns $\vartheta\boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ coincide on their first $m+1$ components, but differ in the last h components. The stable anticipative AR(1) process is a typical example of this phenomenon, which will be studied in the next section. Interestingly, the stable anticipative AR(2) process, and more generally persistent enough processes, avoid this issue, as discussed hereafter.

($\iota\iota\iota$) The tail conditional distribution (4.3) represents an asymptotic behavior as the (semi-) norm of $\boldsymbol{X_t}$ grows infinitely large. It is, therefore, only an approximation of the true dynamics during extreme events. It would be interesting to obtain a more refined asymptotic development in x of the above convergence in order to gauge the approximation error of the true conditional distribution. We partially investigate this issue by means of Monte Carlo simulation. In particular, we numerically quantify how far from„ the predicted patterns the future path can be, depending on the quantile of the observed realizations.

[8] The considerations developed in this remark focus solely on the probabilistic uncertainty of the prediction, assuming that the process (X_t) is entirely known; that is, no parameter nor any sequence $(d_{j,k})$ needs to be inferred from the data. We leave for future research issues related to statistical uncertainty.

4.2 The anticipative AR(1)

We now consider (X_t) the stable anticipative AR(1) process defined by

$$X_t = \rho X_{t+1} + \varepsilon_t, \quad 0 < |\rho| < 1, \tag{4.4}$$

where $(\varepsilon_t)_{t\in\mathbb{Z}} \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta, 1, 0)$. The moving average coefficient is of the form $(\rho^k \mathbb{1}_{\{k\geq 0\}})_k$, and thus, $m_0 = 0$ as stated in (3.9). By Corollary (3.2), we know that for any $m \geq 0$, $h \geq 1$, (X_t) is (m, h)-past-representable. The spectral measures of paths $\boldsymbol{X}_t$ simplify and charge finitely many points. Their forms are given in the following lemma.

Lemma 4.1 *Let (X_t) be an α-stable anticipative AR(1) processes as in (4.4). Letting $\boldsymbol{X_t}$ as in (3.5) for $m \geq 0$, $h \geq 1$, its spectral measure on $C^{\|\cdot\|}_{m+h+1}$ for a seminorm satisfying (3.1) is given by*

$$\Gamma^{\|\cdot\|} = \sum_{\vartheta\in S_1} \left[w_\vartheta \delta_{\{(\vartheta,0,\dots,0)\}} + w_\vartheta \sum_{k=-m+1}^{h-1} \|\boldsymbol{d}_k\|^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}} + \frac{\bar{w}_\vartheta}{1-|\rho|^\alpha} \|\boldsymbol{d}_h\|^\alpha \delta_{\left\{\frac{\vartheta\, \boldsymbol{d}_h}{\|\boldsymbol{d}_h\|}\right\}} \right], \tag{4.5}$$

where for all $\vartheta \in S_1$ and $-m+1 \leq k \leq h$,

$$\begin{aligned}
\boldsymbol{d}_k &= (\rho^{k+m}\mathbb{1}_{\{k\geq -m\}}, \dots, \rho^k \mathbb{1}_{\{k\geq 0\}}, \rho^{k-1}\mathbb{1}_{\{k\geq 1\}}, \dots, \rho^{k-h}\mathbb{1}_{\{k\geq h\}}),\\
w_\vartheta &= (1+\vartheta\beta)/2,\\
\bar{w}_\vartheta &= (1+\vartheta\bar{\beta})/2,\\
\bar{\beta} &= \beta \frac{1-\rho^{<\alpha>}}{1-|\rho|^\alpha},
\end{aligned}$$

and if $h = 1$ and $m = 0$, the sum $\sum_{k=-m+1}^{h-1}$ vanishes by convention.

The next proposition provides the tail conditional distribution of future paths in the case where ρ is positive. Let us first introduce useful neighborhoods of the distinct charged points of $\Gamma^{\|\cdot\|}$. Denote $\boldsymbol{d}_{0,-m} = (\overbrace{1, 0, \dots, 0}^{m+h+1})$ so that the charged points of $\Gamma^{\|\cdot\|}$ are all of the form $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ with indexes (ϑ, k) in the set $\mathcal{I} := S_1 \times \Big(\{-m, h\} \cup \{(0, -m)\}\Big)$. With f as in (4.1), define for any $(\vartheta_0, k_0) \in \mathcal{I}$, the set V_0 as any closed neighborhood of $\vartheta_0 f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\|$ such that

$$\forall (\vartheta', k') \in \mathcal{I}, \qquad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0 \implies \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}, \tag{4.6}$$

In other terms, $V_0 \times \mathbb{R}^d$ is a subset of $C^{\|\cdot\|}_{m+h+1}$ where the only points charged by $\Gamma^{\|\cdot\|}$ have their first $(m+1)^{\text{th}}$ components coinciding with $\vartheta_0 f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\|$. Define $A_{\vartheta,k}$ for any (ϑ, k) as any closed neighborhood of $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ that does not contain any other charged point of $\Gamma^{\|\cdot\|}$, meaning:

$$\forall (\vartheta', k') \in \mathcal{I}, \qquad \frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} \implies (\vartheta', k') = (\vartheta, k). \tag{4.7}$$

Proposition 4.2 *Let (X_t) be an α-stable anticipative AR(1) processes as in (4.4) with $\rho \in (0,1)$. Let $\boldsymbol{X_t}$, the $\boldsymbol{d}_k$'s and the spectral measure of $\boldsymbol{X_t}$ be as given in Lemma 4.1, for any $m \geq 0$, $h \geq 1$. Let V_0 be any small closed neighborhood of $\vartheta_0 f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\|$ in the sense of (4.6) for some $(\vartheta_0, k_0) \in \mathcal{I}$ and let $B(V_0) = V_0 \times \mathbb{R}^h$. Then, with $A_{\vartheta,k}$ an arbitrarily small neighborhood of some $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ as in (4.7), the following hold.*

(ι) *Case $m \geq 1$.*

(*a*) *If $0 \leq k_0 \leq h$:*

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A_{\vartheta,k}\Big|B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \begin{cases} |\rho|^{\alpha k}(1-|\rho|^\alpha)\delta_{\vartheta_0}(\vartheta), & 0 \leq k \leq h-1, \\ |\rho|^{\alpha h}\delta_{\vartheta_0}(\vartheta), & k = h. \end{cases}$$

(*b*) *If $-m \leq k_0 \leq -1$:*

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A_{\vartheta,k}\Big|B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \delta_{\vartheta_0}(\vartheta)\delta_{k_0}(k).$$

($\iota\iota$) *Case $m = 0$.*

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A_{\vartheta,k}\Big|B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \begin{cases} \dfrac{w_{\vartheta_0}}{p_{\vartheta_0}}\delta_{\{\vartheta_0\}}(\vartheta), & k = 0 \\ |\rho|^{\alpha k}(1-|\rho|^\alpha)\delta_{\{\vartheta_0\}}(\vartheta), & 1 \leq k \leq h-1, \\ |\rho|^{\alpha h}\delta_{\{\vartheta_0\}}(\vartheta), & k = h, \end{cases}$$

with $p_{\vartheta_0} = w_{\vartheta_0}/(1-|\rho|^\alpha)$.

For $m \geq 1$, meaning the observed path is assumed to be of length at least 2, there is a significant distinction depending on whether $k_0 \in \{0, \ldots, h\}$ or $k_0 \in \{-m, \ldots, -1\}$. In the latter case, this means that the shock is already observed. This implies that, given the observed path, the shape of the future trajectory is completely determined, as the asymptotic probability of the entire path $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ being in an arbitrarily small neighborhood of $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ is 1 when $\vartheta = \vartheta_0$, $k = k_0$. For the former case, this probability is strictly positive if and only if $\vartheta = \vartheta_0$, but the observed pattern is compatible with several distinct future trajectories. This can be understood by examining the form of the sequences $\boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ and their restrictions to the first

$m+1$ components $f(\boldsymbol{d}_k)/\|\boldsymbol{d}_k\|$. On the one hand (omitting ϑ),

$$\frac{\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} = \begin{cases} \dfrac{(\overbrace{\rho^{k+m},\dots,\rho^k}^{m+1},\overbrace{\rho^{k-1},\dots,\rho,1,0,\dots,0}^{h})}{\|(\rho^{k+m},\dots,\rho^k,\rho^{k-1},\dots,\rho,1,0,\dots,0)\|}, & \text{for} \quad k\in\{0,\dots,h\}, \\ \dfrac{(\rho^{k+m},\dots,\rho,1,0,\dots,0,0,\dots,0)}{\|(\underbrace{\rho^{k+m},\dots,\rho,1,0,\dots,0}_{m+1},\underbrace{0,\dots,0}_{h})\|}, & \text{for} \quad k\in\{-m,\dots,-1\}. \end{cases}$$

We can observe that all the above sequences are segments of explosive exponential functions, each truncated at a certain coordinate. For $k \in \{0,\dots,h\}$, the first zero component, representing the "crash of the bubble", is located at or after the $(m+2)^{\text{th}}$ component, whereas for $k \in \{-m,\dots,-1\}$, it is located at or before the $(m+1)^{\text{th}}$ component. Using the homogeneity of the seminorm and (1.3), we have on the other hand that

$$\frac{f(\boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} = \begin{cases} \dfrac{(\overbrace{\rho^m,\dots,\rho,1}^{m+1})}{\|(\underbrace{\rho^m,\dots,\rho,1}_{m+1},\underbrace{0,\dots,0,0,\dots,0}_{h})\|}, & \text{for} \quad k\in\{0,\dots,h\}, \\ \dfrac{(\overbrace{\rho^{k+m},\dots,\rho,1,0,\dots,0}^{m+1})}{\|(\underbrace{\rho^{k+m},\dots,\rho,1,0,\dots,0}_{m+1},\underbrace{0,\dots,0}_{h})\|}, & \text{for} \quad k\in\{-m,\dots,-1\}. \end{cases}$$

Thus, conditioning the trajectory on the event $\{f(\boldsymbol{X_t})/\|\boldsymbol{X_t}\| \approx f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\|\}$ for some $k_0 \in \{-m,\dots,-1\}$ amounts to conditioning on the burst of a bubble being observed in the past trajectory, with no new bubble forming yet. This allows for the exact identification of the position of the pattern within the moving average's coefficient sequence.

When conditioning with $k_0 \in \{0,\dots,h\}$ however, the crash date is not observed and can occur either within the next $h-1$ periods or after h. The shape of the observed path, though, corresponds to a segment of exponential growth with a rate of ρ^{-1} regardless of how much time remains before the burst. This leaves several potential future paths. The likelihood of each possible scenario can be quantified as follows: the quantity $|\rho|^{\alpha k}(1-|\rho|^{\alpha})$ represents the probability that the bubble will peak in exactly k periods ($0 \le k < h$), while $|\rho|^{\alpha h}$ represents the probability that the bubble will last at least h more periods. This confirms the interpretation of the conditional moments proposed in Fries (2021). Additionally, it extends this interpretation by considering entire paths rather than just point predictions. For $m=0$, meaning only the present value is observed, no pattern can be identified, only the sign of the shock. Hence, the growth rate ρ^{-1} of the ongoing event becomes unidentifiable. This is reflected in the fact that

the asymptotic probabilities of paths with growth rates ρ^{-1}, are positive (as shown in case $(\iota\iota)$ of Proposition 4.2).

4.3 The anticipative AR(2) and fractionally integrated white noise

We focus here on two processes that both share the interesting property of having a 0-1 tail conditional distribution whenever the observed path has length at least 2 (i.e., $m \geq 1$): the anticipative AR(2) and the anticipative fractionally integrated white noise (FWN). With an appropriate choice of parameters, the former can generate bubble-like trajectories with accelerating or decelerating growth rates, while the latter can accommodate hyperbolic bubbles. In contrast with the anticipative AR(1), these bubbles do not display an exponential profile, but they still exhibit an inflation-peak-collapse behavior. Any extension of these two minimal specifications should preserve the following properties.

Anticipative AR(2)

The anticipative AR(2) is the strictly stationary solution of

$$(1-\lambda_1 F)(1-\lambda_2 F)X_t = \varepsilon_t, \qquad \varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha,\beta,\sigma,0), \tag{4.8}$$

where $\lambda_i \in \mathbb{C}$ and $0 < |\lambda_i| < 1$ for i=1,2. In case $\lambda_i \in \mathbb{C}\setminus\mathbb{R}$, $i = 1,2$, we impose that $\lambda_1 = \bar{\lambda}_2$ to ensure (X_t) is real-valued. We further assume that $\lambda_1 + \lambda_2 \neq 0$, to exclude the cases where (X_{2t}) and (X_{2t+1}) are independent anticipative AR(1) processes. The solution of (4.8) admits the moving average representation $X_t = \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k}$ with

$$d_k = \begin{cases} \dfrac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2}\, \mathbb{1}_{\{k\geq 0\}}, & \text{if} \quad \lambda_1 \neq \lambda_2, \\ (k+1)\lambda^k\, \mathbb{1}_{\{k\geq 0\}}, & \text{if} \quad \lambda_1 = \lambda_2 = \lambda. \end{cases} \tag{4.9}$$

Anticipative fractionally integrated white noise

The anticipative FWN process can be defined as the stationary solution of

$$(1-F)^d X_t = \varepsilon_t, \qquad \varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha,\beta,\sigma,0), \tag{4.10}$$

with $\alpha(d-1) < -1$. The solution of (4.10) admits the moving average representation $X_t = \sum_{k=0}^{+\infty} d_k \varepsilon_{t+k}$ with $d_0 = 1$ and

$$d_k = \frac{\Gamma_f(k+d)}{\Gamma_f(d)\Gamma_f(k+1)}\, \mathbb{1}_{\{k\geq 0\}}, \quad \text{for} \quad k \neq 0, \tag{4.11}$$

where $\Gamma_f(\,\cdot\,)$ denotes the Gamma function.

It can be shown that both process are necessarily (m,h)-past-representable for $m \geq 1$ and $h \geq 1$. The 0-1 tail conditional distribution property, when the observed path has length at least 2, is demonstrated in the next proposition:

Proposition 4.3 *Let (X_t) be the α-stable anticipative AR(2) or anticipative fractionally integrated white noise. For any $m \geq 1$ and $h \geq 1$, let $\boldsymbol{X_t}$ as in (3.5) and $\boldsymbol{d}_k = (d_{k+m}, \ldots, d_k, d_{k-1}, \ldots, d_{k-h})$ where (d_k) is as in (4.9) or (4.11). Also V_0 is a small neighborhood of $\vartheta_0 \boldsymbol{d}_{k_0}/\|\boldsymbol{d}_{k_0}\|$ as in (4.6) for some $\vartheta_0 \in S_1$, $k_0 \in \{-m, \ldots, h\}$, and $B(V_0) = V_0 \times \mathbb{R}^h$. Then,*

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A \Big| B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \begin{cases} 1, & \text{if } \dfrac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|} \in A, \\ 0, & \text{otherwise,} \end{cases}$$

for any closed neighborhood $A \subset C_{m+h+1}^{\|\cdot\|}$ such that

$$\partial A \cap \{\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\| : \vartheta \in S_1,\ k \in \{-m, \ldots, h\}\} = \emptyset.$$

Remark 4.2 Contrary to the anticipative AR(1), the trajectories of the anticipative AR(2) and fractionally integrated processes do not leave room for indeterminacy of the future path. Asymptotically, given any observed path of length at least 2, the shape of the future trajectory can be deduced deterministically. This holds even if the peak/collapse of a bubble is not yet present in the observed piece of trajectory. Therefore, provided the current pattern is properly identified,[9] it appears possible in the framework of these models to infer in advance the peak and crash dates of bubbles with very high confidence, in principle, with certainty.

5 Monte Carlo study and numerical analysis

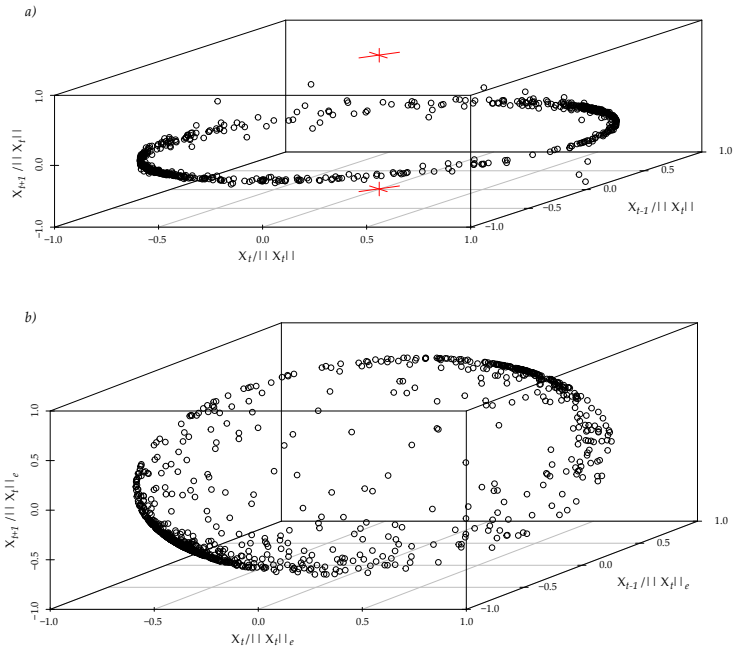
In this section, we take advantage of our theoretical results in different ways. In particular, we suggest two forecasting procedures and demonstrate their performances in finite samples. We also use numerical simulations to provide a visual illustration of the unit cylinder in the particular case of a $\{0,1\}$ tail conditional distribution.

5.1 Visualisation of the unit cylinder

In the spirit of the Remark 2.1, we consider an α-stable vector $\boldsymbol{X_t} = (X_{t-1}, X_t, X_{t+1})$ where X_t is an anticipative AR(2) specified as in 5.2. X_t being past-representable, it admits a representation

[9] See point (ι) of Remark 4.1.

Figure 1: Unit cylinder and unit sphere representations of $X_t = 0.7X_{t+1} + 0.1X_{t+2} + \varepsilon_t$



on the unit-cylinder, as the Theorem 3.1 applies. Furthermore, as discussed in 4.3, its spectral measure exhibits the following asymptotic behavior

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A\Big|B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}\left(A \cap \left\{\dfrac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right)}{\Gamma^{\|\cdot\|}\left(\left\{\dfrac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right)}.$$

and hence $\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A\Big|B(V_0)\Big)$ is either 1 or 0. This peculiar $\{0,1\}$ tail conditional distribution leads to the following graphical representation on the unit-cylinder (see Figure 1.a). The simulation of X_t is performed for a sample size $n = 1000$.

We clearly see that $C_3^{\|\cdot\|}$ spans all directions of $\mathbb{R}^3$ but the ones of $(0,0,-1)$ and $(0,0,+1)$. This is of no consequence as the representability property holds and implies that $\Gamma\Big(\{(0,0,-1),(0,0,+1)\}\Big) = 0$ as $x \to \infty$. In other words, the seminorm representability reflect the fact that extreme realizations of X_{t+1} never occur conditionally to small realizations of X_{t-1} and X_t. Those inaccessible coordinates are indicated by the two red cross. In the opposite case where we represent $\boldsymbol{X_t}$ on the unit sphere, S_3 spans all directions of $\mathbb{R}^3$ and describes any potential tail dependence of (X_{t-1}, X_t, X_{t+1}). This includes the tail dependence between X_{t+1} and the past, which reflects the odd (and rare, as depicted in Figure 1.b) situation where the realisation of X_{t+1} is extreme whereas immediate past realizations are not.

5.2 Forecasting crash probabilities

In this section, we illustrate through simulations that the probability on the left-hand side of Proposition 4.1, converges to the right-hand side when the conditioning value $\|\boldsymbol{X_t}\|$ is large. For the anticipative AR(2), this proposition boils down to Proposition 4.3, with the intriguing result that the future path is purely deterministic. To illustrate these two propositions, we simulate 100 trajectories of $N = 10^6$ observations for four different specifications of MAR processes, as defined in Corollary 3.2. The results of each simulation are reported in the associated tables: the purely anticipative AR(2) (or MAR(0,2)) process in Table 1, the MAR(1,1) process in Table 2, the MAR(1,2) process in Table 3, and the purely anticipative AR(3) (or MAR(0,3)) process in Table 4.[10] As in the empirical section, where the best retained specification is an anticipative AR(2), we begin by presenting the MC simulation procedure for this process. The procedure for the other three specifications is similar. The simulated process, is the following: $X_t = 0.7X_{t+1} + 0.1X_{t+2} + \varepsilon_t$ where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$. We focus here on the case $m = 2$ and evaluate the crash probabilities at different forecasting horizons $h = (1, 5, 10)$. The choice of m is set a priori. According to Proposition 4.3, the only theoretical condition is that $m \geq 1$. To assess whether the choice of m has a practical impact on the predictability of crash probability, we perform another simulation with $m = 7$ for the AR(2) process. The results of this simulation is provided in Table 12 in Appendix A. This table shows results similar to those in Table 1, suggesting that the choice of m has no significant effect on the predictability of crash probabilities and that the theoretical condition appears to hold. However, as discussed in the next section, when predicting the crash date, the choice of m does matter.

The left-hand side of Proposition 4.3 needs two types of conditioning. First we condition on $\|\boldsymbol{X_t}\|$ to be large, and we choose $X_t \geq q$ with q is a theoretical quantile of the marginal distribution of X_t, ranging from 0.90 to 0.9999 quantiles. Second we define the conditioning of the small neighborhood $B(V_0)$:

$$B(V_0) = \left[\frac{\vartheta_0 \boldsymbol{d}_{k_0-m}}{\|\boldsymbol{d}_{k_0}\|} - e, \frac{\vartheta_0 \boldsymbol{d}_{k_0-m}}{\|\boldsymbol{d}_{k_0}\|} + e\right] \times \cdots \times \left[\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|} - e, \frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|} + e\right].$$

Where $\vartheta_0 = 1$, as we investigate the crash probability of a 'positive' bubble. Specifically, we condition $|\boldsymbol{X_t}|$ on $X_t \geq q$, where q is the upper quantile, and e is a very small neighborhood. In practice, e is determined by the minimum Euclidean distance between $\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}$ and $\frac{\boldsymbol{X_t}}{\|\boldsymbol{X_t}\|}$. We also set $A = B(V_0) \times [-\delta, \delta]$, with $\delta = 0.2$. This is equivalent to estimating the probability of a

[10] The length N of one simulated trajectory seems large because we are investigating the probability when $|\boldsymbol{X_t}|$ is very large, greater than the 99.99% quantile. To do this, we need enough observations to ensure we have sufficient data above this quantile.

crash (a decrease of more than 80%) at horizon h.[11] For each simulated trajectory, we compute the two following estimators, one for the probability on the left-hand side of Proposition 4.3 defined as

$$\hat{p}_q = \frac{\sum_{t=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-m},\dots,X_t)}{\|\boldsymbol{X_t}\|}\in B(V_0)\right\}\cap\{\frac{X_{t+h}}{\|\boldsymbol{X_t}\|}\leq\delta\}\cap\{X_t>q\}\right)}}{\sum_{t=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-m},\cdots,X_t)}{\|\boldsymbol{X_t}\|}\in B(V_0)\right\}\cap\{X_t>q\}\right)}} \tag{5.1}$$

and the other one for the probability on the right-hand side of Proposition 4.3 p_q. The latter is computed as follows:

$$p_q = \frac{\sum_{t=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-m},\cdots,X_t)}{\|\boldsymbol{X_t}\|}\in B(V_0)\right\}\cap\{\frac{d_{k_0+h}}{\|\boldsymbol{d}_{k_0}\|}\leq\delta\}\cap\{X_t>q\}\right)}}{\sum_{i=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-m},\cdots,X_t)}{\|\boldsymbol{X_t}\|}\in B(V_0)\right\}\cap\{X_t>q\}\right)}} \tag{5.2}$$

According to Proposition 4.3, these two probabilities have to converge to the same value, because $\boldsymbol{X_t}/\|\boldsymbol{X_t}\| \in A$, is equivalent to $\vartheta_0\boldsymbol{d}_{k_0}/\|\boldsymbol{d}_{k_0}\| \in A$. To estimate the $\boldsymbol{d}_{k_0}$, and check whether $\boldsymbol{X_t}/\|\boldsymbol{X_t}\| \in B(V_0)$, we determine the sample of size m of the d_k deterministic path which $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ is in $B(V_0)$. To do so, first, we compute $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ for $m = 2$. Second, we evaluate the Euclidian disctance between $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ and $\vartheta_0\boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ for $k \in (\underline{k}, \bar{k})$ where $\underline{k} = 30$ and $\bar{k} = 0$ and we check whether some $\vartheta_0\boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ belongs to a small neighborhood of $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$.Once we have determined k_0, we simply extend the deterministic path to k_0+h and calculate $d_{k_0+h}/|\boldsymbol{d}_{k_0}|$.

Table 1 gathers the average of p_q and $\hat{p}_q$ empirical probabilities across the M simulations along empirical 95% confidence. One notices that the empirical probabilities indeed come very close to the theoretical ones as q increases. However, as h increases, there is a loss in accuracy, and it tends to overestimate the crash probability.

Table 1: Comparison of theoretical and empirical crash probabilities at horizons $h = 1, 5, 10$ of bubbles generated by the anticipative AR(2)

	$h = 1$	$h = 5$	$h = 10$
$p_{0.9}\backslash\hat{p}_{0.9}$	46.362\17.247 (17.061-17.414)	99.946\52.369 (51.946-52.794)	100.000\69.239 (68.742-69.677)
$p_{0.99}\backslash\hat{p}_{0.99}$	35.476\27.413 (26.490-28.191)	99.825\73.396 (71.932-74.640)	100.000\90.438 (89.254-91.511)
$p_{0.999}\backslash\hat{p}_{0.999}$	30.093\29.490 (26.763-33.122)	99.415\77.712 (72.964-82.796)	100.000\94.708 (92.124-97.136)
$p_{0.9999}\backslash\hat{p}_{0.9999}$	30.155\30.141 (24.000-38.017)	98.311\79.247 (69.003-89.492)	100.000\95.782 (87.993-100.000)

Notes: The theoretical crash probabilities p_q are computed using (5.2). Empirical average (Mean) and 95% confidence intervals (95%-CI) of the estimated probabilities are computed using (5.1) on M = 100 simulated trajectories of $N = 10^6$ observations, for $q = q_a$ several a-quantiles of the marginal distribution of X_t. These probabilities are reported in percent. We set $m = 2$. The simulated process is the following: $X_t = 0.7X_{t+1} + 0.1X_{t+2} + \varepsilon_t$ where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$

[11] Determining this parameter can be challenging, as it raises the question of what drop size constitutes a bubble. For this simulation exercise, we have selected a significant drop size. Further numerical results could help in defining this parameter more accurately.

Proposition 4.1 is more general than Proposition 4.3; the former includes any process, whereas the latter only applies to processes with at least two leads ($q > 2$). The MAR(1,1) process is not included in the latter proposition, as it does not have the 0-1 tail deterministic conditional distribution, and it is subject to different types of uncertainty (see Remark 4.1). It is important to highlight whether this distinction matters in practice for predicting crash probabilities. To investigate this, we apply the same procedure as for the AR(2) process to the following simulated process: $X_t = 0.9X_{t+1} + 0.1X_{t-1} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$. The choice of 0.9 for the anticipative parameter is not arbitrary; we aim for a long and gradual increase in the bubble. However, the results remain similar even when choosing a parameter that models a faster increase in the bubble's growth rate. Table 2 presents the average of p_q and $\hat{p_q}$ empirical probabilities across the M simulations, along with the empirical 95% confidence intervals.

Table 2: Comparison of theoretical and empirical crash probabilities at horizons $h = 1, 5, 10$ of bubbles generated by the MAR(1,1)

	$h = 1$	$h = 5$	$h = 10$
$p_{0.9} \backslash \hat{p}_{0.9}$	0.000\0.478 (0.441-0.518)	0.000\19.961 (19.563-20.414)	79.101\53.569 (52.96-54.266)
$p_{0.99} \backslash \hat{p}_{0.99}$	0.000\0.064 (0.017-0.126)	0.000\5.005 (4.181-5.728)	70.413\64.874 (63.214-66.255)
$p_{0.999} \backslash \hat{p}_{0.999}$	0.000\0.007 (0.000-0.109)	0.000\0.521 (0.000-1.559)	69.347\69.284 (64.843-72.694)
$p_{0.9999} \backslash \hat{p}_{0.9999}$	0.000\0.006 (0.000-0.010)	0.000\0.053 (0.000-0.567)	67.580\67.561 (53.115-75.074)

Notes: The theoretical crash probabilities p_q are computed using (5.2). Empirical average (Mean) and 95% confidence intervals (95%-CI) of the estimated probabilities are computed using (5.1) on M = 100 simulated trajectories of $N = 10^6$ observations, for $q = q_a$ several a-quantiles of the marginal distribution of X_t. These probabilities are reported in percent. We set m=2. The simulated process is the following $X_t = 0.9X_{t+1} + 0.1X_{t-1} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$

Table 2 shows that we can accurately capture the probability of a crash for a short-term forecast horizon ($h = 1$) compared to longer horizons ($h > 1$), except when considering the asymptotic distribution of $\boldsymbol{X_t}$, $q > q_{0.999}$. This suggests that Proposition 4.1 is verified. It also appears that, in practice, the uncertainty discussed in Remark 4.1 is manageable. One could hypothesize that increasing the persistence in the lead part of the processes should improve the accuracy of predicting crash probability, and/or introducing a persistence decrease in the non-anticipative part of the process could make it more challenging to assess crash probability. To explore this, we simulate and apply our procedure to two processes: the purely anticipative AR(3), following the equation: $X_t = 0.8X_{t+1} + 0.2X_{t+2} - 0.1X_{t+3} + \varepsilon_t$ and the MAR(1,2), following the equation: $X_t = 0.7X_{t+1} + 0.1X_{t+2} + 0.4X_{t-1} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$. Table 3 presents the results for the MAR(1,2), which show similar, and even better, outcomes compared to those in Table 1.

Table 3: Comparison of theoretical and empirical crash probabilities at horizons $h = 1, 5, 10$ of bubbles generated by the MAR(1,2)

	$h = 1$	$h = 5$	$h = 10$
$p_{0.9}\backslash\hat{p}_{0.9}$	9.997\10.429 (10.133-10.743)	96.264\59.71 (58.816-60.517)	100.000\82.535 (81.343-83.754)
$p_{0.99}\backslash\hat{p}_{0.99}$	8.915\9.066 (8.436-9.815)	89.724\66.83 (63.998-69.511)	99.998\90.299 (87.706-92.689)
$p_{0.999}\backslash\hat{p}_{0.999}$	8.522\8.534 (5.782-10.661)	76.974\67.486 (59.013-78.226)	100.000\91.695 (83.716-99.396)
$p_{0.9999}\backslash\hat{p}_{0.9999}$	7.815\7.815 (0.000-13.437)	71.505\68.23 (45.776-100.000)	100.000\92.863 (70.983-100.000)

Notes: The theoretical crash probabilities p_q are computed using (5.2). Empirical average (Mean) and 95% confidence intervals (95%-CI) of the estimated probabilities are computed using (5.1) on M = 100 simulated trajectories of $N = 10^6$ observations, for $q = q_a$ several a-quantiles of the marginal distribution of X_t. These probabilities are reported in percent. We set m=2. The simulated process is the following $X_t = 0.7X_{t+1} + 0.1X_{t+2} + 0.4X_{t-1} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$

Table 4 primarily confirms the results on the asymptotic convergence of probability. However, surprisingly, increasing persistence in the anticipative part of the process does not improve the accuracy of predicting the crash probability rather, it has a small opposite effect.

Table 4: Comparison of theoretical and empirical crash probabilities at horizons $h = 1, 5, 10$ of bubbles generated by the anticipative AR(3)

	$h = 1$	$h = 5$	$h = 10$
$p_{0.9}\backslash\hat{p}_{0.9}$	24.300\8.599 (8.433-8.743)	91.947\34.573 (34.128-35.056)	98.791\54.792 (54.087-55.440)
$p_{0.99}\backslash\hat{p}_{0.99}$	16.321\13.740 (13.093-14.345)	85.883\50.585 (48.61-52.659)	98.962\75.168 (73.035-77.219)
$p_{0.999}\backslash\hat{p}_{0.999}$	14.821\14.612 (12.53-16.693)	68.756\53.400 (47.320-58.863)	98.823\78.857 (72.304-85.104)
$p_{0.9999}\backslash\hat{p}_{0.9999}$	15.323\15.287 (10.716-22.733)	57.840\54.813 (41.167-72.491)	97.982\79.757 (63.015-99.580)

Notes: The theoretical crash probabilities p_q are computed using (5.2). Empirical average (Mean) and 95% confidence intervals (95%-CI) of the estimated probabilities are computed using (5.1) on M = 100 simulated trajectories of $N = 10^6$ observations, for $q = q_a$ several a-quantiles of the marginal distribution of X_t. These probabilities are reported in percent. We set m=3. The simulated process is the following $X_t = 0.8X_{t+1} + 0.2X_{t+2} - 0.1X_{t+3} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$

5.3 Forecasting crash dates

One can also apply Proposition 4.1 to infer information on future paths from the observed trajectory, as long as it deviates far enough from central values. We focus on the case where k belongs to $\{-m, \ldots, -1\}$ and document that in practice, for large values of x and sufficiently persistence anticipative process, the approximation

$$\boldsymbol{X_t}/\|\boldsymbol{X_t}\| \approx \vartheta(d_{k+m}, \ldots, d_{k+1}, d_k)/\|\boldsymbol{d}_k\|, \quad \boldsymbol{X_t} = (X_{t-m}, \ldots, X_{t-1}, X_t),$$

can be used to derive the next crash date and then estimate the future path up to $t+h$. We also discuss to what extent the sources of uncertainty listed in Remark 4.1 affect the performance

of our procedure in the presence of finitely large realizations. As for a range of realizations, we ignore to which piece of the moving average trajectory it corresponds, we pay particular attention to the selection of k_0 and the impact of m. ϑ_0 is assumed to be known here as in general it can be deduced from the data.

Our forecasting procedure proceeds in 4 steps. First, we compute $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ for a given $m \geq 1$. Second, we evaluate $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ for $k \in (\underline{k}, \bar{k})$ where $\underline{k} = 30$ and $\bar{k} = 0$. Third, we check whether some $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ belong to a small neighborhood A of $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$. If k_0 cannot be identified because several values k satisfy this condition, we reduce the neighborhood until a unique $k = k_0$ remains. At this stage, we structure $\boldsymbol{X_t}$ as in (3.5) and $\boldsymbol{d}_{k_0}$ as $(d_{k+m}, \ldots, d_k, d_{k-1}, \ldots, d_{k-h})$. The last step simply consists of using the deterministic trajectory of $\boldsymbol{d}_{k_0}$ to iterate up to $d_{k-h} = 0$ and hence obtain the bubble burst date. From Proposition 4.3 we know that if X_t is anticipative enough, its future path will follow the one of $\boldsymbol{d}_{k_0}$ with a very high level of confidence such that

$$(X_{t-m}, \ldots, X_{t-1}, X_t, X_{t+1}, \ldots, X_{t+h})/\|\boldsymbol{X_t}\| \approx \vartheta_0 (d_{k+m}, \ldots, d_{k+1}, d_k, d_{k-1}, \ldots, d_{k-h})/\|\boldsymbol{d}_k\|,$$

hence offering the possibility to predict $X_{t+1}, \ldots, X_{t+h}$.

This procedure is likely to be sensible to the selection of m. We investigate this issue by considering $m = \{1, 3, 5, 7, 9, 11\}$. We also anticipate that, how far we deviate from the Gaussian distribution, in terms of tail index, is likely to affect the results, and hence we consider $\alpha = \{0.9, 1.2, 1.5, 1.8\}$ (see Remark 4.1 (*iii*)). As in the previous section, we simulated three out of the four different processes that we considered, and the associated results of each simulation are reported in the corresponding tables: the purely anticipative AR(2) (or MAR(0,2)) process in Table 6, the MAR(1,2) process in Table 7, and the purely anticipative AR(3) (or MAR(0,3)) process in Table 8. The MAR(1,1) process was ruled out because Proposition 4.3 does not apply directly, as we do not obtain certainty in the path but rather some probabilities. Let start by presenting the simulation procedure for the anticipative AR(2). Each simulated path is governed by a SαS anticipative AR(2) of the following form: $X_t = 0.7X_{t+1} + 0.1X_{t+2} + \varepsilon_t$ where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, 0, 0.1, 0)$. For a given artificial time series x_t, we identify a positive bubble peak as $max(x_t)$ and treat as unobserved the remaining values of the series and the $\lceil N \times 0.01 \rceil$ periods preceding the bubble burst. We then explore all these scenarios for $N = \{250, 500, 1000\}$ (i.e. $k_0 = \{3, 5, 10\}$) and 1000 trajectories. In theory, N should not impact the prediction performance but we use it here to control the quantile of the last in-sample observation. More precisely, our simulation framework results in the quantiles reported in Table 5 and allows us to investigate the impact of departing from the asymptotic theory ($x \to \infty$). For instance, we can see that for $N = 1000$, the last in-sample observation used to predict an extreme event

that surge 10 periods ahead, actually corresponds to the quantile 0.91 when $\alpha = 1.5$. In such a configuration, the realizations of X_t are likely to be only moderately large compared to the asymptotic requirements ($x \to \infty$).

Table 5: Quantile of the last in-sample observation

N/α	0.9	1.2	1.5	1.8
250	0.99	0.99	0.99	0.94
500	0.98	0.98	0.94	0.89
1000	0.97	0.96	0.91	0.78

Accordingly, in the simulation results, we report the labels "High", "Quite High", "Moderately High", rather than the sample sizes. For each simulation, we compute the bias as the difference between the predicted crash date and the true simulated date.

The results are reported in Table 6. First, the results shed light on the crucial role of limit theory, as the predicted crash date is considerably more biased when the shape of the trajectory is inferred from an observation that corresponds to a moderately high quantile. Second, for a given m, the performance deteriorates as α increases, thereby involving quantiles far from the asymptotic theory (e.g., $q_{X_t} \approx 0.78$ when $N = 1000$ and $\alpha = 1.8$) and introducing more noise. Our theory states that when $x \to \infty$, $m = 1$ can be sufficient. However, in practice, the simulation study reveals that the optimal selection of m is not obvious, as it interacts in a complex manner with the tail index α. For instance, when X_t is very high and the tail index is close to 1, a shorter m improves the performance of the forecasting procedure. Conversely, if the tail index is close to 2, it is slightly better to select a medium-range m. The same analysis holds for X_t that is large or moderately large.

Table 7 presents the results for the MAR(1,2). We want to check if adding a causal part impacts the estimation of the crash date. Theoretically, it should lead to no difference, as only the persistence in the anticipative part matters in estimating the crash date. This is verified, as Table 6 shows similar results except for the choice of m. For the MAR(1,2), choosing a lower m is always better. We also apply the same procedure to an anticipative AR(3) process to check if adding persistence in the anticipative part improves the estimation of the crash date. Table 8 showcases the results for this process. Overall, the bias is smaller in any case compared to the anticipative AR(2). However, in comparison to AR(2) and even for the MAR(1,2), the choice of m is not clear-cut. It seems better to choose a large or medium-range m when α is close to 2, and for small α, it is better to choose a small m.

Table 6: Bias for the crash date predictor for the purely anticipative AR(2) process

High						
α/m	1	3	5	7	9	11
0.9	0.1540	0.3590	0.5310	0.5570	0.6730	0.7280
1.2	0.5850	0.7950	0.8500	0.9880	0.9860	1.0040
1.5	1.0170	1.0510	1.1280	1.1500	1.2330	1.2470
1.8	1.2660	1.2480	1.3130	1.3040	1.3200	1.3570
Quite High						
α/m	1	3	5	7	9	11
0.9	1.7670	2.0210	2.2130	2.3010	2.3880	2.4680
1.2	2.5120	2.6530	2.7960	2.8500	2.9580	2.9460
1.5	2.9550	3.0560	3.1160	3.1160	3.1470	3.1750
1.8	3.2970	3.2250	3.2760	3.3070	3.3310	3.2860
Moderately High						
α/m	1	3	5	7	9	11
0.9	6.8500	7.1450	7.2460	7.3340	7.4360	7.5420
1.2	7.6960	7.8270	7.9320	7.9790	8.0530	8.0810
1.5	8.1690	8.1730	8.2160	8.2390	8.2910	8.2860
1.8	8.4230	8.3550	8.3640	8.3790	8.4210	8.4700

Notes: The simulated process is the following:$X_t = 0.7X_{t+1} + 0.1X_{t+2} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, 0, 0.5, 0)$. High, Quite High, and Moderately High correspond to a number of simulated observations of 250, 500, and 1000, respectively.

6 Forecasting climate anomalies

A growing body of literature highlights the impact of climate variables on economic performance (Dell et al., 2014). A key variable for identifying this impact is the occurrence of El Niño (and, correspondingly, La Niña) weather shocks. These shocks are known to affect various economic indicators, including growth, inflation, energy markets, and agricultural commodity returns (Brenner, 2002; Cashin et al., 2017; Makkonen et al., 2021). Providing a forecast of El Niño (and La Niña) weather shocks is of primary importance, as it offers numerous societal benefits, ranging from extreme weather warnings to agricultural planning (Alley et al., 2019). El Niño (or La Niña) intensity is defined as a value constructed from the Southern Oscillation Index

Table 7: Bias for the crash date predictor for the MAR(1,2) process

High						
α/m	1	3	5	7	9	11
0.9	0.3170	0.5870	0.6650	0.7870	0.8420	0.8880
1.2	0.7240	0.9630	1.0490	1.1050	1.1620	1.1180
1.5	1.0880	1.2100	1.2760	1.2320	1.2530	1.2990
1.8	1.2670	1.3400	1.3860	1.3630	1.3720	1.3740
Quite High						
α/m	1	3	5	7	9	11
0.9	1.9840	2.3310	2.5250	2.5570	2.5540	2.6780
1.2	2.6130	2.8560	2.9550	2.9940	3.0280	3.0480
1.5	3.0080	3.2010	3.2180	3.2230	3.3120	3.2940
1.8	3.3040	3.3480	3.3440	3.3550	3.3930	3.4100
Moderately High						
α/m	1	3	5	7	9	11
0.9	7.1650	7.2600	7.5560	7.5900	7.7090	7.6910
1.2	7.8420	7.9810	8.1190	8.1320	8.1760	8.1810
1.5	8.1990	8.2750	8.2870	8.3690	8.3500	8.3910
1.8	8.4080	8.4180	8.4180	8.4300	8.4470	8.4670

Notes: The simulated process is the following:$X_t = 0.7X_{t+1} + 0.1X_{t+2} + 0.4X_{t-1} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, 0, 0.5, 0)$. High, Quite High, and Moderately High correspond to a number of simulated observations of 250, 500, and 1000, respectively.

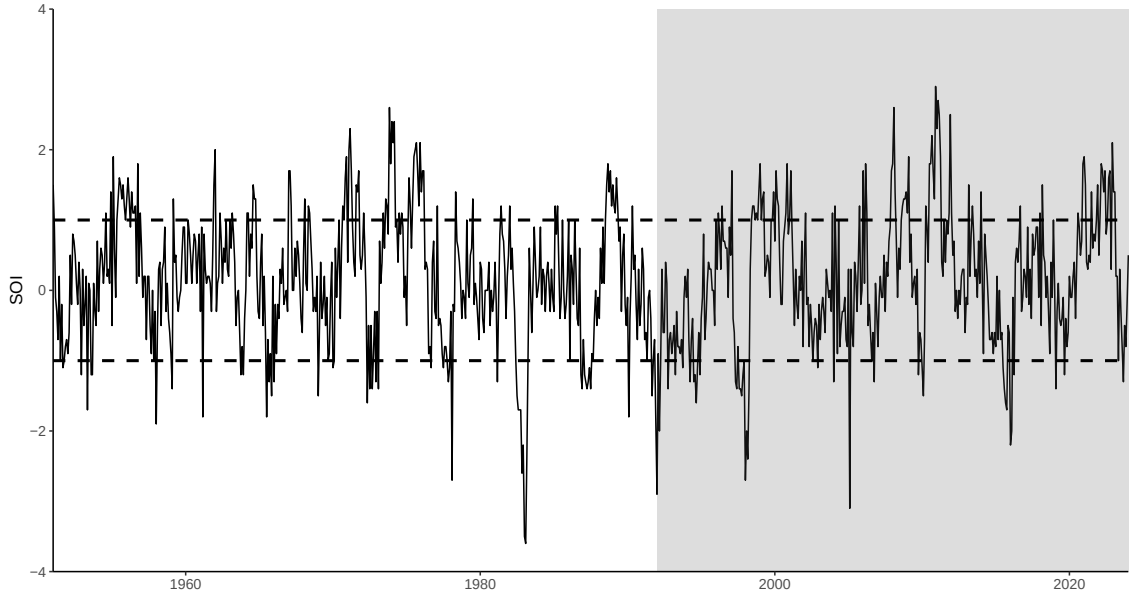
(SOI).[12]

This section discusses the performance of our forecasting procedures in detecting the peak and the end of an El Niño (and La Niña) shock, and assesses the probability of staying in these episodes h periods ahead. We split the data into an in-sample period (from 01/1951 to 12/1991) and an out-of-sample period (from 01/1992 to 01/2024) to test the robustness of our forecasting procedure. Figure 6 displays the data sample, where the shaded area corresponds to the out-of-sample data. The alternation of boom and bust, which appears to be an identifiable deterministic pattern, is clearly distinguishable.

Figure 6 shows that SOI values, range approximately from -4 to 2. Most observations

[12] Data and methodology for constructing the SOI are available at https://www.ncei.noaa.gov/access/monitoring/enso/soi. The SOI is a monthly variable derived from air-pressure differentials in the South Pacific, measured between Tahiti and Darwin.

Figure 2: Southern Oscillation Index (SOI)



Notes: The shaded area represents the out-of-sample data. The horizontal black dashed lines indicate values of 1 and -1. An El Niño event is defined as an SOI below -1, while a La Niña event corresponds to an SOI above 1, each persisting for at least three periods.

Figure 3: Distribution of the Southern Oscillation Index (SOI)

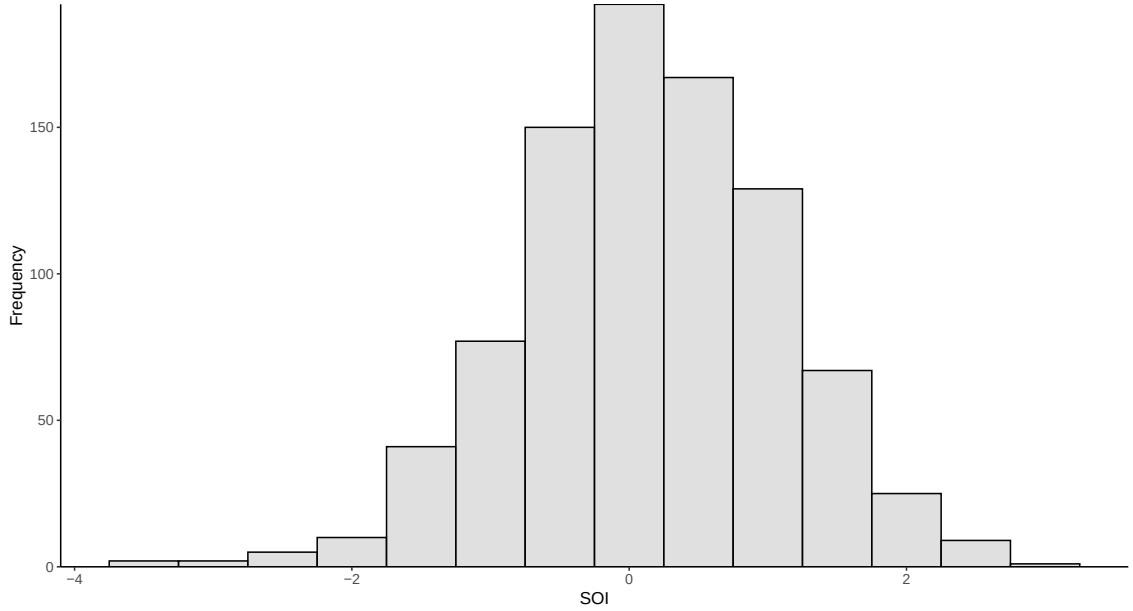


Table 8: Bias for the crash date predictor for the purely anticipative AR(3) process

High						
α/m	1	3	5	7	9	11
0.9	-0.3180	-0.3170	-0.4670	-0.6250	-0.6740	-0.5570
1.2	-0.1440	-0.3650	-0.5310	-0.4950	-0.4850	-0.3820
1.5	0.3630	-0.1000	-0.2420	-0.2150	-0.2960	-0.1490
1.8	0.8220	0.2710	-0.0890	-0.1710	0.0480	0.0160
Quite High						
α/m	1	3	5	7	9	11
0.9	0.2200	0.2790	0.4240	0.3490	0.4730	0.5830
1.2	0.9180	0.8680	0.9250	1.0030	1.0500	1.1030
1.5	1.8640	1.6100	1.4960	1.4720	1.5320	1.5520
1.8	2.6780	2.2010	1.8550	1.8870	1.9600	2.0040
Moderately High						
α/m	1	3	5	7	9	11
0.9	3.7370	4.2240	4.4710	4.6080	4.7000	5.0470
1.2	5.5510	5.7180	5.7680	5.8700	5.9220	5.9730
1.5	6.7790	6.5850	6.5550	6.5830	6.6000	6.6310
1.8	7.7710	7.2760	7.0830	7.1390	7.1680	7.1380

Notes: The simulated process is the following:$X_t = 0.8X_{t+1} + 0.2X_{t+2} - 0.1X_{t+3} + \varepsilon_t$, where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, 0, 0.5, 0)$. High, Quite High, and Moderately High correspond to a number of simulated observations of 250, 500, and 1000, respectively.

are clustered around 0, reflecting the index's central tendency, while fewer are found at the extremes, underscoring the rarity of extreme El Niño (SOI below -1) or La Niña (SOI above 1) events. Notably, an SOI of -1 corresponds to the 10th quantile, while 1 is equal to the 80th quantile. To assess the non-normality of the SOI, we compute the skewness (-0.2024) and the kurtosis (3.4823), and conduct a Jarque-Bera test. The results, summarized in Table 13 in the Appendix, yield a p-value below 0.05, strongly rejecting the null hypothesis of normality. This deviation is likely due to the slight skewness and excess kurtosis

Estimating mix-causal processes is challenging, and only a few estimation procedures are available. For the sake of robustness, we rely on two different methods: the semi-parametric Generalized Covariance Estimator (see Gourieroux and Jasiak, 2023), referred to as GCoV, and the approximate maximum likelihood method (AML), introduced by Lanne and Saikkonen (2011). Both methods face limitations due to the assumption that the error term follows a

$\mathcal{S}(\alpha, \beta, \sigma, 0)$ distribution. For AML, Andrews et al. (2009) demonstrates that while asymptotic results exist, the limit distribution is largely intractable because the rate of convergence depends on the tail α parameter.[13] Gourieroux and Jasiak (2023) shows that the consistency and asymptotic properties of the semi-parametric estimator require the existence of the first four moments of the residuals, which is not the case for α-stable laws, as only the first $2\alpha + 1$ unconditional moments exist. However, as pointed out in Remark 2 on p. 1318 of Gourieroux and Jasiak (2023), if the error term has no finite fourth-order moment, the consistency and asymptotic properties of the GCoV estimator are preserved if some nonlinear transformation of the error terms exists. In both cases, we extend the procedure to accommodate an α-stable distribution. However, a formal analysis of these extensions, along with their associated asymptotic theory, is deferred to future research. To maintain a reasonable length for the paper, we report only the GCoV results and include the AML results in the online appendix. Nonetheless, the retained specification, a purely anticipative AR(2), and the coefficients estimated using AML are similar.

Let us begin by recalling the $MAR(p, q)$ model from Lanne and Saikkonen (2011), which imposes a multiplicative representation of the two-sided alpha-stable MA(∞) form in equation (3.2).[14] This representation corresponds to Corollary 3.2, where $\Theta = H = 1$. It is referred to as a mixed-causal autoregressive process, $MAR(p, q)$, assuming an α-stable distributed error term. This process is defined as follows:

$$\Psi(F)\Phi(B)X_t = \varepsilon_t, \qquad \varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta, \sigma, 0), \tag{6.1}$$

where $\Phi(B) = 1 - \phi_1 B - \cdots - \phi_q B^q$, $\Psi(F) = 1 - \psi_1 F - \cdots - \psi_p F^p$, have their roots outside the unit circle so that

$$\Phi(z) \neq 0 \quad \text{for } |z| \leq 1 \quad \text{and} \quad \Psi(z) \neq 0 \quad \text{for } |z| \leq 1.$$

This assumption ensures stationarity in both the causal and noncausal components. A key challenge in this context is the identification of p and q. One solution, proposed by Lanne and Saikkonen (2011) and Hecq et al. (2017b, 2020), is to first select $p_{ap} = p + q$ by relying on the ACF and PACF of the SOI in the training sample, and then confirm the choice using the BIC criterion.[15]

[13] Bootstrap procedures are proposed in Andrews et al. (2009) and further extended in Cavaliere et al. (2017). However, these procedures are applicable only to purely anticipative processes, not to mixed processes.

[14] In the approach of Gourieroux and Jasiak (2023), no such restriction is imposed. In the univariate case, both representations are equivalent. However, in the multivariate case, certain conditions must be met for the equivalence of the two representations (see Giancaterini, 2023).

[15] p_{ap} corresponds to the number of lags in the all-pass representation of the $MAR(p, q)$ model, which is known to correspond exactly to $p + q$ (see Fries and Zakoian, 2019).

Figure 4: ACF and PACF of the in-sample SOI Data

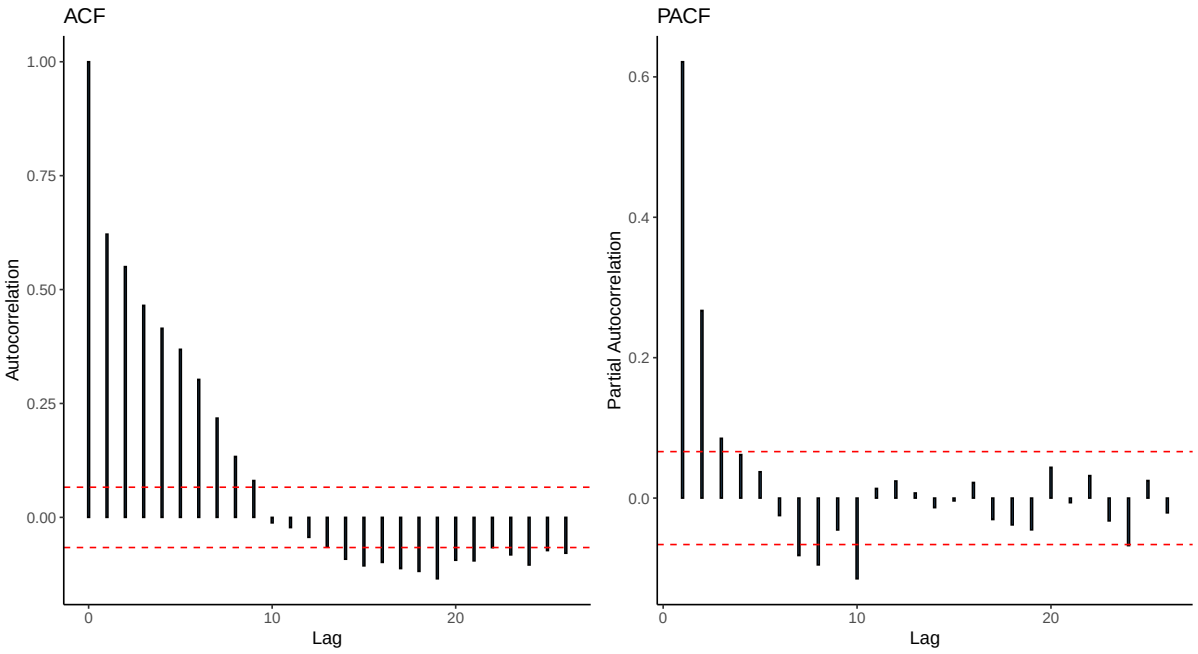


Figure 4 displays the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) for the in-sample SOI data. Based on the PACF, we set $p_{ap} = p+q$, which is equal to 2 in this case. We also apply standard information criteria, such as BIC, to validate this choice.[16] Once p_{ap} is selected, we follow the procedure outlined by Lanne and Saikkonen (2011) and (see Hecq et al., 2020) to select the best specification for the SOI. This is done by estimating all the $2^{p_{ap}} - 1$ possible combinations for $MAR(p,q)$ using the likelihood, along with the ACF of both the residuals and the squared residuals.[17] Theoretically, only the true specification should result in the i.i.d. nature of the residuals. This is reflected in a clear ACF and PACF for both the residuals and the squared residuals. Achieving this is the primary goal of the GCoV estimator. Specifically, this estimator relies on the i.i.d. assumption for the errors as the parameters of interest and minimizes a residual-based portmanteau statistic. It is a one-step estimator that has been shown to be consistent, asymptotically normally distributed, and semi-parametrically efficient. More precisely, the estimator minimizes a portmanteau-type objective function involving the autocovariances of nonlinear transformations of model errors, viewed as functions of the model parameters. If $\theta = (\Phi, \Psi) \in \Theta$ represents the set of parameters of our MAR model from equation 6.1, where Θ is the entire parameter space, the GCoV estimator minimizes the following portmanteau statistic:

[16] Table 14 in Appendix B.3 confirms the selection of 2.

[17] See again Table 14, in the Appendix B.3, which confirms the choice of purely anticipative AR(2)

$$\hat{\theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{h=1}^{\tilde{H}} Tr\left[\hat{\gamma}_g(h;\theta)\hat{\gamma}_g(0;\theta)^{-1}\hat{\gamma}_g(h;\theta)'\hat{\gamma}_g(0;\theta)^{-1}\right] \tag{6.2}$$

where $\tilde{H}$ is the highest selected lag, $\hat{\gamma}_g(h;\theta)$ is the sample autocovariance between $g(\varepsilon_t)$ and $g(\varepsilon_{t-h})$, with $g(\varepsilon_t) = [g_1(\varepsilon_t), \cdots, g_K(\varepsilon_t)]$, where ε_t, is the residuals obtained using equation (6.1). Since only $2\alpha+1$ moments exist for the α-stable law, it is impossible to rely on nonlinear transformations such as ε_t^k with $k > 1$ and k being an integer. This also violates the assumption of the existence of fourth-order moments, as required by the GCoV approach, to be consistent and asymptotically normally distributed. To address the issue of moment existence, as pointed out in Remark 2 on p. 1318 of Gourieroux and Jasiak (2023), if ε_t has no finite fourth-order moment, the consistency and asymptotic properties of the GCoV estimator are preserved if the transformed errors $g_k(\varepsilon_t)$ have finite fourth-order moments. We choose the following nonlinear transformation for the residuals: $g_k = \log(|\varepsilon_t|)^k$, where $k \in \{0, \ldots, K\}$. Indeed, Corollary 3.6 in Nolan (2021) ensures that g_k, $k \in \{0, \ldots, K\}$ exists even if ε_t does not have finite fourth-order moments. With this choice of g_k, we are able to compute the standard errors associated with the parameters in θ using the formula in Corollary 1 of Gourieroux and Jasiak (2023). If θ_0 represents the parameters estimated by GCoV, the standard errors can be computed as described in Gourieroux and Jasiak (2023), requiring the first-order derivative of $\gamma(h;\theta_0)$. We use finite differences to approximate this first-order derivative. Proving and assessing the performance of the GCoV approach for univariate and multivariate α-stable mixed-causal models is left for future research. Recall that the aim of the GCoV approach is to target i.i.d. residuals.[18] We test all $2^{p_{ap}} - 1$ possible combinations for MAR(r, s) with different combinations of $H = \{1, 2, 3\}$ and $K = \{1, 2, 3\}$. Subsequently, we test for the i.i.d.-ness of the residuals, focusing specifically on the autocorrelations of both the residuals and the squared residuals. The only choice of (r, s, K, H) leading to close i.i.d. residuals is the parameter set $(0, 2, 2, 2)$. Figure 6, in Appendix B.2 displays the ACF of the residuals, showing no significant autocorrelation, which is confirmed by the results of the Ljung-Box test on the residuals, as reported in Table 9. Figure 6 also shows the ACF of the squared residuals. A barely significant autocorrelation at lag one is observed, leading to the rejection of the null hypothesis in the Ljung-Box test for the squared residuals (see Table 9). However, we also implement the Portmanteau Test from Jasiak and Neyazi (2023), which is a residual-based specification test for semiparametric models with i.i.d. errors. The i.i.d. nature of the residuals is confirmed by the non-rejection of the null hypothesis for this test (see Table 9).

Table 9 presents the estimated parameters for ψ_1 and ψ_2, which closely resemble those

[18] The residuals are extracted from the SOI data using the estimated parameters and the equation (6.1)

obtained using the AML approach, as detailed in Table 15 in Appendix B.3.

Table 9: AR(2) Estimation for SOI, uisng GCov

ψ_1	ψ_2	α	β	σ
0.4224	0.2924	1.9754	-0.0216	0.4735
(0.0480)	(0.0487)	(0.0018)	(1.6763)	(0.0011)

Residuals test	Stats	$CV_{\alpha=5\%}$	p-value
LB-Test on ε_t (lag=5)	8.5836	7.8147	0.0354
LB-Test on ε_t^2 (lag=5)	17.142	7.8147	0.0007
Jasiak and Neyazi (2023)'s test	5.5933	12.5915	0.4702
JB-Test	21.261	5.9900	0.0000

Notes: Estimated parameters of the α-stable anticipative AR(2) process associated with the SOI series for the period 01/1951–12/1991. Standard deviations are provided in parentheses. The specification tests are as follows: the Ljung-Box test on residuals, the Ljung-Box test on the squared residuals, the Portmanteau test of Jasiak and Neyazi (2023), and a non-Gaussianity test.

GCoV is a semi-parametric approach that does not rely on any distributional assumptions, unlike our approach, which assumes that ε_t is α-stable. First, the Jarque-Bera test in Table 9 confirms that the residuals are indeed non-Gaussian. Next, we fit an α-stable distribution $\mathcal{S}(\alpha, \beta, \sigma, 0)$ to the residuals using the characteristic function-based estimation for α-stable distributions described in Nolan (2021). The estimated parameters are reported in Table 9. These results are consistent with those obtained using the AML approach (see Table 15 in the Appendix B.3). Figure 7, in Appendix B.2, confirms the good fit of the estimated α-stable distribution. However, we encounter an identification issue with the β parameters, though this is not significant. This is unsurprising, as β is inherently difficult to identify. This challenge arises from the fact that it becomes nearly impossible to reliably estimate β when α is close to 2. As α approaches 2, the distribution increasingly resembles a Gaussian distribution, which is symmetric. In this limit, the skewness parameter β has a negligible effect on the shape of the distribution, making it practically unidentifiable due to the dominance of the symmetric properties inherent in the Gaussian limit of α-stable laws.

A La Niña shock is defined as the SOI exceeding 1 for at least three consecutive periods. We then estimate the probability of the SOI returning to central values after h periods, where $h = 3$ and $h = 5$, using the procedure detailed in Section 5.2. We define the same neighborhood $B(V_0)$ of $\frac{\boldsymbol{X_t}}{\|\boldsymbol{X_t}\|}$ and consider $\|\boldsymbol{X_t}\|$ to be large when $X_t \geq q$, where first $q = q_{0.90}$ and second $q = q_{0.95}$ correspond to the 90th and 95th percentiles of the theoretical marginal distribution of X_t, estimated using the parameters from Table 9. Given that we rely on 492 in-sample observations (out of a total of 877), higher quantiles are sparse and cannot be considered here. We choose $\delta = 0.5$. The results are reported in Table 10 and show a high empirical (average)

probability of returning to central values 3 periods ahead. However, this probability is less than unity, indicating that La Niña episodes can occur (or persist if the SOI was already above 1 for several periods). When $h = 5$, $\hat{p}_{0.95} = 84.6154$, meaning that very persistent La Niña occurrences are unlikely to appear.

Table 10: Comparison of theoretical and empirical SOI reversal probabilities

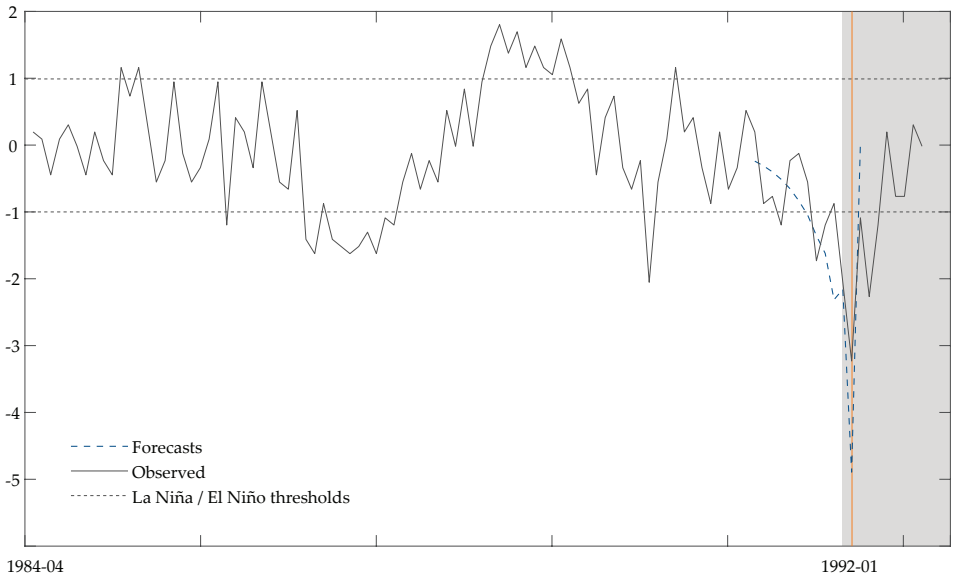
		$h=1$	$h=3$	$h=5$
In-sample	$p_{0.90}\backslash\hat{p}_{0.90}$	32.0755\41.5094	69.8113\64.1509	86.7925 \67.9245
	$p_{0.95}\backslash\hat{p}_{0.95}$	20.8333\54.1667	54.1667\75.0000	83.3333 \87.5000
Out-of-sample	$p_{0.90}\backslash\hat{p}_{0.90}$	32.6733\47.5248	72.2772 \62.3762	93.0693\73.2673
	$p_{0.95}\backslash\hat{p}_{0.95}$	32.6923\53.8462	63.4615 \67.3077	92.3077\84.6154

Notes: The theoretical reversal probabilities p_q are computed using (5.2). Empirical average (Mean) and 95% confidence intervals (95%-CI) of the estimated probabilities are computed using (5.1) on the in-sample period (from 01/1951 to 12/1991) and on out-of-sample period (from 01/1992 to 01/2024) for several a-quantiles of the estimated theoretical marginal distribution of the SOI. These probabilities are reported in percent. We set m=2.

The reversal probabilities are useful to determine the probability of eluding dramatic climatic events such as strong and persistent La Niña or El Niño occurrences. In this context, forecasting the reversal date, that is the end of La Niña or El Niño, is also of particular interest. We hence take advantage of Proposition 4.3 to predict the reversal date of the El Niño occurrence that presumably starts at the end of the in-sample period. As this last observation is below -1, we admit that x is far from central values. Following the methodology of the simulation study (see 5.3), we determine k_0 for various values of $m \in [1, 10]$. As $\hat{\alpha} \approx 1.9$ exhibits light tails, we might encounter some difficulties in applying our pattern recognition procedure: far from the peak (m large) we are more likely to observe values from the center of the distribution. On the other hand, $m = 1$ might lead to imprecise results as few past information is used to determine the piece of trajectory and the process is not strongly anticipative given the estimated coefficients. Our findings offer some robustness in this particular case as for $m = \{1, 2\}$ and $m \in [5, 10]$ our procedure always points toward $k_0 = 1$. For $m = 3$ and $m = 4$ we find $k_0 = 5$ and $k_0 = 3$ respectively. We hence retain $k_0 = 1$ and $m = 10$, therefore implying an imminent reversal date as we are close to the last piece of the trajectory described by $\vartheta_0 \boldsymbol{d}_{k_0}$, with $\vartheta_0 = -1$. The selected piece of the trajectory is represented in Figure 5. We then deduce the reversal date and we compute the future values of X_t up to $X_{t+h} = 0$, with $h = k_0 + 1$, that is when the SOI goes back to its central value. We find that El Niño should reverse just after February 1992, reaching a peak at $\hat{x}_{t+1} = -4.60$. When compared with the out-of-sample period, the reversal date appeared to be very accurately predicted. However, the magnitude of the peak reached during this El Niño occurrence is overestimated as $x_{t+1} = -3.04$.

To ensure the robustness of our approach, following the same procedure used to predict the

Figure 5: El Niño reversal forecast



Notes: The shaded area corresponds to the out-of-sample data.

reversal date in Figure 5, we predict all El Niño and La Niña anomalies in the out-of-sample dataset (from 01/1992 to 01/2024). The results are summarized in Table 11. For an El Niño (or La Niña) event, we start our forecasting procedure at the first date when the SOI is below -1 (or above 1) before the end date of an identified El Niño phenomenon, called the start date in Table 11; the end date is when the SOI returns between -1 and 0 (or 1 and 0). We also forecast the peak date, defined as the minimum (or maximum) value of the SOI before the start and end dates. Table 11 shows that for all El Niño and La Niña occurrences (14) in the out-of-sample dataset, our procedure leads to an average error of 0.42 months in finding the peak date and 0.57 months in finding the end date compared to the true peak and end dates. We also report in Table 11 the selected k_0 and m from our procedure.

Table 11: Forecasting out-of-sample El Niño and La Niña anomalies

Type of anomaly	El Niño	El Niño	La Niña	El Niño	La Niña	La Niña	El Niño	La Niña	La Niña	La Niña	La Niña
Start date	12/1991	07/1994	11/2007	12/2009	07/2010	11/2010	07/2015	11/2021	02/2022	08/2022	11/2022
Peak date	01/1992	09/1994	02/2008	02/2010	09/2010	12/2010	10/2015	01/2021	03/2022	10/2022	12/2022
End date	04/1992	10/1994	03/2008	03/2010	11/2010	04/2011	11/2015	03/2021	05/2022	11/2022	02/2023
Forecasted Peak	01/1992	09/1994	02/2008	03/2010	08/2010	01/2011	09/2015	01/2021	04/2022	10/2022	01/2023
Forecasted End	02/1992	10/1994	03/2008	04/2010	09/2010	02/2011	10/2015	02/2021	05/2022	11/2022	02/2023
Peak forecast error	0	0	0	1	-1	1	-1	0	1	0	-1
End forecast error	-2	0	0	1	-1	-2	-1	-1	0	0	0
k_0	1	2	3	3	1	2	2	2	2	2	2
m	10	10	10	9	10	10	10	10	10	10	10

7 Conclusion

For α-stable infinite moving averages, the conditional distribution of future paths, given the observed past trajectory during extreme events, is derived using a novel spectral representation of stable random vectors on unit cylinders with respect to seminorms. In contrast to conventional norm-based representations, this approach yields a multivariate regularly varying tails property that is well-suited for prediction purposes. However, not all stable random vectors can be represented on seminorm unit cylinders. A representation theorem is provided, showing that predictions are feasible if and only if the process is sufficiently "anticipative." Finite-length paths of α-stable moving averages, which themselves exhibit multivariate α-stable properties, are incorporated into this framework. Our approach also reveals that, despite their appealing "causal" interpretation, non-anticipative processes inherently imply the unpredictability of extreme events. In contrast, anticipative processes operate under the assumption that future events exhibit early visible signs that hint at their forthcoming occurrence. These early signs manifest as emerging trends and patterns that an observer can detect and use to infer potential future outcomes. In certain cases, we demonstrate that the trajectory leaves no room for indeterminacy and can, in theory, be deduced with certainty, and in practice, with a very high level of confidence. We use Monte Carlo simulations to illustrate two applications derived from our theoretical results: forecasting crash probabilities and predicting crash dates. Additionally, we discuss various sources of uncertainty that may arise in finite-sample and non-asymptotic settings. The numerical analysis confirms that both procedures we implement are straightforward to use and perform well across a wide range of scenarios. To provide further insight into the empirical relevance of the seminorm representation of α-stable moving averages, we demonstrate its ability to accurately predict climate anomalies. Specifically, we estimate the probabilities of occurrence for the so-called La Niña and El Niño episodes. We also precisely detect, out-of-sample, the reversal date of these episodes.

References

Alley, R.B., Emanuel, K.A., Zhang, F., 2019. Advances in weather prediction. Science 363, 342–344.

Andrews, B., Calder, M., Davis, R., 2009. Maximum likelihood estimation for α-stable autoregressive process.

Basrak, B., Planinić, H., Soulier, P., 2016. An invariance principle for sums and record times

of regularly varying stationary sequences. Probability Theory and Related Fields , 1–46.

Basrak, B., Segers, J., 2009. Regularly varying multivariate time series. Stochastic processes and their applications 119, 1055–1080.

Behme, A., Lindner, A., Maller, R., 2011. Stationary solutions of the stochastic differential equation with lévy noise. Stochastic Processes and their Applications 121, 91–108.

Behme, A.D., 2011. Distributional properties of solutions of $dv_t = v_t - du_t + dl_t$ with lévy noise. Advances in Applied Probability 43, 688–711.

Brenner, A.D., 2002. El niño and world primary commodity prices: Warm water or hot air? The Review of Economics and Statistics 84, 176–183.

Brockwell, P., Davis, R., 1991. Time Series: Theory and Methods. Springer series in statistics, Springer.

Cashin, P., Mohaddes, K., Raissi, M., 2017. Fair weather or foul? the macroeconomic effects of el niño. Journal of International Economics 106, 37–54.

Cavaliere, G., Nielsen, H.B., Rahbek, A., 2017. Bootstrapping non-causal autoregressions: with applications to explosive bubble modelling. Journal of Business and Economic Statistics .

Chahrour, R., Jurado, K., 2021. Recoverability and expectations-driven fluctuations. The Review of Economic Studies 89, 214–239.

Chen, B., Choi, J., Escanciano, J.C., 2017. Testing for fundamental vector moving average representations. Quantitative Economics 8, 149–180.

Cioczek-Georges, R., Taqqu, M.S., 1994. How do conditional moments of stable vectors depend on the spectral measure? Stochastic processes and their applications 54, 95–111.

Cioczek-Georges, R., Taqqu, M.S., 1998. Sufficient conditions for the existence of conditional moments of stable random variables. Stochastic Processes and Related Topics , 35–67.

Dell, M., Jones, B.F., Olken, B.A., 2014. What Do We Learn from the Weather? The New Climate-Economy Literature. Journal of Economic Literature 52, 740–798.

Dombry, C., Hashorva, E., Soulier, P., 2017. Tail measure and tail spectral process of regularly varying time series. arXiv preprint arXiv:1710.08358 .

Fries, S., 2021. Conditional moments of noncausal alpha-stable processes and the prediction of bubble crash odds. Journal of Business & Economic Statistics 0, 1–21.

Fries, S., Zakoian, J.M., 2019. Mixed causal-noncausal ar processes and the modelling of explosive bubbles. Econometric Theory 35, 1234–1270.

Giancaterini, F., 2023. Essays on univariate and multivariate noncausal processes. Ph.D. thesis. Maastricht University. Netherlands. doi:10.26481/dis.20231204fg.

Giancaterini, F., Hecq, A., Morana, C., 2022. Is climate change time-reversible? Econometrics 10.

Gourieroux, C., Hencic, A., Jasiak, J., 2021a. Forecast performance and bubble analysis in noncausal mar (1, 1) processes. Journal of Forecasting 40, 301–326.

Gourieroux, C., Jasiak, J., 2023. Generalized covariance estimator. Journal of Business & Economic Statistics 41, 1315–1327.

Gourieroux, C., Jasiak, J., 2024. Nonlinear fore(back)casting and innovation filtering for causal-noncausal var models. URL: https://arxiv.org/abs/2205.09922, arXiv:2205.09922.

Gourieroux, C., Jasiak, J., Monfort, A., 2020. Stationary bubble equilibria in rational expectation models. Journal of Econometrics 218, 714–735.

Gourieroux, C., Jasiak, J., Tong, M., 2021b. Convolution-based filtering and forecasting: An application to wti crude oil prices. Journal of Forecasting 40, 1230–1244.

Gouriéroux, C., Jasiak, J., 2016. Filtering, prediction and simulation methods for noncausal processes. Journal of Time Series Analysis 37, 405–430.

Gouriéroux, C., Jasiak, J., 2017. Noncausal vector autoregressive process: Representation, identification and semi-parametric estimation. Journal of Econometrics 200, 118–134.

Gouriéroux, C., Monfort, A., Renne, J.P., 2019. Identification and estimation in non-fundamental structural varma models. The Review of Economic Studies 87, 1915–1953.

Gouriéroux, C., Zakoian, J.M., 2015. On uniqueness of moving average representations of heavy-tailed stationary processes. Journal of Time Series Analysis 36, 876–887.

Gouriéroux, C., Zakoian, J.M., 2017. Local explosion modelling by non-causal process. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 79, 737–756.

Hecq, A., Issler, J., Voisin, E., 2024. A short term credibility index for central banks under inflation targeting: an application to brazil. Journal of International Money and Finance .

Hecq, A., Issler, J.V., Telg, S., 2020. Mixed causal–noncausal autoregressions with exogenous regressors. Journal of Applied Econometrics 35, 328–343.

Hecq, A., Lieb, L., Telg, S.M., 2016. Identification of mixed causal-noncausal models in finite samples. Annals of Economics and Statistics 123/124, 307–331.

Hecq, A., Telg, S., Lieb, L., 2017a. Do seasonal adjustments induce noncausal dynamics in inflation rates? Econometrics 5, 48.

Hecq, A., Telg, S., Lieb, L., 2017b. Simulation, estimation and selection of mixed causal-noncausal autoregressive models: The marx package. SSRN https://ssrn.com/abstract=3015797.

Hecq, A., Voisin, E., 2021. Forecasting bubbles with mixed causal-noncausal autoregressive models. Econometrics and Statistics 20, 29–45.

Hencic, A., Gouriéroux, C., 2015. Noncausal autoregressive model in application to bitcoin/usd exchange rates. Econometrics of Risk , 17–40.

Janßen, A., 2019. Spectral tail processes and max-stable approximations of multivariate regularly varying time series. Stochastic Processes and their Applications 129, 1993–2009.

Janßen, A., Segers, J., 2014. Markov tail chains. Journal of Applied Probability 51, 1133–1153.

Jasiak, J., Neyazi, A.M., 2023. Gcov-based portmanteau test. URL: https://arxiv.org/abs/2312.05373, arXiv:2312.05373.

Lanne, M., Luoto, J., 2013. Autoregression-based estimation of the new Keynesian Phillips curve. Journal of Economic Dynamics and Control 37, 561–570.

Lanne, M., Luoto, J., 2016. Noncausal bayesian vector autoregression. Journal of Applied Econometrics 31, 1392–1406.

Lanne, M., Luoto, J., Saikkonen, P., 2012. Optimal forecasting of noncausal autoregressive time series. International Journal of Forecasting 28, 623–631.

Lanne, M., Saikkonen, P., 2011. Noncausal autogressions for economic time series. Journal of Time Series Econometrics 3.

Lanne, M., Saikkonen, P., 2013. Noncausal vector autoregression. Econometric Theory 29, 447–481.

Makkonen, A., Vallström, D., Uddin, G.S., Rahman, M.L., Haddad, M.F.C., 2021. The effect of temperature anomaly and macroeconomic fundamentals on agricultural commodity futures returns. Energy Economics 100, 105377.

Meinguet, T., Segers, J., 2010. Regularly varying time series in banach spaces. arXiv preprint arXiv:1001.3262 .

Nolan, J., 2021. Univariate Stable Distributions: Models for Heavy Tailed Data. Springer Series in Operations Research and Financial Engineering, Springer International Publishing.

Planinić, H., Soulier, P., 2017. The tail process revisited. Extremes , 1–29.

Samorodnitsky, G., Taqqu, M.S., 1994. Stable non-Gaussian random processes. Chapman & Hall, London.

Velasco, C., Lobato, I.N., 2018. Frequency domain minimum distance inference for possibly noninvertible and noncausal ARMA models. The Annals of Statistics 46, 555 – 579.

A Complementary results for the Monte-Carlo simulation

Table 12: Comparison of theoretical and empirical crash probabilities at horizons $h = 1, 5, 10$ of bubbles generated by the anticipative AR(2)

	$h = 1$	$h = 5$	$h = 10$
$p_{0.9} \backslash \hat{p}_{0.9}$	51.149\20.777 (20.566-21.031)	99.959\56.494 (56.045-56.923)	100.000\73.029 (72.552-73.488)
$p_{0.99} \backslash \hat{p}_{0.99}$	38.023\27.945 (26.966-28.896)	99.856\74.154 (72.688-75.712)	100.000\91.18 (90.062-92.466)
$p_{0.999} \backslash \hat{p}_{0.999}$	30.888\29.826 (26.747-32.895)	99.457\77.736 (73.216-83.068)	100.000\94.492 (90.776-97.368)
$p_{0.9999} \backslash \hat{p}_{0.9999}$	30.076\30.042 (22.697-42.028)	98.379\78.115 (63.28-90.865)	100.000\94.371 (82.953-100.000)

Notes: The theoretical crash probabilities p_q are computed using (5.2). Empirical average (Mean) and 95% confidence intervals (95%-CI) of the estimated probabilities are computed using (5.1) on M = 100 simulated trajectories of $N = 10^6$ observations, for $q = q_a$ several a-quantiles of the marginal distribution of X_t. These probabilities are reported in percent. We set $m = 7$. The simulated process is the following: $X_t = 0.7X_{t+1} + 0.1X_{t+2} + \varepsilon_t$ where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(1.5, 0, 0.5, 0)$

B Complementary results for the empirical application

B.1 Distribution property of the SOI

Table 13: Descriptive Statistics and Jarque-Bera Test Results for SOI Data

Statistic	**Value**
Skewness	-0.2024
Kurtosis	3.4823
Jarque-Bera Test Statistic (X^2)	14.4890
Degrees of Freedom (df)	2
p-value	0.0007141

B.2 GCov Estimation Residuals Analysis

In Figure 6, the left panel shows the ACF of the residuals computed using the GCoV approach, confirming their lack of serial correlation, while the right panel presents the ACF of the squared residuals, indicating only a barely significant autocorrelation at lag one. This confirms the good accuracy of the anticipative AR(2) model.

Figure 7, helps to visually assess the fit of the α-stable model to the residuals. The histogram and the density curve align well, it suggests that the model is appropriately capturing the underlying distribution of the residuals.

Figure 6: ACF of Residuals and Squared Residual using GCoV

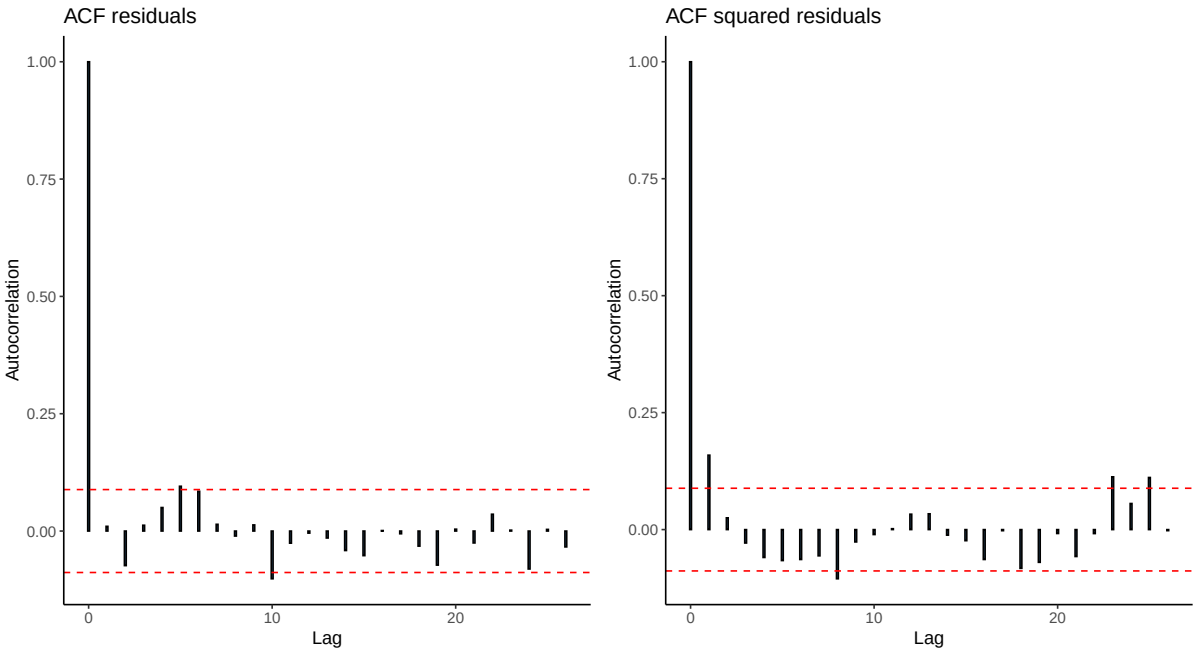
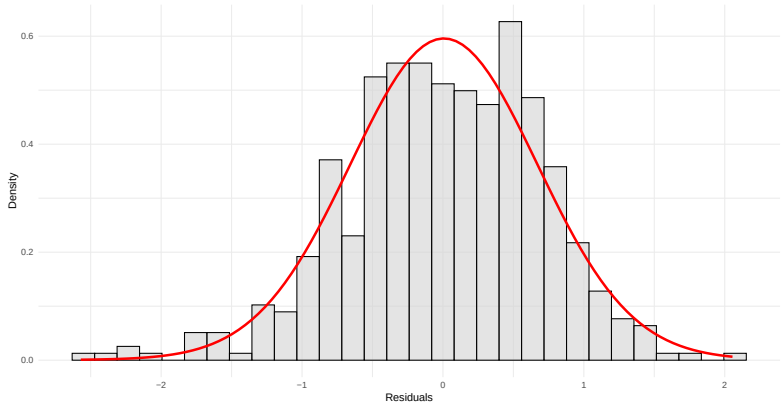


Figure 7: Histogram of residuals using GCoV



Notes: The figure illustrates the histogram of residuals with an overlaid density curve. The red line represents the density of an α-stable law with the parameters: $\alpha = 1.9754$, $\beta = -0.0216$, $\sigma = 0.4735$.

B.3 AML estimation of AR(2)

This section, presents the results on in-sample SOI data using the AML estimation procedure from, Lanne and Saikkonen (2011) and Hecq et al. (2020, 2017b). As in 6, we set $p_{ap} = 2$, and we estimate all the possible $2^{p_{ap}}) - 1$ combinaison of $MAR(p, q)$. More precisely, the parameters of all the $MAR(p, q)$ are subsequently estimated using a modified version of the MARX package suitable for α-stable laws (Hecq et al., 2020, 2017b). For the $MAR(p, q)$ as in eq (6.1) in section 6, the ensemble of parameters to be estimate is $\theta = (\Psi, \Phi, \alpha, \beta, \sigma) \in \Theta$ then the Lanne and

Saikkonen (2011) AML estimator, is defined as:

$$\hat{\theta} = \underset{\Theta}{\operatorname{argmax}} \sum_{t=p+1}^{T-q} \ln f^{-1} \left[\Psi(F)\Phi(B)X_t; (\alpha, \beta, \sigma, 0)\right] \tag{B.1}$$

where $f(.;(\alpha,\beta,\sigma,0))$ denotes the probability function of the ε_t.[19] Standard deviations are estimated using finite differences gradient and Hessian for the parameters in the right space. Table 14, according to the likelihood criteria, shows that the best specification for causal-non-causal models is the anticipative AR(2)

$$(1 - \psi_1 F - \psi_2 F^2)X_t = \varepsilon_t, \ \varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta, \sigma, 0)$$

Table 14: Identification of Non-Causal Processes for SOI

BIC	MAR(2,0)	MAR(1,1)	MAR(0,2)
p_{pseudo}		Likelihood	
2	-513.7138	-511.3617	-508.8777

Table 15 reports the estimation results for the anticipative AR(2) parameters and the associated parameters of the α-stable distribution for the error term. Additionally, we present a set of descriptive statistics and validation tests on the residuals and the squared residuals. From Table 15, the Ljung-Box (LB) tests indicate that all autocorrelation in the residuals of the AR(2) model has been removed. This conclusion is further supported by the ACFs of the residuals, shown in Figure 8. However, Figure 8 reveals a barely significant autocorrelation at lag one in the squared residuals, which is confirmed by the rejection of the null hypothesis for the Ljung-Box test on the squared residuals, as reported in Table 15.

The Jarque-Bera (JB) test in Table 15 indicates that the residuals are indeed non-Gaussian ($\alpha < 2$), this is consistent with an estimated $\alpha = 1.93$ $(1.90E-4)$. However, we have an estimated β of -0.99 $(2.68E-2)$ which is barely significant. This is not surprising as β is hard to identify, explained by the fact that it is impossible to reliably identify the coefficient β when α is too close to 2. This is because, as α gets close to 2, the distribution increasingly resembles a Gaussian distribution, which is symmetric. In this limit, the skewness parameter β has a diminishing effect on the shape of the distribution, rendering it practically unidentifiable due to the dominance of the symmetric properties inherent in the Gaussian limit of α-stable laws. The goodness of fit of the estimated parameters for the α-stable law is confirmed by the associated estimated density and the histogram of the residuals shown in Figure 9.

[19] Lanne and Saikkonen (2011) shows that the AML approach is consistent if f admits a Lebesgue representation, which is the case for the α-stable law.

Table 15: AR(2) AML estimation for SOI

ψ_1	ψ_2	α	β	σ
0.4476	0.2750	1.9390	-0.9970	0.4696
(0.0000)	(0.0000)	(0.0001)	(0.0268)	(0.0000)

Residuals test	Stats	$CV_{\alpha=5\%}$	p-value
LB-Test on ε_t (lag=5)	7.6878	7.8147	0.0529
LB-Test on ε_t^2 (lag=5)	16.641	7.8147	0.0008
JB-Test	20.6780	5.9900	0.0000

Notes: Estimated parameters of α-stable anticipative AR(2) process associated with the SOI series for the period 01/1951 - 12/1991. Standard deviations are in parentheses.

In Figure 8, the left panel shows the ACF of the residuals computed using the AML approach, confirming their lack of serial correlation, while the right panel presents the ACF of the squared residuals, indicating only a barely significant autocorrelation at lag one. This confirms the good accuracy of the anticipative AR(2) model.

Figure 8: ACF of Residuals and Squared Residual using AML

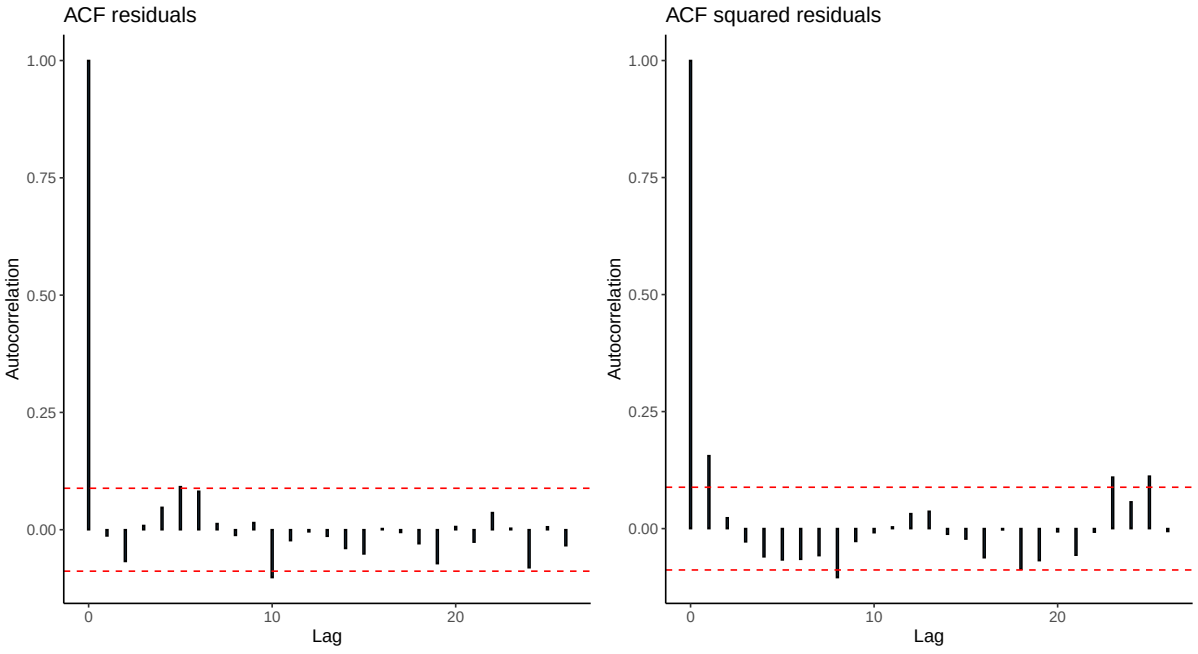
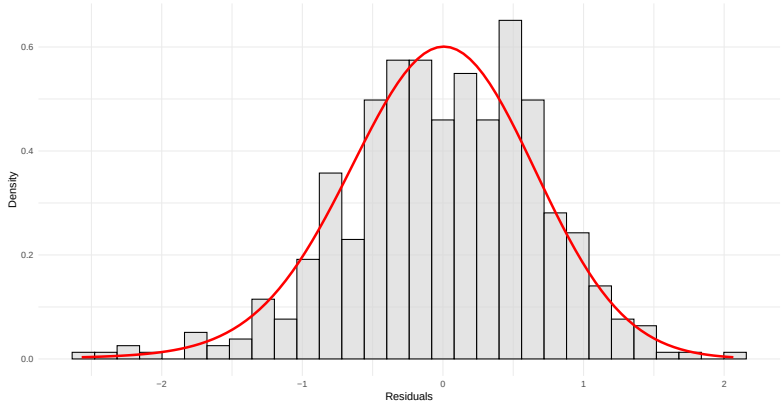


Figure 9, helps to visually assess the fit of the α-stable model to the residuals. The histogram and the density curve align well, it suggests that the model is appropriately capturing the underlying distribution of the residuals.

Figure 9: Histogram of residuals using AML



Notes: The figure illustrates the histogram of residuals with an overlaid density curve. The red line represents the density of an α-stable law with the parameters: $\alpha = 1.9390$, $\beta = -0.9970$, $\sigma = 0.4696$.

C Proofs

C.1 Proof of Proposition 2.1

Consider first the case where either $\alpha \neq 1$ or $\boldsymbol{X}$ is S1S. We only provide the proof for $\alpha \neq 1$ as it is similar under both assumptions.

Assume that $\Gamma(K^{\|\cdot\|}) = 0$ and let us show that $\boldsymbol{X}$ admits a representation of the unit cylinder $C_d^{\|\cdot\|}$ relative to the seminorm $\|\cdot\|$. The characteristic function of $\boldsymbol{X}$ writes for any $\boldsymbol{u} \in \mathbb{R}^d$, with $a = \mathrm{tg}(\pi\alpha/2)$,

$$
\begin{aligned}
\varphi_{\boldsymbol{X}}(\boldsymbol{u}) &= \exp\left\{ -\int_{S_d} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}\rangle)^{<\alpha>}\Big)\Gamma(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \right\} \\
&= \exp\left\{ -\int_{S_d\setminus K^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}\rangle)^{<\alpha>}\Big)\Gamma(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \right\} \\
&= \exp\left\{ -\int_{S_d\setminus K^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|}\rangle)^{<\alpha>}\Big)\|\boldsymbol{s}\|^\alpha\Gamma(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \right\} \\
&= \exp\left\{ -\int_{T_{\|\cdot\|}(S_d\setminus K^{\|\cdot\|})} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}'\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}'\rangle)^{<\alpha>}\Big)\left\|\frac{\boldsymbol{s}'}{\|\boldsymbol{s}'\|_e}\right\|^\alpha \Gamma\circ T_{\|\cdot\|}^{-1}(d\boldsymbol{s}') + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \right\} \\
&= \exp\left\{ -\int_{C_d^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}\rangle)^{<\alpha>}\Big) \underbrace{\|\boldsymbol{s}\|_e^{-\alpha}\Gamma\circ T_{\|\cdot\|}^{-1}(d\boldsymbol{s})}_{\Gamma^{\|\cdot\|}(d\boldsymbol{s})} + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \right\}
\end{aligned}
$$

where we used the change of variable $\boldsymbol{s}' = T_{\|\cdot\|}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|$ between the third and fourth lines, which yields the representation on $C_d^{\|\cdot\|}$.

Reciprocally, assume that $\boldsymbol{X}$ is representable on $C_d^{\|\cdot\|}$. By definition of the representability of $\boldsymbol{X}$ on $C_d^{\|\cdot\|}$, there exists a measure $\gamma^{\|\cdot\|}$ on $C_d^{\|\cdot\|}$ and a non-random vector $\boldsymbol{m}^0_{\|\cdot\|} \in \mathbb{R}^d$ such that

$$
\varphi_{\boldsymbol{X}}(\boldsymbol{u}) = \exp\left\{ -\int_{C_d^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}\rangle)^{<\alpha>}\Big)\gamma^{\|\cdot\|}(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|}\rangle \right\}.
$$

With the change of variable $\boldsymbol{s}' = T^{-1}_{\|\cdot\|}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|_e$,

$$
\begin{aligned}
\varphi_{\boldsymbol{X}}(\boldsymbol{u}) &= \exp\left\{ -\int_{C_d^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e}\rangle)^{<\alpha>}\Big)\|\boldsymbol{s}\|_e^\alpha\gamma^{\|\cdot\|}(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|}\rangle \right\} \\
&= \exp\left\{ -\int_{T^{-1}_{\|\cdot\|}(C_d^{\|\cdot\|})} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}'\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}'\rangle)^{<\alpha>}\Big)\left\|\frac{\boldsymbol{s}'}{\|\boldsymbol{s}'\|}\right\|_e^\alpha \gamma^{\|\cdot\|}\circ T_{\|\cdot\|}(d\boldsymbol{s}') + i\,\langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|}\rangle \right\} \\
&= \exp\left\{ -\int_{S_d\setminus K^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}\rangle)^{<\alpha>}\Big)\|\boldsymbol{s}\|^{-\alpha}\gamma^{\|\cdot\|}\circ T_{\|\cdot\|}(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|}\rangle \right\} \\
&= \exp\left\{ -\int_{S_d\setminus K^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha - ia(\langle \boldsymbol{u}, \boldsymbol{s}\rangle)^{<\alpha>}\Big)\gamma(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|}\rangle \right\},
\end{aligned}
$$

where $\gamma(d\boldsymbol{s}) := \|\boldsymbol{s}\|^{-\alpha}\gamma^{\|\cdot\|} \circ T_{\|\cdot\|}(d\boldsymbol{s})$. Letting now $\overline{\gamma}(A) := \gamma(A \cap (S_d \setminus K^{\|\cdot\|}))$ for any Borel set A of S_d, we have

$$\varphi_{\boldsymbol{X}}(\boldsymbol{u}) = \exp\left\{ -\int_{S_d} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^{\alpha} - ia(\langle \boldsymbol{u}, \boldsymbol{s}\rangle)^{<\alpha>}\Big)\overline{\gamma}(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|}\rangle \right\}.$$

By the unicity of the spectral representation of $\boldsymbol{X}$ on S_d, we necessarily have $(\Gamma, \boldsymbol{\mu}^0) = (\overline{\gamma}, \boldsymbol{m}^0_{\|\cdot\|})$. Thus, $\overline{\gamma}$ and Γ have to coincide, and in particular

$$\Gamma(K^{\|\cdot\|}) = \overline{\gamma}(K^{\|\cdot\|}) = \gamma(K^{\|\cdot\|} \cap (S_d \setminus K^{\|\cdot\|})) = \gamma(\emptyset) = 0.$$

Given that $\Gamma = \overline{\gamma}$ and $\Gamma(K^{\|\cdot\|}) = 0$, we can follow the initial steps of the proof to show that $\gamma^{\|\cdot\|} = \Gamma^{\|\cdot\|}$.

Consider now the case where $\alpha = 1$ and $\boldsymbol{X}$ is not symmetric. Assume first that $\int_{S_d} \Big|\ln\|\boldsymbol{s}\|\Big|\Gamma(d\boldsymbol{s}) < +\infty$, that is, $\Gamma(K^{\|\cdot\|}) = 0$ and $\int_{S_d\setminus K^{\|\cdot\|}} \Big|\ln\|\boldsymbol{s}\|\Big|\Gamma(d\boldsymbol{s}) < +\infty$. With $a = 2/\pi$,

$$\begin{aligned}
\varphi_{\boldsymbol{X}}(\boldsymbol{u}) &= \exp\left\{ -\int_{S_d} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle| + ia\langle \boldsymbol{u}, \boldsymbol{s}\rangle \ln|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|\Big)\Gamma(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \right\} \\
&= \exp\left\{ -\int_{S_d\setminus K^{\|\cdot\|}} \left(|\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|}\rangle| + ia\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|}\rangle \ln|\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|}\rangle|\right)\|\boldsymbol{s}\|\Gamma(d\boldsymbol{s}) \right. \\
&\qquad\qquad \left. + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle - ia\int_{S_d\setminus K^{\|\cdot\|}} \langle \boldsymbol{u}, \boldsymbol{s}\rangle \ln\|\boldsymbol{s}\|\Gamma(d\boldsymbol{s}) \right\}.
\end{aligned}$$

We have $\int_{S_d\setminus K^{\|\cdot\|}} \langle \boldsymbol{u}, \boldsymbol{s}\rangle \ln\|\boldsymbol{s}\|\Gamma(d\boldsymbol{s}) = \sum_{i=1}^d u_i \int_{S_d\setminus K^{\|\cdot\|}} s_i \ln\|\boldsymbol{s}\|\Gamma(d\boldsymbol{s}) = \langle \boldsymbol{u}, \tilde{\boldsymbol{\mu}}\rangle$, and thus,

$$i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle - ia\int_{S_d\setminus K^{\|\cdot\|}} \langle \boldsymbol{u}, \boldsymbol{s}\rangle \ln\|\boldsymbol{s}\|\Gamma(d\boldsymbol{s}) = i\langle \boldsymbol{u}, \boldsymbol{\mu}^0_{\|\cdot\|}\rangle.$$

The condition $\int_{S_d\setminus K^{\|\cdot\|}} \Big|\ln\|\boldsymbol{s}\|\Big|\Gamma(d\boldsymbol{s}) < +\infty$, ensures that $|\boldsymbol{\mu}^0_{\|\cdot\|}| < +\infty$. Again with the change of variable $\boldsymbol{s}' = T_{\|\cdot\|}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|$, we get

$$\begin{aligned}
\varphi_{\boldsymbol{X}}(\boldsymbol{u}) &= \exp\left\{ -\int_{T_{\|\cdot\|}(S_d\setminus K^{\|\cdot\|})} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}'\rangle| + ia\langle \boldsymbol{u}, \boldsymbol{s}'\rangle \ln|\langle \boldsymbol{u}, \boldsymbol{s}'\rangle|\Big) \left\|\frac{\boldsymbol{s}'}{\|\boldsymbol{s}'\|_e}\right\|^{\alpha} \Gamma \circ T^{-1}_{\|\cdot\|}(d\boldsymbol{s}') + i\langle \boldsymbol{u}, \boldsymbol{\mu}^0_{\|\cdot\|}\rangle \right\} \\
&= \exp\left\{ -\int_{C_d^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle| + ia\langle \boldsymbol{u}, \boldsymbol{s}\rangle \ln|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|\Big) \underbrace{\|\boldsymbol{s}\|_e^{-\alpha}\Gamma \circ T^{-1}_{\|\cdot\|}(d\boldsymbol{s})}_{\Gamma^{\|\cdot\|}(d\boldsymbol{s})} + i\langle \boldsymbol{u}, \boldsymbol{\mu}^0_{\|\cdot\|}\rangle \right\}
\end{aligned}$$

Reciprocally, assume there exists a measure $\gamma^{\|\cdot\|}$ on $C_d^{\|\cdot\|}$ satisfying (2.4) and a non-random vector $\boldsymbol{m}^0_{\|\cdot\|} \in \mathbb{R}^d$ such that

$$\varphi_{\boldsymbol{X}}(\boldsymbol{u}) = \exp\left\{ -\int_{C_d^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle| + ia\langle \boldsymbol{u}, \boldsymbol{s}\rangle \ln|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|\Big)\gamma^{\|\cdot\|}(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|}\rangle \right\}.$$

First, we can see that

$$\varphi_{\boldsymbol{X}}(\boldsymbol{u}) = \exp\Bigg\{ - \int_{C_d^{\|\cdot\|}} \Big[\Big(|\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e} \rangle| + ia \langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e} \rangle \ln |\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e} \rangle| \Big) \|\boldsymbol{s}\|_e + ia \langle \boldsymbol{u}, \boldsymbol{s} \rangle \ln \|\boldsymbol{s}\|_e \Big] \gamma^{\|\cdot\|}(d\boldsymbol{s}) + i \langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|} \rangle \Bigg\}.$$

We will later show the following result:

Lemma C.1 *Let $\gamma^{\|\cdot\|}$ a Borel measure on $C_d^{\|\cdot\|}$ satisfying (2.4). Then,*

$$\int_{C_d^{\|\cdot\|}} \|\boldsymbol{s}\|_e \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty. \tag{C.1}$$

Assuming Lemma C.1 holds, then by the Cauchy-Schwarz inequality, we have $\int_{C_d^{\|\cdot\|}} |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty$, and thus

$$\begin{aligned}
\varphi_{\boldsymbol{X}}(\boldsymbol{u}) &= \exp\Bigg\{ - \int_{C_d^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e} \rangle| + ia \langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e} \rangle \ln |\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e} \rangle| \Big) \|\boldsymbol{s}\|_e \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\
&\qquad + i \langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|} \rangle - ia \int_{C_d^{\|\cdot\|}} \langle \boldsymbol{u}, \boldsymbol{s} \rangle \ln \|\boldsymbol{s}\|_e \gamma^{\|\cdot\|}(d\boldsymbol{s}) \Bigg\}, \\
&= \exp\Bigg\{ - \int_{S_d \setminus K^{\|\cdot\|}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}' \rangle| + ia \langle \boldsymbol{u}, \boldsymbol{s}' \rangle \ln |\langle \boldsymbol{u}, \boldsymbol{s}' \rangle| \Big) \gamma(d\boldsymbol{s}') \\
&\qquad + i \langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|} \rangle - ia \int_{S_d \setminus K^{\|\cdot\|}} \langle \boldsymbol{u}, \boldsymbol{s}' \rangle \ln \|\boldsymbol{s}'\| \gamma(d\boldsymbol{s}') \Bigg\},
\end{aligned}$$

where we used the change of variable $\boldsymbol{s}' = T^{-1}_{\|\cdot\|}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|_e$, and $\gamma(d\boldsymbol{s}) := \|\boldsymbol{s}\|^{-1} \gamma^{\|\cdot\|} \circ T_{\|\cdot\|}(d\boldsymbol{s})$. Letting then $\overline{\gamma}(A) := \gamma(A \cap (S_d \setminus K^{\|\cdot\|}))$ for any Borel set A of S_d and $\tilde{\boldsymbol{m}} := (\tilde{m}_i)$ with $\tilde{m}_i = \int_{S_d \setminus K^{\|\cdot\|}} s_i \ln \|\boldsymbol{s}\| \overline{\gamma}(d\boldsymbol{s})$, $j = 1, \ldots, d$, we get

$$\varphi_{\boldsymbol{X}}(\boldsymbol{u}) = \exp\Bigg\{ - \int_{S_d} \Big(|\langle \boldsymbol{u}, \boldsymbol{s} \rangle| + ia \langle \boldsymbol{u}, \boldsymbol{s} \rangle \ln |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big) \overline{\gamma}(d\boldsymbol{s}) + i \langle \boldsymbol{u}, \boldsymbol{m}^0_{\|\cdot\|} - a\tilde{\boldsymbol{m}} \rangle \Bigg\},$$

and $\boldsymbol{X}$ admits the pair $(\overline{\gamma}, \boldsymbol{m}^0_{\|\cdot\|} - a\tilde{\boldsymbol{m}})$ for spectral representation on the Euclidean unit sphere. The unicity of the spectral representation of $\boldsymbol{X}$ on S_d implies that $(\Gamma, \boldsymbol{\mu}^0) = (\overline{\gamma}, \boldsymbol{m}^0_{\|\cdot\|} - a\tilde{\boldsymbol{m}})$. Thus, $\overline{\gamma}$ and Γ have to coincide, and in particular

$$\begin{aligned}
\Gamma(K^{\|\cdot\|}) = \overline{\gamma}(K^{\|\cdot\|}) = \gamma(K^{\|\cdot\|} \cap (S_d \setminus K^{\|\cdot\|})) = \gamma(\emptyset) = 0, \\
\tilde{m}_i = \int_{S_d \setminus K^{\|\cdot\|}} s_i \ln \|\boldsymbol{s}\| \Gamma(d\boldsymbol{s}), \quad i = 1, \ldots, d.
\end{aligned}$$

Last, as $\int_{C_d^{\|\cdot\|}} \|\boldsymbol{s}\|_e \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty$ (Lemma C.1) and $\Gamma(K^{\|\cdot\|}) = 0$, we have by a change

of variable

$$\begin{aligned}\int_{C_d^{\|\cdot\|}} \|\boldsymbol{s}\|_e \Big|\ln \|\boldsymbol{s}\|_e\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) &= \int_{S_d \setminus K^{\|\cdot\|}} \Big|\ln \|\boldsymbol{s}\|\Big| \|\boldsymbol{s}\|^{-1} \gamma^{\|\cdot\|} \circ T_{\|\cdot\|}(d\boldsymbol{s}) \\ &= \int_{S_d \setminus K^{\|\cdot\|}} \Big|\ln \|\boldsymbol{s}\|\Big| \gamma(d\boldsymbol{s}) \\ &= \int_{S_d} \Big|\ln \|\boldsymbol{s}\|\Big| \Gamma(d\boldsymbol{s}) \\ &< +\infty,\end{aligned}$$

which concludes the proof of Proposition 2.1.

Proof of Lemma C.1

Notice that there exists a positive real number b such that for all $\boldsymbol{s} \in C_d^{\|\cdot\|}$, $\|\boldsymbol{s}\|_e \geq b$ because $\|\boldsymbol{s}\| = 1$. Letting $M > 0$, we have for all $\boldsymbol{u} \in \mathbb{R}^d$

$$\int_{C_d^{\|\cdot\|}} \|\boldsymbol{s}\|_e \Big|\ln \|\boldsymbol{s}\|_e\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) = \int_{C_d^{\|\cdot\|} \cap \{b \leq \|\boldsymbol{s}\|_e \leq M\}} + \int_{C_d^{\|\cdot\|} \cap \{\|\boldsymbol{s}\|_e > M\}} := I_1 + I_2.$$

We will show that both I_1 and I_2 are finite. Focus first on I_2. From (2.4), we know that for all $\boldsymbol{u} \in \mathbb{R}^d$

$$\int_{C_d^{\|\cdot\|}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big|\ln |\langle \boldsymbol{u}, \boldsymbol{s}\rangle|\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) = \int_{C_d^{\|\cdot\|} \cap \{b \leq \|\boldsymbol{s}\|_e \leq M\}} + \int_{C_d^{\|\cdot\|} \cap \{\|\boldsymbol{s}\|_e > M\}} < +\infty. \tag{C.2}$$

and thus, in particular

$$\begin{aligned}&\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big|\ln |\langle \boldsymbol{u}, \boldsymbol{s}\rangle|\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\ &\qquad = \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big|\ln \|\boldsymbol{s}\|_e + \ln |\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e}\rangle|\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty.\end{aligned} \tag{C.3}$$

By the triangular inequality, for all $\boldsymbol{u} \in \mathbb{R}^d$,

$$\begin{aligned}&\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big|\ln \|\boldsymbol{s}\|_e + \ln |\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e}\rangle|\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\ &\qquad = \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big|\ln \|\boldsymbol{s}\|_e\Big| \Big|1 + \frac{\ln |\langle \boldsymbol{u}, \boldsymbol{s}/\|\boldsymbol{s}\|_e\rangle|}{\ln \|\boldsymbol{s}\|_e}\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\ &\qquad \geq \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big|\ln \|\boldsymbol{s}\|_e\Big| \Big|1 - \Big|\frac{\ln |\langle \boldsymbol{u}, \boldsymbol{s}/\|\boldsymbol{s}\|_e\rangle|}{\ln \|\boldsymbol{s}\|_e}\Big|\Big| \gamma^{\|\cdot\|}(d\boldsymbol{s})\end{aligned} \tag{C.4}$$

Let us now partition the space $\mathbb{R}^d$ into subsets $R_1, \ldots, R_d$ such that, for any $i = 1, \ldots, d$ and any $\boldsymbol{s} = (s_1, \ldots, s_d) \in R_i$, $\sup_j |s_j| = |s_i|$.[20] We have by (C.3)-(C.4) that for any $i = 1, \ldots, d$,

[20] Strictly speaking, $(R_1, \ldots, R_d)$ is not a partition of $\mathbb{R}^d$ as the R_i's may intersect because of ties in the components of vectors. This will not affect the proof.

any $\boldsymbol{u} \in \mathbb{R}^d$,

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ \|\boldsymbol{s}'\|_e > M\} \cap R_i} |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \ln \|\boldsymbol{s}\|_e \Big| \Big| 1 - \left| \frac{\ln |\langle \boldsymbol{u}, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|}{\ln \|\boldsymbol{s}\|_e} \right| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty.$$

Denoting $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d)$ the canonical orthonormal basis of $\mathbb{R}^d$, evaluate now the above at $\boldsymbol{u} = \boldsymbol{e}_i$. We get that

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ \|\boldsymbol{s}'\|_e > M\} \cap R_i} |\langle \boldsymbol{e}_i, \boldsymbol{s} \rangle| \Big| \ln \|\boldsymbol{s}\|_e \Big| \Big| 1 - \left| \frac{\ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|}{\ln \|\boldsymbol{s}\|_e} \right| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty. \tag{C.5}$$

Let us show that $\boldsymbol{s} \longmapsto \ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|$ is a bounded function for $\boldsymbol{s} \in \{\boldsymbol{s}' \in C_d^{\|\cdot\|} : \|\boldsymbol{s}'\|_e > M\} \cap R_i$. *Ad absurdum*, if it is not bounded, then for any $A > 0$, there exists $\boldsymbol{s} \in \{\boldsymbol{s}' \in C_d^{\|\cdot\|} : \|\boldsymbol{s}'\|_e > M\} \cap R_i$ such that

$$\Big| \ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle| \Big| > A.$$

Taking the sequence $A_n = n$ for any $n \geq 1$, we get that there exists a sequence $(\boldsymbol{s}_n)$, $\boldsymbol{s}_n \in \{\boldsymbol{s}' \in C_d^{\|\cdot\|} : \|\boldsymbol{s}'\|_e > M\} \cap R_i$ such that

$$\Big| \ln |\langle \boldsymbol{e}_i, \boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \rangle| \Big| > n.$$

Thus, for all $n \geq 1$

$$0 \leq |\langle \boldsymbol{e}_i, \boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \rangle| \leq e^{-n}.$$

and

$$|\langle \boldsymbol{e}_i, \boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \rangle| \underset{n \to +\infty}{\longrightarrow} 0.$$

Consider now the decomposition of $\boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e$ in the orthonormal basis $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d)$,

$$\boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e = \sum_{j=1}^{d} \langle \boldsymbol{e}_j, \boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \rangle \boldsymbol{e}_j.$$

As $\boldsymbol{s}_n \in R_i$ for all $n \geq 1$, we also have that $\boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \in R_i$ for all $n \geq 1$, and thus, for any $j = 1, \ldots, d$

$$0 \leq |\langle \boldsymbol{e}_j, \boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \rangle| \leq |\langle \boldsymbol{e}_i, \boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \rangle| \underset{n \to +\infty}{\longrightarrow} 0.$$

Hence, $\boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \underset{n \to +\infty}{\longrightarrow} 0$, which is impossible since $\Big\| \boldsymbol{s}_n/\|\boldsymbol{s}_n\|_e \Big\|_e = 1$ for all $n \geq 1$. The function $\boldsymbol{s} \longmapsto \ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|$ is thus bounded on $\{\boldsymbol{s} \in C_d^{\|\cdot\|} : \|\boldsymbol{s}\|_e > M\} \cap R_i$, say $\Big| \ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle| \Big| \leq A$ for some $A > 0$. Provided M is taken large enough (e.g., $M > 2A$), we will have in (C.5)

$$\left| 1 - \left| \frac{\ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|}{\ln \|\boldsymbol{s}\|_e} \right| \right| = 1 - \left| \frac{\ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|}{\ln \|\boldsymbol{s}\|_e} \right| \geq 1 - \frac{A}{M} > 0,$$

which thus yields for all $i = 1, \dots, d$

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ \|\boldsymbol{s}'\|_e > M\} \cap R_i} |\langle \boldsymbol{e}_i, \boldsymbol{s} \rangle| \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty.$$

As $|\langle \boldsymbol{e}_i, \boldsymbol{s} \rangle| \geq \|\boldsymbol{s}\|_e e^{-A}$, we further get that

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ \|\boldsymbol{s}'\|_e > M\} \cap R_i} \|\boldsymbol{s}\|_e \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty,$$

and because $\bigcup_{i=1,\dots,d} R_i = \mathbb{R}^d$,

$$\begin{aligned} I_2 &= \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ \|\boldsymbol{s}'\|_e > M\}} \|\boldsymbol{s}\|_e \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\ &\leq \sum_{i=1}^{d} \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ \|\boldsymbol{s}'\|_e > M\} \cap R_i} \|\boldsymbol{s}\|_e \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty. \end{aligned}$$

Let us now show that I_1 is finite. Assuming for a moment that

$$\gamma^{\|\cdot\|}\Big(\{\boldsymbol{s}' \in C_d^{\|\cdot\|} :\ b \leq \|\boldsymbol{s}'\|_e \leq M\}\Big) < +\infty,$$

we get

$$\begin{aligned} I_1 &= \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ b \leq \|\boldsymbol{s}'\|_e \leq M\}} \|\boldsymbol{s}\|_e \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\ &\leq \Big(\max_{x \in [b,M]} x |\ln x| \Big)\ \gamma^{\|\cdot\|}\Big(\{\boldsymbol{s}' \in C_d^{\|\cdot\|} :\ b \leq \|\boldsymbol{s}'\|_e \leq M\}\Big), \end{aligned}$$

because $x \longmapsto x|\ln x|$ is a bounded function on $[b, M]$, and thus $I_1 < +\infty$. We now show that $\gamma^{\|\cdot\|}$ is indeed finite on the set $\{\boldsymbol{s}' \in C_d^{\|\cdot\|} :\ b \leq \|\boldsymbol{s}'\|_e \leq M\}$.
Proceeding as in the case of I_2, it can be obtained that for $i = 1, \dots, d$, the function $\boldsymbol{s} \longmapsto \ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|$ is bounded on the set $\{\boldsymbol{s}' \in C_d^{\|\cdot\|} :\ b \leq \|\boldsymbol{s}'\|_e \leq M\} \cap R_i$. Say, again, that $\Big| \ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle| \Big| \leq A$ for some $A > 0$. Then, $|\langle \boldsymbol{e}_i, \boldsymbol{s} \rangle| \geq \|\boldsymbol{s}\|_e e^{-A}$, and for any $\lambda > 2b^{-1}e^A$, we have

$$|\langle \lambda \boldsymbol{e}_i, \boldsymbol{s} \rangle| \geq 2,$$

for any $i = 1, \dots, d$, $\boldsymbol{s} \in \{\boldsymbol{s}' \in C_d^{\|\cdot\|} :\ b \leq \|\boldsymbol{s}'\|_e \leq M\} \cap R_i$. From (C.2), we have for any $\boldsymbol{u} \in \mathbb{R}^d$

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ b \leq \|\boldsymbol{s}'\|_e \leq M\}} |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \ln |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty,$$

and thus, for any $\boldsymbol{u} \in \mathbb{R}^d$,

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\ b \leq \|\boldsymbol{s}'\|_e \leq M\} \cap R_i} |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \ln |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty,$$

for any $i = 1, \dots, d$. Evaluating the above in particular at $\boldsymbol{u} = \lambda \boldsymbol{e}_i$, for any $\lambda > 2b^{-1}e^A$, we get

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, b \le \|\boldsymbol{s}'\|_e \le M\} \cap R_i} |\langle \lambda \boldsymbol{e}_i, \boldsymbol{s} \rangle| \Big| \ln |\langle \lambda \boldsymbol{e}_i, \boldsymbol{s} \rangle| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty.$$

Noticing that $x \longmapsto x|\ln x|$ is increasing on $[1, +\infty)$ and that $|\langle \lambda \boldsymbol{e}_i, \boldsymbol{s} \rangle| \ge 2$ for any $\boldsymbol{s}$ in the domain of integration, we have $|\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \ln |\langle \boldsymbol{u}, \boldsymbol{s} \rangle| \Big| \ge 2 \ln 2$, and

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, b \le \|\boldsymbol{s}'\|_e \le M\} \cap R_i} \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty,$$

for any $i = 1, \dots, d$. Hence,

$$\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, b \le \|\boldsymbol{s}'\|_e \le M\}} \gamma^{\|\cdot\|}(d\boldsymbol{s}) \le \sum_{i=1}^{d} \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|}:\, b \le \|\boldsymbol{s}'\|_e \le M\} \cap R_i} \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty,$$

and $\gamma^{\|\cdot\|}\Big(\{\boldsymbol{s}' \in C_d^{\|\cdot\|} :\, b \le \|\boldsymbol{s}'\|_e \le M\}\Big)$ is finite. □

C.2 Proof of Proposition 2.2

The proposition is an immediate consequence of Bayes formula and of the following result, which is an adaptation of Theorem 4.4.8 by Samorodnitsky and Taqqu (1994) Samorodnitsky and Taqqu (1994) to seminorms.

Proposition C.1 *Let $\boldsymbol{X} = (X_1, \dots, X_d)$ be an α-stable random vector and let $\|\cdot\|$ be a seminorm on $\mathbb{R}^d$ such that $\boldsymbol{X}$ is representable on $C_d^{\|\cdot\|}$. Then, for every Borel set $A \subseteq C_d^{\|\cdot\|}$ with $\Gamma^{\|\cdot\|}(\partial A) = 0$,*

$$\lim_{x \to +\infty} x^\alpha \mathbb{P}\Big(\|\boldsymbol{X}\| > x, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big) = C_\alpha \Gamma^{\|\cdot\|}(A), \tag{C.6}$$

with $C_\alpha = \dfrac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)}$ if $\alpha \neq 1$, and $C_1 = 2/\pi$.

Proof.

We follow the proof of Theorem 4.4.8 by Samorodnitsky and Taqqu (1994). The main hurdle is to show that, with $\|\cdot\|$ a seminorm, $K^{\|\cdot\|} = \{\boldsymbol{s} \in S_d :\, \|\boldsymbol{s}\| = 0\}$, and $\Gamma^{\|\cdot\|}(K^{\|\cdot\|}) = 0$, we have the series representation of $\boldsymbol{X}$, $(X_1, \dots, X_d) \overset{d}{=} (Z_1, \dots, Z_d)$ where

$$Z_k = (C_\alpha \Gamma^{\|\cdot\|}(C_d^{\|\cdot\|}))^{1/\alpha} \sum_{i=1}^{\infty} [\Gamma_i^{-1/\alpha} S_i^{(k)} - b_{i,k}(\alpha)], \quad k = 1, \dots, d, \tag{C.7}$$

with $\boldsymbol{S}_i = (S_i^{(1)}, \dots, S_i^{(d)})$, $i \ge 1$, are i.i.d. $C_d^{\|\cdot\|}$-valued random vectors with common law $\Gamma^{\|\cdot\|}/\Gamma^{\|\cdot\|}(C_d^{\|\cdot\|})$ and the $b_{i,k(\alpha)}$'s are constants.

By Proposition 2.1, we know that $\boldsymbol{X}$ admits a characteristic function of the form (2.1). This allows to restate the integral representation Theorem 3.5.6 in Samorodnitsky and Taqqu

(1994) on the seminorm unit cylinder as follows: with the measurable space $(E, \mathcal{E}) = (C_d^{\|\cdot\|}, \text{Borel } \sigma\text{-algebra on } C_d^{\|\cdot\|})$, let M be an α-stable random measure on $(E, \mathcal{E})$ with control measure $m = \Gamma^{\|\cdot\|}$, skewness intensity $\beta(\cdot) \equiv 1$ (see Definition 3.3.1 in Samorodnitsky and Taqqu (1994) for details). Letting also $f_j : C_d^{\|\cdot\|} \longrightarrow \mathbb{R}$ defined by $f_j\Big((s_1, \ldots, s_d)\Big) = s_j$, $j = 1, \ldots, d$, then

$$\boldsymbol{X} \stackrel{d}{=} \left(\int_{C_d^{\|\cdot\|}} f_1(\boldsymbol{s}) M(d\boldsymbol{s}), \ldots, \int_{C_d^{\|\cdot\|}} f_d(\boldsymbol{s}) M(d\boldsymbol{s}) \right) + \boldsymbol{\mu}^{\|\cdot\|}.$$

This representation can be checked directly by comparing the characteristic functions of the left-hand and right-hand sides. We can now apply Theorem 3.10.1 in Samorodnitsky and Taqqu (1994) to the above integral representation with $(E, \mathcal{E}, m)$ the measure space as described before, and $\hat{m} = \Gamma^{\|\cdot\|}/\Gamma^{\|\cdot\|}(C_d^{\|\cdot\|})$. This establishes (C.7). The rest of the proof is similar to that of Theorem 4.4.8 in Samorodnitsky and Taqqu (1994). We rely on the triangle inequality property of seminorms and the fact that any norm is finer than any seminorm in finite dimension.[21] □

C.3 Proof of Theorem 3.1

From Proposition 2.1, we know that a necessary condition for the representability of $\boldsymbol{X_t}$ on $C_{m+h+1}^{\|\cdot\|}$ is $\Gamma(K^{\|\cdot\|}) = 0$, where $K^{\|\cdot\|} = \{\boldsymbol{s} \in S_{m+h+1} : \|\boldsymbol{s}\| = 0\}$. This condition is also sufficient when either $\alpha \neq 1$ or $\alpha = 1$, $\beta = 0$. Using the fact that Γ only charges discrete atoms on $C_{m+h+1}^{\|\cdot\|}$,

$$\begin{aligned}
\Gamma(K^{\|\cdot\|}) = 0 &\iff \{\boldsymbol{s} \in S_{m+h+1} : \Gamma(\{\boldsymbol{s}\}) > 0\} \cap K^{\|\cdot\|} = \emptyset \\
&\iff \forall \boldsymbol{s} \in S_{m+h+1}, \quad \Big[\Gamma(\{\boldsymbol{s}\}) > 0 \Longrightarrow \|\boldsymbol{s}\| > 0\Big] \\
&\iff \forall k \in \mathbb{Z}, \quad \Big[\|\boldsymbol{d}_k\|_e > 0 \Longrightarrow \|\boldsymbol{d}_k\| > 0\Big] \\
&\iff \forall k \in \mathbb{Z}, \quad \Big[\|\boldsymbol{d}_k\| = 0 \Longrightarrow \|\boldsymbol{d}_k\|_e = 0\Big] \\
&\iff \forall k \in \mathbb{Z}, \quad \Big[\|\boldsymbol{d}_k\| = 0 \Longrightarrow \boldsymbol{d}_k = 0\Big] \\
&\iff \forall k \in \mathbb{Z}, \quad \Big[(d_{k+m}, \ldots, d_k) = \boldsymbol{0} \Longrightarrow (d_{k+m}, \ldots, d_{k-h}) = \boldsymbol{0}\Big],
\end{aligned}$$

by (3.1). Now assume that the following holds:

$$\forall k \in \mathbb{Z}, \quad \Big[(d_{k+m}, \ldots, d_k) = \boldsymbol{0} \Longrightarrow (d_{k+m}, \ldots, d_{k-h}) = \boldsymbol{0}\Big]. \tag{C.8}$$

Then, if for some particular $k_0 \in \mathbb{Z}$, we have

$$(d_{k_0+m}, \ldots, d_{k_0}) = \boldsymbol{0}.$$

[21] We say that a norm N is finer than a seminorm N_s if there is a positive constant C such that $N_s(x) \leq CN(x)$ for any $x \in \mathbb{R}^d$.

It implies that

$$(d_{k_0+m}, \dots, d_{k_0-h}) = \mathbf{0},$$

and especially, as we assume $h \geq 1$,

$$(d_{(k_0-1)+m}, \dots, d_{k_0-1}) = \mathbf{0}.$$

Invoking (C.8), we deduce by recurrence that for any $n \geq 0$,

$$(d_{(k_0-n)+m}, \dots, d_{k_0-n}) = \mathbf{0}.$$

Therefore, (C.8) implies

$$\forall k \in \mathbb{Z}, \quad \Big[(d_{k+m}, \dots, d_k) = \mathbf{0} \Longrightarrow \forall \ell \leq k-1, \quad d_\ell = 0\Big]$$

The reciprocal is clearly true. This establishes that (3.7) is a necessary and sufficient condition for $\boldsymbol{X}_t$ to be representable on $C_d^{\|\cdot\|}$ in the cases where either $\alpha \neq 1$, or $\alpha = 1$, $\beta = 0$.

In the case $\alpha = 1$, $\beta \neq 0$, Proposition 2.1 states that the necessary and sufficient condition for representability reads $\int_{S_d} \Big|\ln \|\boldsymbol{s}\|\Big| \Gamma(d\boldsymbol{s}) < +\infty$. That is

$$\Gamma(K^{\|\cdot\|}) = 0 \qquad \text{and} \qquad \int_{S_d \setminus K^{\|\cdot\|}} \Big|\ln \|\boldsymbol{s}\|\Big| \Gamma(d\boldsymbol{s}) < +\infty.$$

Substituting Γ by its expression in (3.6), the above condition holds if and only if (3.7) is true and

$$\sigma \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_\vartheta \|\boldsymbol{d}_k\|_e \left| \ln \left\| \frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|_e} \right\| \right| < +\infty,$$

the latter being equivalent to

$$\sum_{k \in \mathbb{Z}} \|\boldsymbol{d}_k\|_e \left| \ln \frac{\|\boldsymbol{d}_k\|}{\|\boldsymbol{d}_k\|_e} \right| < +\infty.$$

C.4 Proof of Proposition 3.1

By Definition 3.1, (X_t) is past-representable if and only if there exists $m \geq 0$, $h \geq 1$ such that the vector $(X_{t-m}, \dots, X_t, X_{t+1}, \dots, X_{t+h})$ is representable on $C_{m+h+1}^{\|\cdot\|}$. Consider first point $(\iota)(a)$, that is, the case $\alpha \neq 1$, $(\alpha, \beta) = (1, 0)$. By Theorem 3.1,

$$\begin{aligned}(X_t) \text{ is past-representable} &\iff \text{There exist } m \geq 0,\ h \geq 1, \text{ such that (3.7) holds} \\ &\iff \exists m \geq 0, \forall k \in \mathbb{Z}, \Big[d_{k+m} = \ldots = d_k = 0 \implies \forall \ell \leq k-1,\ d_\ell = 0\Big].\end{aligned}$$

Thus,

$$\begin{aligned}(X_t) \text{ not past-representable} &\iff \forall m \geq 0, \exists k \in \mathbb{Z}, d_{k+m} = \ldots = d_k = 0 \text{ and } \exists \ell \leq k-1,\ d_\ell \neq 0 \\ &\iff \forall m \geq 0, \exists k \in \mathbb{Z}, d_{k+m} = \ldots = d_k = 0 \text{ and } d_{k-1} \neq 0 \\ &\iff \forall m \geq 1, \exists k \in \mathbb{Z}, d_{k+m} = \ldots = d_{k+1} = 0 \text{ and } d_k \neq 0 \\ &\iff \sup\{m \geq 1 :\ \exists k \in \mathbb{Z},\ d_{k+m} = \ldots = d_{k+1} = 0,\ d_k \neq 0\} = +\infty,\end{aligned}$$

hence (3.10).

Regarding the last statement of point $(\iota)(a)$, assume first that $m_0 < +\infty$ and $m \geq m_0$. Property (3.7) necessarily holds with m_0. Indeed, if it did not, there would exist $k \in \mathbb{Z}$ such that

$$d_{k+m_0} = \ldots = d_k = 0, \quad \text{and} \quad d_\ell \neq 0, \quad \text{for some} \quad \ell \leq k-1,$$

and we would have found a sequence of consecutive zero values of length at least m_0+1 preceded by a non-zero value, contradicting the fact that

$$m_0 = \sup\{m \geq 1 :\ \exists\, k \in \mathbb{Z},\ d_{k+m} = \ldots = d_{k+1} = 0, \quad \text{and} \quad d_k \neq 0\}.$$

As (3.7) holds with m_0, it holds *a fortiori* for any $m' \geq m_0$. Thus, $\boldsymbol{X}_t = (X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$ is representable for any $m' \geq m_0$, $h \geq 1$ by Theorem 3.1, and (X_t) is in particular (m,h)-past-representable.

Reciprocally let $m \geq 0$, $h \geq 1$ and assume that (X_t) is (m,h)-past-representable. The process (X_t) is thus in particular past-representable, which as we have shown previously, implies that $m_0 < +\infty$. *Ad absurdum*, suppose now that $0 \leq m < m_0 < +\infty$. If $m_0 = 0$, there is nothing to do. Otherwise if $m_0 \geq 1$, by definition, there exists a $k \in \mathbb{Z}$ such that

$$d_{k+m_0} = \ldots = d_{k+1} = 0, \quad \text{and} \quad d_k \neq 0. \tag{C.9}$$

Because (X_t) is (m,h)-past-representable, we have by Theorem 3.1 that (3.7) holds with m. As $m < m_0$ and $d_{k+m_0} = \ldots = d_{k+1} = 0$, we thus have that $d_\ell = 0$ for all $\ell \leq k+1$, and in particular $d_k = 0$, hence the contradiction. We conclude that $m \geq m_0$.

Consider now point $(\iota)(b)$, i.e., the case $\alpha = 1$ and $\beta \neq 0$. From Theorem 3.1,

$$(X_t) \text{ is past-representable} \iff \text{There exist } m \geq 0,\ h \geq 1, \text{ such that (3.7) and (3.8) hold}$$

From the previous proof, we moreover have that

$$\exists\, m \geq 0, \text{ such that (3.7) holds} \iff m_0 < +\infty \iff \begin{cases} m_0 < +\infty \\ \forall\, m' \geq m_0,\ \text{(3.7) holds} \\ \forall\, m' < m_0,\ \text{(3.7) does not hold} \end{cases}$$

Hence

$$\exists\, m \geq 0,\ h \geq 1, \text{ such that (3.7) and (3.8) hold}$$
$$\iff \begin{cases} m_0 < +\infty \\ \forall\, m' \geq m_0, \text{ (3.7) holds} \\ \forall\, m' < m_0, \text{ (3.7) does not hold} \\ \exists\, m \geq 0,\ h \geq 1, \text{ such that (3.7) and (3.8) hold.} \end{cases}$$

The latter in particular implies $m_0 < +\infty$ and the existence of $m \geq m_0$, $h \geq 1$ such that (3.8) holds. Reciprocally,

$$\begin{cases} m_0 < +\infty \\ \exists\, m \geq m_0,\ h \geq 1, \text{ such that (3.8) holds} \end{cases}$$
$$\implies \begin{cases} m_0 < +\infty \\ \forall\, m' \geq m_0, \text{ (3.7) holds} \\ \exists\, m \geq m_0,\ h \geq 1, \text{ such that (3.8) holds,} \end{cases}$$

which in particular implies that there exists $m \geq m_0$, $h \geq 1$ such that both (3.7) and (3.8) hold. Hence the past-representability of (X_t).

In view of Definition 3.1, point $(\iota\iota)$ is a direct consequence of the second part of Proposition 2.1.

C.5 Proof of Corollary 3.1

Letting k_0 be the greatest integer such that $d_{k_0} \neq 0$ (such an index exists by (3.3)), then immediately, for any $m \geq 1$, $d_{k_0+m} = \ldots = d_{k_0+1} = 0$ and therefore $m_0 = +\infty$.

C.6 Proof of Corollary 3.2

We first show that $\deg(\Psi) \geq 1$ if and only if $m_0 < +\infty$.

Clearly, if $\deg(\Psi) = 0$, then $X_t = \sum_{k=-\infty}^{k_0} d_k \varepsilon_{t+k}$ for some k_0 in $\mathbb{Z}$ and $m_0 = +\infty$.

Reciprocally, assume $\deg(\Psi) = p \geq 1$. Let us first show that (3.10) holds.
Denote $\Psi(F)\Phi(B) = \sum_{i=-q}^{p} \varphi_i F^i$ and $\Theta(F)H(B) = \sum_{k=-r}^{s} \theta_i F^i$, for any non-negative degrees $q = \deg(\Phi)$, $r = \deg(H)$, $s = \deg(\Theta)$. From the recursive equation satisfied by (X_t), we have

that

$$\begin{aligned} & \sum_{i=-q}^{p} \varphi_i X_{t+i} = \sum_{k=-r}^{s} \theta_k \varepsilon_{t+k} \\ \iff \quad & \sum_{i=-q}^{p} \varphi_i \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k+i} = \sum_{k=-r}^{s} \theta_k \varepsilon_{t+k} \\ \iff \quad & \sum_{k\in\mathbb{Z}} \Big(\sum_{i=-q}^{p} \varphi_i d_{k-i} \Big) \varepsilon_{t+k} = \sum_{k=-r}^{s} \theta_k \varepsilon_{t+k}. \end{aligned} \tag{C.10}$$

Proceeding by identification using the uniqueness of representation of heavy-tailed moving averages (see Gouriéroux and Zakoian (2015)), we get that for $|k| > \max(r, s)$,

$$\sum_{i=-q}^{p} \varphi_i d_{k-i} = 0. \tag{C.11}$$

Ad absurdum, if (X_t) is not past-representable, then by Proposition 3.1

$$\sup\{m \geq 1 : \ \exists k \in \mathbb{Z}, \ d_{k+m} = \ldots = d_{k+1} = 0, \ d_k \neq 0\} = +\infty.$$

Thus, there exists a sequence $\{m_n : \ n \geq 0\}$, $m_n \geq 1$, $\lim_{n\to+\infty} = +\infty$, satisfying: for any $n \geq 0$, there is an index $k \in \mathbb{Z}$ such that

$$d_{k-p} \neq 0 \quad \text{and} \quad d_{k-p+1} = d_{k-p+2} = \ldots = d_{k+m_n} = 0.$$

We can therefore construct a sequence (k_n) such that the above relation holds for all $n \geq 0$. This sequence of integers in $\mathbb{Z}$ is either bounded or unbounded. We will show that both cases lead to a contradiction.

First case: $\sup\{|k_n| : n \geq 0\} = +\infty$

There are two subsequences such that $m_{g(n)} \longrightarrow +\infty$ and $|k_{g(n)}| \longrightarrow +\infty$. For some n large enough such that (C.11) holds and $m_{g(n)} \geq p + q$, we have both

$$\sum_{i=-q}^{p} \varphi_i d_{k_{g(n)}-i} = 0.$$

and

$$d_{k_{g(n)}-p} \neq 0, \quad d_{k_{g(n)}-p+1} = \ldots = d_{k_{g(n)}+q} = 0.$$

Hence,

$$\varphi_p d_{k_{g(n)}-p} = 0,$$

which is impossible given that $d_{k_{g(n)}-p} \neq 0$ and $\varphi_p \neq 0$. Indeed, denoting $\Psi(z) = 1 + \psi_1 z + \ldots + \psi_p z^p$, $\psi_p \neq 0$ because $\deg(\Psi) = p$, it can be shown that $\varphi_p = \psi_p$.

Second case: $\sup\{|k_n| : n \geq 0\} < +\infty$

Given that (k_n) is a bounded sequence, there exists by the Bolzano-Weierstrass theorem a convergent subsquence $(k_{g(n)})$. As $(k_{g(n)})$ takes only discrete values, it necessarily holds that $(k_{g(n)})$ reaches its limit at a finite integer $n_0 \geq 1$, that is, for all $n \geq n_0$, $k_{g(n)} = \lim_{n\to+\infty} k_{g(n)} := \bar{k} \in \mathbb{Z}$. Thus, for all $n \geq n_0$

$$d_{\bar{k}} \neq 0, \quad \text{and} \quad d_{\bar{k}+m_{g(n)}} = 0,$$

and as $m_{g(n)} \to +\infty$, we deduce that

$$d_{\bar{k}} \neq 0, \quad \text{and} \quad d_{\bar{k}+\ell} = 0, \quad \text{for all} \quad \ell \geq 1.$$

The process (X_t) hence admit a moving average representation of the form

$$X_t = \sum_{k=-\infty}^{\bar{k}} d_k \varepsilon_{t+k}, \qquad t \in \mathbb{Z}. \tag{C.12}$$

However, we also have by partial fraction decomposition

$$\begin{aligned}
X_t &= \frac{\Theta(F)H(B)}{\Psi(F)\Phi(B)}\varepsilon_t \\
&= \Theta(F)H(B)\frac{B^p}{B^p\Psi(F)\Phi(B)}\varepsilon_t \\
&= \Theta(F)H(B)B^p\left[\frac{b_1(B)}{B^p\Psi(F)} + \frac{b_2(B)}{\Phi(B)}\right]\varepsilon_t \\
&= \Theta(F)H(B)\left[\frac{b_1(B)}{\Psi(F)} + \frac{B^p b_2(B)}{\Phi(B)}\right]\varepsilon_t,
\end{aligned}$$

for some polynomials $b1$ and b_2 such that $0 \leq \deg(b_1) \leq p-1$, $0 \leq \deg(b_2) \leq q-1$ and $\Phi(B)b_1(B) + B^p b_2(B)\Psi(F) = 1$. We can write in general

$$\begin{aligned}
\frac{\Theta(F)H(B)b_1(B)}{\Psi(F)} &= \sum_{k=-\ell_1}^{+\infty} c_k \varepsilon_{t+k}, \\
\frac{\Theta(F)H(B)B^p b_2(B)}{\Phi(B)} &= \sum_{k=-\infty}^{\ell_2} e_k \varepsilon_{t+k},
\end{aligned}$$

for some sequences of coefficients (c_k), (e_k), and where ℓ_1 is the degree of the largest order monomial in B of $\Theta(F)H(B)b_1(B)$ (recall that $F = B^{-1}$) and ℓ_2 is the degree of the largest monomial in F of $B^p\Theta(F)H(B)b_2(B)$. By (C.12), we deduce by identification that there is some $\bar{\ell} \in \mathbb{Z}$ such that $c_k = 0$ for all $k \geq \bar{\ell} + 1$ and

$$\frac{\Theta(F)H(B)b_1(B)}{\Psi(F)} = \sum_{k=-\ell_1}^{\bar{\ell}} c_k F^k.$$

Necessarily, $\bar{\ell} \geq \ell_1$, otherwise $\Theta(F)H(B)b_1(B)\Psi^{-1}(F) = 0$ which is impossible as all the polynomials involved have non-negative degrees. Thus, we deduce that there exist two polynomials

P and Q of non-negative degrees such that

$$\frac{\Theta(z^{-1})H(z)b_1(z)}{\Psi(z^{-1})} = \sum_{k=-\ell_1}^{\bar{\ell}} c_k z^k := P(z^{-1}) + Q(z), \qquad z \in \mathbb{C}.,$$

which yields

$$\Theta(z^{-1})H(z)b_1(z) = \Psi(z^{-1})(P(z^{-1}) + Q(z)), \qquad z \in \mathbb{C}. \tag{C.13}$$

As $\deg(\Psi) = p$ and $\Psi(z) = 0$ if and only if $|z| > 1$, we know that there are p complex numbers $z_1, \ldots, z_p$ such that $0 < |z_i| < 1$ and $\Psi(z_i^{-1}) = 0$ for $i = 1, \ldots, p$. Evaluating (C.13) at the z_i's, we get that

$$\Theta(z_i^{-1})b_1(z_i) = 0, \qquad \text{for} \quad i = 1, \ldots, p,$$

because H has no roots inside the unit circle and P and Q are of finite degrees. From the fact that $\deg(b_1) \leq p - 1$, we also know that for some z_{i_0}, $b(z_{i_0}) \neq 0$ which finally yields

$$\Theta(z_{i_0}^{-1}) = 0.$$

We therefore obtain that Ψ and Θ have a common root, which is ruled out by assumption, hence the contradiction. The sequence (k_n) can thus be neither bounded nor unbounded, which is absurd. We conclude that

$$m_0 = \sup\{m \geq 1 : \ \exists k \in \mathbb{Z}, \ d_{k+m} = \ldots = d_{k+1} = 0, \ d_k \neq 0\} < +\infty.$$

Hence the equivalence between $(\iota\iota)$ and $(\iota\iota\iota)$.

Let us now show that whenever $m_0 < +\infty$, then (3.8) holds for any $m \geq m_0$.

As $m_0 < +\infty$, we have that for any $m \geq m_0$ and $h \geq 1$, $\|\boldsymbol{d}_k\| > 0$ as soon as $\boldsymbol{d}_k \neq \boldsymbol{0}$, for all $k \in \mathbb{Z}$ (recall $\boldsymbol{d}_k = (d_{k+m}, \ldots, d_k, d_{k+1}, \ldots, d_{k-h})$). For ARMA processes, the non-zero coefficients d_k of the moving average necessarily decay geometrically (times a monomial) as $k \to \pm\infty$. To fix ideas, say $d_k \underset{k\to\pm\infty}{\sim} ak^b\lambda^k$, for constants $a \neq 0$, b a non-negative integer, and $0 < |\lambda| < 1$, which may change according to whether $k \to +\infty$ or $k \to -\infty$ (if $\deg(\Phi) = 0$, then $d_{-k} = 0$ for $k \geq 0$ large enough, however, since we assume $\deg(\Psi) \geq 1$, it always holds that $|d_k| \underset{k\to+\infty}{\sim} ak^b\lambda^k$, for the non-zero terms d_k). Hence,

$$\boldsymbol{d}_k \underset{k\to\pm\infty}{\sim} ak^b\lambda^k\boldsymbol{d}_*,$$

for some constant vector $\boldsymbol{d}_*$ such that $\|\boldsymbol{d}_*\| > 0$ (which may change according to whether $k \to +\infty$ or $k \to -\infty$). We then have that

$$\frac{\|\boldsymbol{d}_k\|}{\|\boldsymbol{d}_k\|_e} \underset{k\to\pm\infty}{\longrightarrow} \frac{\|\boldsymbol{d}_*\|}{\|\boldsymbol{d}_*\|_e} > 0,$$

and

$$\|\boldsymbol{d}_k\|_e \left| \ln\left(\|\boldsymbol{d}_k\|/\|\boldsymbol{d}_k\|_e\right)\right| \underset{k\to\pm\infty}{\sim} \text{const } k^b \lambda^k.$$

Therefore, for any $m \geq m_0$, $h \geq 1$,

$$\sum_{k\in\mathbb{Z}} \|\boldsymbol{d}_k\|_e \left| \ln\left(\|\boldsymbol{d}_k\|/\|\boldsymbol{d}_k\|_e\right)\right| < +\infty$$

The equivalence between (ι) and $(\iota\iota\iota)$ is now clear: on the one hand, if $m_0 < +\infty$, then (3.8) holds for all $m \geq m_0$, $h \geq 1$, which yields the (m,h)-past-representability of $(X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$ for any $m \geq m_0$, $h \geq 1$, by Theorem 3.1. In particular, (X_t) is past-representable. On the other hand, assuming (X_t) is past-representable, then necessarily $m_0 < +\infty$.

Regarding the last statement, it follows from the above proof that the condition $m_0 < +\infty$ and $m \geq m_0$ is sufficient for (m,h)-past-representability. It is also necessary, as (3.7) never holds with $m < m_0$ (*a fortiori*, with $m < m_0 = +\infty$), concluding the proof.

C.7 Proof of Proposition 4.1

By Proposition 2.2

$$\mathbb{P}_x^{\|\cdot\|}(\boldsymbol{X_t}, A|B) \underset{x\to+\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}(A \cap B(V))}{\Gamma^{\|\cdot\|}(B(V))}.$$

The conclusion follows by considering the points of $B(V)$ and $A \cap B(V)$ that are charged by the spectral measure $\Gamma^{\|\cdot\|}$ in (4.2).

C.8 Proof of Lemma 4.1

By Proposition 3.1, we have

$$\Gamma^{\|\cdot\|} = \sum_{\vartheta\in S_1} \sum_{k\in\mathbb{Z}} \|\boldsymbol{d}_k\|^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}},$$

with $\boldsymbol{d}_k = (\rho^{k+m}\mathbb{1}_{\{k+m\geq 0\}}, \ldots, \rho^{k-h}\mathbb{1}_{\{k-h\geq 0\}})$ and $k \in \mathbb{Z}$. Thus,

$$\boldsymbol{d}_k = \begin{cases} \boldsymbol{0}, & \text{if} \quad k \leq -m-1, \\ (\rho^{k+m}, \ldots, \rho, 1, 0, \ldots, 0), & \text{if} \quad -m \leq k \leq h, \\ \rho^{k-h}\boldsymbol{d}_h, & \text{if} \quad k \geq h. \end{cases}$$

Therefore,

$$\Gamma^{\|\cdot\|} = \sum_{\vartheta\in S_1}\left[\sum_{k=-m}^{h-1}\|\boldsymbol{d}_k\|^\alpha\delta_{\left\{\frac{\vartheta\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}} + \sum_{k=h}^{+\infty}|\rho|^{\alpha(k-h)}\|\boldsymbol{d}_h\|^\alpha\delta_{\left\{\frac{\vartheta\rho^{k-h}\boldsymbol{d}_h}{|\rho|^{k-h}\|\boldsymbol{d}_h\|}\right\}}\right].$$

Moreover,

$$\begin{aligned}
&\sum_{\vartheta\in S_1}\sum_{k=h}^{+\infty}|\rho|^{\alpha(k-h)}\|\boldsymbol{d}_h\|^\alpha\delta_{\left\{\mathrm{sign}(\rho)^{k-h}\frac{\vartheta\boldsymbol{d}_h}{\|\boldsymbol{d}_h\|}\right\}}\\
&\qquad=\sum_{\vartheta\in S_1}\|\boldsymbol{d}_h\|^\alpha\frac{1}{2}\left[\sum_{k=h}^{+\infty}|\rho|^{\alpha(k-h)}+\vartheta\beta\sum_{k=h}^{+\infty}(\rho^{<\alpha>})^{k-h}\right]\delta_{\left\{\frac{\vartheta\boldsymbol{d}_h}{\|\boldsymbol{d}_h\|}\right\}}\\
&\qquad=\sum_{\vartheta\in S_1}\frac{1}{1-|\rho|^\alpha}\|\boldsymbol{d}_h\|^\alpha\bar{w}_\vartheta\delta_{\left\{\frac{\vartheta\boldsymbol{d}_h}{\|\boldsymbol{d}_h\|}\right\}}.
\end{aligned}$$

Finally, noticing that for $k=-m$ and $\boldsymbol{d}_k=(1,0,\dots,0)$,

$$\begin{aligned}
\Gamma^{\|\cdot\|} &= \sum_{\vartheta\in S_1}\left[w_\vartheta\sum_{k=-m}^{h-1}\|\boldsymbol{d}_k\|^\alpha\delta_{\left\{\frac{\vartheta\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}}+\frac{\bar{w}_\vartheta}{1-|\rho|^\alpha}\|\boldsymbol{d}_h\|^\alpha\delta_{\left\{\frac{\vartheta\,\boldsymbol{d}_h}{\|\boldsymbol{d}_h\|}\right\}}\right]\\
&= \sum_{\vartheta\in S_1}\left[w_\vartheta\left(\delta_{\{(\vartheta,0,\dots,0)\}}+\sum_{k=-m+1}^{h-1}\|\boldsymbol{d}_k\|^\alpha\delta_{\left\{\frac{\vartheta\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}}\right)+\frac{\bar{w}_\vartheta}{1-|\rho|^\alpha}\|\boldsymbol{d}_h\|^\alpha\delta_{\left\{\frac{\vartheta\,\boldsymbol{d}_h}{\|\boldsymbol{d}_h\|}\right\}}\right]\\
&= \sum_{\vartheta\in S_1}\left[w_\vartheta\delta_{\{(\vartheta,0,\dots,0)\}}+\left(w_\vartheta\sum_{k=-m+1}^{h-1}\|\boldsymbol{d}_k\|^\alpha\delta_{\left\{\frac{\vartheta\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}}+\frac{\bar{w}_\vartheta}{1-|\rho|^\alpha}\|\boldsymbol{d}_h\|^\alpha\delta_{\left\{\frac{\vartheta\,\boldsymbol{d}_h}{\|\boldsymbol{d}_h\|}\right\}}\right)\right].
\end{aligned}$$

C.9 Proof of Proposition 4.2

Lemma C.2 *Let $\Gamma^{\|\cdot\|}$ be the spectral measure given in Lemma 4.1 and assume that the ρ is positive.*

Letting $(\vartheta_0,k_0)\in\mathcal{I}$, consider

$$I_0 := \left\{\frac{\vartheta'\boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|}:\quad\frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|}=\frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}\ \ \textit{for}\ \ (\vartheta',k')\in\mathcal{I}\right\}.$$

For $m\geq 1$, and $0\leq k_0\leq h$, then

$$I_0=\left\{\frac{\vartheta_0\boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|}:\quad 0\leq k'\leq h\right\}.$$

For $m\geq 1$, and $-m\leq k_0\leq -1$, then

$$I_0=\begin{cases}\left\{\dfrac{\vartheta_0\boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}, & \textit{if}\quad -m+1\leq k_0\leq -1\\[2ex] \left\{\dfrac{\vartheta_0\boldsymbol{d}_{0,k_0}}{\|\boldsymbol{d}_{0,k_0}\|}\right\}=\{(\vartheta_0,0,\dots,0)\}, & \textit{if}\quad k_0=-m.\end{cases}$$

For $m=0$, then

$$I_0=\left\{\frac{\vartheta_0\boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|}:\quad k'\in\{1,\dots,h\}\cup\{(0,0)\}\right\}.$$

Proof.

Case $m \geq 1$ and $k_0 \in \{0, \ldots, h\}$

If $k' \in \{-m, \ldots, -1\}$, the $(m+1)^{\text{th}}$ component of $f(\boldsymbol{d}_{k'})$ is zero, whereas the $(m+1)^{\text{th}}$ component of $f(\boldsymbol{d}_{k_0})$ is $\rho^{k_0} \neq 0$. Necessarily, $\vartheta' f(\boldsymbol{d}_{k'})/\|\boldsymbol{d}_{k'}\| \neq \vartheta_0 f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\|$ and

$$I_0 = \left\{ \frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|} \text{ for } (\vartheta', k') \in \{-1, +1\} \times \{0, \ldots, h\} \right\}.$$

Now, with $k' \in \{0, \ldots, h\}$, we have that

$$\begin{aligned} f(\boldsymbol{d}_{k'}) &= (\rho^{k'+m}, \ldots, \rho^{k'+1}, \rho^{k'}), \\ f(\boldsymbol{d}_{k_0}) &= (\rho^{k_0+m}, \ldots, \rho^{k_0+1}, \rho^{k_0}), \end{aligned}$$

and by (3.1) we also have that

$$\begin{aligned} \|\boldsymbol{d}_{k'}\| &= \|(\rho^{k'+m}, \ldots, \rho^{k'+1}, \rho^{k'}, \overbrace{0, \ldots, 0}^{h})\|, \\ \|\boldsymbol{d}_{k_0}\| &= \|(\rho^{k_0+m}, \ldots, \rho^{k_0+1}, \rho^{k_0}, \underbrace{0, \ldots, 0}_{h})\|. \end{aligned}$$

Thus,

$$\begin{aligned} \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|} & \\ \iff \quad & \frac{\vartheta' \rho^{k'} f(\boldsymbol{d}_0)}{|\rho|^{k'} \|\boldsymbol{d}_0\|} = \frac{\vartheta_0 \rho^{k_0} f(\boldsymbol{d}_0)}{|\rho|^{k_0} \|\boldsymbol{d}_0\|} \\ \iff \quad & \frac{\vartheta' \rho^{\ell}}{\|\boldsymbol{d}_0\|} = \frac{\vartheta_0 \rho^{\ell}}{\|\boldsymbol{d}_0\|}, \qquad \ell = 0, \ldots, m \\ \iff \quad & \vartheta' \vartheta_0 \frac{\|\boldsymbol{d}_0\|}{\|\boldsymbol{d}_0\|} = \left(\frac{\rho}{\rho}\right)^{\ell}, \quad \ell = 0, \ldots, m \\ \iff \quad & \vartheta' \vartheta_0 = 1 \\ \iff \quad & \vartheta' = \vartheta_0, \end{aligned}$$

because $\rho \neq 0$ is assumed.

Case $m \geq 1$ and $k_0 \in \{-m, \ldots, -1\}$

By comparing the place of the first zero component, it is easy to see that

$$\frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|} \Longrightarrow k' = k_0.$$

$$\begin{aligned} f(\boldsymbol{d}_{k'}) &= (\overbrace{\rho^{k'+m}, \ldots, \rho, 1, 0, \ldots, 0}^{m+1}, \overbrace{0, \ldots, 0}^{h}), \\ f(\boldsymbol{d}_{k_0}) &= (\underbrace{\rho^{k_0+m}, \ldots, \rho, 1, 0, \ldots, 0}_{m+1}, \underbrace{0, \ldots, 0}_{h}), \end{aligned}$$

and we also have that

$$\|\boldsymbol{d}_{k'}\| = \|(\overbrace{\rho^{k'+m},\dots,\rho,1,0,\dots,0}^{m+1},\overbrace{0,\dots,0}^{h})\|,$$
$$\|\boldsymbol{d}_{k_0}\| = \|(\underbrace{\rho^{k_0+m},\dots,\rho,1,0,\dots,0}_{m+1},\underbrace{0,\dots,0}_{h})\|.$$

As $k' = k_0 \le -1$,

$$\begin{aligned}\frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|} \\ \iff \frac{\vartheta'\rho^\ell}{\|\boldsymbol{d}_{k_0}\|} = \frac{\vartheta_0\rho^\ell}{\|\boldsymbol{d}_{k_0}\|}, \quad \ell = 0,\dots,m+k_0, \text{ and } k' = k_0 \\ \iff \vartheta'\vartheta_0\frac{\|\boldsymbol{d}_{k_0}\|}{\|\boldsymbol{d}_{k_0}\|} = \left(\frac{\rho}{\rho}\right)^\ell, \quad \ell = 0,\dots,m+k_0, \text{ and } k' = k_0.\end{aligned}$$

Now if $-m+1 \le k_0 \le -1$,

$$\begin{aligned}\vartheta'\vartheta_0\frac{\|\boldsymbol{d}_{k_0}\|}{\|\boldsymbol{d}_{k_0}\|} = \left(\frac{\rho}{\rho}\right)^\ell, \quad \ell = 0,1,\dots,m+k_0, \text{ and } k' = k_0 \\ \iff \vartheta' = \vartheta_0 \text{ and } k' = k_0.\end{aligned}$$

If $k_0 = -m$, given that $(\vartheta_0, k_0) \in \mathcal{I} = S_1 \times \Big(\{-m,\dots,-1,0,1,\dots,h\} \cup \{(0,-m)\}\Big)$, and as $k' = k_0 = -m$, we have that $\boldsymbol{d}_{k_0} = \boldsymbol{d}_{0,-m} = (1,0,\dots,0)$. Hence

$$\begin{aligned}\vartheta'\vartheta_0\frac{\|\boldsymbol{d}_{k_0}\|}{\|\boldsymbol{d}_{k_0}\|} = \left(\frac{\rho}{\rho}\right)^\ell, \quad \ell = 0, \text{ and } k' = k_0 = -m \ , \\ \iff \vartheta' = \vartheta_0 \text{ and } k' = k_0 = -m\end{aligned}$$

Case $m = 0$

If $k_0 \in \{1,\dots,h\}$ then $f(\boldsymbol{d}_{k_0}) = \rho^{k_0}$ and by (3.1), $\|\boldsymbol{d}_{k_0}\| = |\rho|^{k_0}$. Thus, $\vartheta_0 f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\| = \vartheta_0$. If $k_0 = -m = 0$, then $f(\boldsymbol{d}_{k_0}) = 1$ and $\vartheta_0 f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\| = \vartheta_0$. The same holds for $(\vartheta', k') \in \mathcal{I}$ and we obtain that

$$\frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|} \iff \vartheta' = \vartheta_0.$$

□

Let us now prove Proposition 4.2. By Proposition 4.1,

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A_{\vartheta,k}\Big|B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta'\boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \ \frac{\vartheta' f(\boldsymbol{d}_{,k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right)}{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta'\boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \ \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right)}. \tag{C.14}$$

Focusing on the denominator, we have by (4.6)

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta'\boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \ \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right) = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta'\boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \ \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}\right\}\right)$$

We will now distinguish the cases arising from the application of Lemma C.2. Recall that we assume for this proposition that the ρ is positive. Thus, $\text{sign}(\rho) = 1$ and $\bar{\beta} = \beta \dfrac{1-|\rho|^\alpha}{1-\rho^{<\alpha>}} = \beta$ and $\bar{w}_\vartheta = w_\vartheta$ in (4.5) for $\vartheta \in \{-1, +1\}$.

Case $m \geq 1$ and $0 \leq k_0 \leq h$

By Lemma C.2,

$$\begin{aligned}\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) \\ = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} : \quad 0 \leq k' \leq h\right\}\right) \\ = \left[w_{\vartheta_0} \sum_{k'=0}^{h-1} \|\boldsymbol{d}_{k'}\|^\alpha + \frac{\bar{w}_{\vartheta_0}}{1-|\rho|^\alpha}\|\boldsymbol{d}_h\|^\alpha\right]\end{aligned}$$

By (3.1), for $k' \in \{0, 1, \ldots, h\}$

$$\begin{aligned}\|\boldsymbol{d}_{k'}\| &= \|(\rho^{k'+m}, \ldots, \rho^{k'+1}, \rho^{k'}, \underbrace{0, \ldots, 0}_{h})\| \\ &= |\rho|^{k'-h}\|(\rho^{m+h}, \ldots, \rho^{h+1}, \rho^h, \underbrace{0, \ldots, 0}_{h})\| \\ &= |\rho|^{k'-h}\|\boldsymbol{d}_h\|.\end{aligned}$$

Thus,

$$\begin{aligned}\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) &= w_{\vartheta_0}\|\boldsymbol{d}_h\|^\alpha\left[\sum_{k'=0}^{h-1}\rho^{\alpha(k'-h)} + \frac{1}{1-|\rho|^\alpha}\right] \\ &= w_{\vartheta_0}\|\boldsymbol{d}_h\|^\alpha \frac{|\rho|^{-\alpha h}}{1-|\rho|^\alpha}.\end{aligned}$$

Similarly for the numerator in (C.14), by (4.7),

$$\begin{aligned}&\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right) \\ &\qquad = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \quad 0 \leq k' \leq h\right\}\right)\end{aligned}$$

$$= \begin{cases} \Gamma^{\|\cdot\|}\left(\left\{\dfrac{\vartheta_0 \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}\right), & \text{if} \quad \vartheta = \vartheta_0, \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if} \quad \vartheta \neq \vartheta_0, \end{cases}$$

$$= \begin{cases} w_{\vartheta_0}\|\boldsymbol{d}_h\|^\alpha |\rho|^{\alpha(k-h)}\delta_{\{\vartheta_0\}}(\vartheta), & \text{if} \quad 0 \leq k \leq h-1, \\ w_{\vartheta_0}\|\boldsymbol{d}_h\|^\alpha \dfrac{1}{1-|\rho|^\alpha}\delta_{\{\vartheta_0\}}(\vartheta), & \text{if} \quad k = h. \end{cases}$$

The conclusion follows.

Case $m \geq 1$ and $-m \leq k_0 \leq -1$

We have by Lemma C.2

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C_{m+h+1}^{\|\cdot\|} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right).$$

If $-m+1 \leq k_0 \leq -1$,

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) = w_{\vartheta_0}\|\boldsymbol{d}_{k_0}\|^{\alpha},$$

and

$$\begin{aligned}
&\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right) \\
&\qquad = \Gamma^{\|\cdot\|}\left(A_{\vartheta,k} \cap \left\{\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) \\
&\qquad = \begin{cases} \Gamma^{\|\cdot\|}\left(\left\{\dfrac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right), & \text{if } \vartheta = \vartheta_0, \text{ and } k = k_0, \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if } \vartheta \neq \vartheta_0 \text{ or } k \neq k_0, \end{cases} \\
&\qquad = w_{\vartheta_0}\|\boldsymbol{d}_{k_0}\|^{\alpha}\delta_{\{\vartheta_0\}}(\vartheta)\delta_{\{k_0\}}(k).
\end{aligned}$$

If $k_0 = -m$, then $\boldsymbol{d}_{k_0} = \boldsymbol{d}_{0,-m} = (1, 0, \ldots, 0)$, and

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) = \Gamma^{\|\cdot\|}\Big(\{\vartheta_0(1, 0, \ldots, 0)\}\Big) = w_{\vartheta_0},$$

and

$$\begin{aligned}
&\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right) \\
&\qquad = \Gamma^{\|\cdot\|}\left(A_{\vartheta,k} \cap \left\{\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) \\
&\qquad = \begin{cases} \Gamma^{\|\cdot\|}\Big(A_{\vartheta,k} \cap \{\vartheta_0(1, 0, \ldots, 0)\}\Big), & \text{if } \vartheta = \vartheta_0, \text{ and } k = k_0 = -m, \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if } \vartheta \neq \vartheta_0 \text{ or } k \neq k_0, \end{cases} \\
&\qquad = w_{\vartheta_0}\delta_{\{\vartheta_0\}}(\vartheta)\delta_{\{k_0\}}(k).
\end{aligned}$$

The conclusion follows as previously.

Case $m = 0$

By Lemma C.2, as the ρ is positive

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \ \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}\right\}\right)$$
$$= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \ k' \in \{0, \dots, h\} \cup \{(0,0)\}\right\}\right)$$

Given that $\|\boldsymbol{d}_{k'}\| = |\rho|^{k'}$, for any $1 \le k' \le h$,

$$\begin{aligned}
\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \ \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|}\right\}\right) &= w_{\vartheta_0} + w_{\vartheta_0}\left[\sum_{k'=1}^{h-1} \|\boldsymbol{d}_{k'}\|^{\alpha} + \frac{\|\boldsymbol{d}_h\|^{\alpha}}{1-|\rho|^{\alpha}}\right] \\
&= w_{\vartheta_0}\left[1 + \sum_{k'=1}^{h-1} |\rho|^{\alpha k'} + \frac{|\rho|^{\alpha h}}{1-|\rho|^{\alpha}}\right] \\
&= w_{\vartheta_0}\left[\frac{1-|\rho|^{\alpha h}}{1-|\rho|^{\alpha}} + \frac{|\rho|^{\alpha h}}{1-|\rho|^{\alpha}}\right] \\
&= w_{\vartheta_0}\frac{1}{1-|\rho|^{\alpha}}.
\end{aligned}$$

Similarly, by (4.7),

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \ \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right)$$
$$= \Gamma^{\|\cdot\|}\left(A_{\vartheta,k} \cap \left\{\frac{\vartheta_0 \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \ k' \in \{0, \dots, h\} \cup \{(0,0)\}\right\}\right)$$

$$= \begin{cases} \Gamma^{\|\cdot\|}\left(\left\{\dfrac{\vartheta_0 \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}\right), & \text{if } \vartheta = \vartheta_0, \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if } \vartheta \neq \vartheta_0, \end{cases}$$

$$= \begin{cases} w_{j\vartheta_0}\delta_{\{\vartheta_0\}}(\vartheta), & \text{if } \ k = 0, \\ w_{\vartheta_0}|\rho|^{\alpha k}\delta_{\{\vartheta_0\}}(\vartheta), & \text{if } \ 1 \le k \le h-1, \\ w_{\vartheta_0}\dfrac{|\rho|^{\alpha h}}{1-|\rho|^{\alpha}}\delta_{\{\vartheta_0\}}(\vartheta), & \text{if } \ k = h. \end{cases}$$

The conclusion follows.

C.10 Proof of Proposition 4.3

Lemma C.3 *Let X_t be the α-stable anticipative AR(2) (resp. fractionally integrated AR) as in (4.8) (resp. (4.10)). With f as in (4.1), and for any $m \geq 1$, $h \geq 0$,*

$$\forall k, \ell \geq -m, \ \ \forall \vartheta_1, \vartheta_2 \in S_1, \quad \left[\frac{f(\vartheta_1 \boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} = \frac{f(\vartheta_2 \boldsymbol{d}_\ell)}{\|\boldsymbol{d}_\ell\|} \implies k = \ell \ \text{ and } \ \vartheta_1 = \vartheta_2 \right].$$

Proof.

The result is clear for both processes for $-m \leq k, \ell \leq -1$. For $k, \ell \geq 0$,

$$\begin{aligned} \frac{f(\vartheta_1 \boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} = \frac{f(\vartheta_2 \boldsymbol{d}_\ell)}{\|\boldsymbol{d}_\ell\|} &\iff \left[\forall\, i = 0, \dots, m, \quad \frac{\vartheta_1 d_{k+i}}{\|\boldsymbol{d}_k\|} = \frac{\vartheta_2 d_{\ell+i}}{\|\boldsymbol{d}_\ell\|} \right] \\ &\iff \frac{d_k}{d_\ell} = \frac{d_{k+1}}{d_{\ell+1}} = \dots = \vartheta_1 \vartheta_2 \frac{\|\boldsymbol{d}_k\|}{\|\boldsymbol{d}_\ell\|}. \end{aligned} \tag{C.15}$$

The last statement in particular implies that $\dfrac{d_k}{d_\ell} = \dfrac{d_{k+1}}{d_{\ell+1}}$.

For the anticipative AR(2), if $\lambda_1 \neq \lambda_2$, we then have

$$\begin{aligned} \frac{d_k}{d_\ell} = \frac{d_{k+1}}{d_{\ell+1}} &\iff \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1^{\ell+1} - \lambda_2^{\ell+1}} = \frac{\lambda_1^{k+2} - \lambda_2^{k+2}}{\lambda_1^{\ell+2} - \lambda_2^{\ell+2}} \\ &\iff \lambda_1^{k-\ell} = \lambda_2^{k-\ell} \\ &\iff k = \ell. \end{aligned}$$

This case $\lambda_1 = \lambda_2 = \lambda$ is similar. For the anticipative fractionally integrated AR, given that $\Gamma(z+1) = z\Gamma(z)$ for any $z \in \mathbb{C}$, we have

$$\begin{aligned} \frac{d_k}{d_\ell} = \frac{d_{k+1}}{d_{\ell+1}} &\iff \frac{\Gamma(k+d)\Gamma(\ell+1)}{\Gamma(\ell+d)\Gamma(k+1)} = \frac{\Gamma(k+d+1)\Gamma(\ell+2)}{\Gamma(\ell+d+1)\Gamma(k+2)} \\ &\iff \frac{\Gamma(\ell+d+1)\Gamma(k+2)}{\Gamma(\ell+d)\Gamma(k+1)} = \frac{\Gamma(k+d+1)\Gamma(\ell+2)}{\Gamma(k+d)\Gamma(\ell+1)} \\ &\iff (k-\ell)(d-1) = 0 \\ &\iff k = \ell. \end{aligned}$$

Therefore, in all cases,

$$\frac{d_k}{d_\ell} = \frac{d_{k+1}}{d_{\ell+1}} = \dots = \vartheta_1 \vartheta_2 \frac{\|\boldsymbol{d}_k\|}{\|\boldsymbol{d}_\ell\|} \implies k = \ell \ \text{ and } \ \vartheta_1 \vartheta_2 = 1.$$

□

Let us now prove Proposition 4.3. The spectral measure of $\boldsymbol{X}_t$ writes

$$\Gamma^{\|\cdot\|} = \sigma^\alpha \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_\vartheta \|\boldsymbol{d}_k\|^\alpha \delta_{\left\{ \frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \right\}},$$

where the sequences (d_k) are given respectively by (4.9) and (4.11) for the anticipative AR(2) and fractionally integrated processes. By Proposition 2.2,

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A\Big|B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}(A\cap B(V_0))}{\Gamma^{\|\cdot\|}(B(V_0))}.$$

On the one hand, we have by definition of $B(V_0)$, V_0 and Lemma C.3,

$$\begin{aligned}
\Gamma^{\|\cdot\|}(B(V_0)) &= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \in B(V_0) :\ (\vartheta,k)\in\{-1,+1\}\times\mathbb{Z}\right\}\right)\\
&= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \in C_{m+h+1}^{\|\cdot\|} :\ \frac{\vartheta f(\boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} \in V_0,\ (\vartheta,k)\in\{-1,+1\}\times\mathbb{Z}\right\}\right)\\
&= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} \in C_{m+h+1}^{\|\cdot\|} :\ \frac{\vartheta f(\boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{k_0})}{\|\boldsymbol{d}_{k_0}\|},\ (\vartheta,k)\in\{-1,+1\}\times\mathbb{Z}\right\}\right)\\
&= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right).
\end{aligned}$$

Similarly, it is easily shown that

$$\Gamma^{\|\cdot\|}(A\cap B(V_0)) = \Gamma^{\|\cdot\|}\left(A\cap\left\{\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|}\right\}\right).$$

The conclusion follows.

WORKING PAPER

PREVIOUS ISSUES

Working Paper Publication Directors :

Marc Baudry, Philippe Delacote, Olivier Massol

The Climate Economics Chair is a joint initiative by Paris-Dauphine University, CDC, TOTAL and EDF, under the aegis of the European Institute of Finance.