# Forecasting extreme trajectories using semi-norm representations

Gilles de Truchis*, Sébastien Fries†, Arthur Thomas‡

PRELIMINARY DRAFT - PLEASE DO NOT CIRCULATE

May 31, 2024


### Abstract

For $(X_t)$ a two-sided $\alpha$-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 $\boldsymbol{X_t} = (X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$, $m \geq 0$, $h \geq 1$, are multivariate $\alpha$-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 sets $\{\boldsymbol{s} \in \mathbb{R}^{m+h+1} : \quad \|\boldsymbol{s}\| = 1\}$ for $\|\cdot\|$ an adequate semi-norm is proposed to describe the tail behaviour of vectors $\boldsymbol{X_t}$ 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 $(X_t)$ will have to be "anticipative enough" for $\boldsymbol{X_t}$ 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.


*Keywords:* Prediction, Stable random vectors, Spectral representation, Pattern identification

*Université d'Orléans, Laboratoire d'Economie d'Orléans (LEO), 45000 PARIS, FRANCE, gilles.detruchis@univ-orleans.fr.

†Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

‡Université Paris Dauphine, Université PSL, LEDA, CNRS, IRD, 75016 PARIS, FRANCE, arthur.thomas@dauphine.psl.eu. We are grateful to the Agence Nationale de la Recherche (ANR), which supported this work via Project MultiRisk (ANR CE26 2016 - CR)

# 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. This gain of interest is driven in particular by their convenience for modelling exotic patterns in time series, such as explosive bubbles in financial prices (Cavaliere et al., 2017; Cubadda et al., 2019; Fries and Zakoian, 2019; Fries, 2021; Gourieroux et al., 2020; Gouriéroux and Zakoian, 2017; Hecq et al., 2016, 2017a,b, 2020; Hecq and Voisin, 2021; Hencic and Gouriéroux, 2015).[1] The attractive flexibility of anticipative processes cannot yet be fully leveraged, however, as their dynamics, and especially the conditional distribution of future paths given the observed past trajectory, remain largely mysterious. A remarkable exception is that of the anticipative $\alpha$-stable AR(1) for which partial results were obtained in Gouriéroux and Zakoian (2017) and further completed in Fries (2021). Even in this simplest case within the family of anticipative processes, however, future realisations feature a complex dependence on the observed past, which is reflected in the functional forms of the conditional moments obtained in Fries (2021). Interestingly, the dynamics of the anticipative stable AR(1) simplifies during extreme events where it appears to follow an explosive exponential path with a determined killing probability. This naturally raises the question of whether and under which form such behaviour could be found in more general stable processes.

For $X_t = \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k}$ a two-sided moving average with $(\varepsilon_t)$ an i.i.d. $\alpha$-stable sequence and $(d_k)$ a non-random coefficients sequence, this paper analyses the conditional distribution of future paths given the observed trajectory, say $(X_{t+1}, \dots, X_{t+h})$ given $(X_{t-m}, \dots, X_t)$, $m \geq 0$, $h \geq 1$, when the process is far from its central values. Only mild summability conditions are assumed on the sequence $(d_k)$ and, in particular, we do

[1] see also Andrews et al. (2009); Behme (2011); Behme et al. (2011); Chen et al. (2017); Gouriéroux and Jasiak (2016, 2017); Lanne and Saikkonen (2011, 2013); Saikkonen and Sandberg (2016)

not presume anything upfront on the anticipativeness or non-anticipativeness of $(X_t)$.[2] Under this framework, any vector of the form $\boldsymbol{X_t} = (X_{t-m}, \ldots, X_{t+h})$ is multivariate $\alpha$-stable and its distribution is characterised by a unique finite measure $\Gamma$ 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 $\Gamma$ in particular completely describes the conditional distribution of the normalised paths $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|_e$, the "shape" of the trajectory, when $\boldsymbol{X_t}$ is large according to the Euclidean norm and given some information about the observed first $m+1$ components. A straightforward application of Theorem 4.4.8 by Samorodnitsky and Taqqu (1994) indeed shows that

$$\mathbb{P}\Big(\boldsymbol{X}_t/\|\boldsymbol{X}_t\|_e \in A \ \Big| \ \|\boldsymbol{X}_t\|_e > x \ \text{ and } \ \boldsymbol{X}_t/\|\boldsymbol{X}_t\|_e \in B\Big) \underset{x\longrightarrow\infty}{\longrightarrow} \frac{\Gamma(A\cap B)}{\Gamma(B)}, \tag{1.1}$$

for any appropriately chosen Borel sets $A, B \subset S_{m+h+1}$. As such however, (1.1) is of little value for prediction purposes where only $X_{t-m}, \ldots, X_t$ are assumed to be observed, given that the conditioning generally depends on the future realisations $X_{t+1}, \ldots, X_{t+h}$, mainly through the Euclidean norm of $\boldsymbol{X}_t$. The idea developed here is to obtain a version of (1.1) where the Euclidean norm is replaced by a semi-norm $\|\cdot\|$ satisfying

$$\|(x_{-m}, \ldots, x_0, x_1, \ldots, x_h)\| = \|(x_{-m}, \ldots, x_0, 0, \ldots, 0)\|, \tag{1.2}$$

for any $(x_{-m}, \ldots, x_h) \in \mathbb{R}^{m+h+1}$. In this view, a new representation of stable random vectors on the "unit cylinder" $C^{\|\cdot\|}_{m+h+1} := \{\boldsymbol{s} \in \mathbb{R}^{m+h+1} : \ \|\boldsymbol{s}\| = 1\}$ is thus explored, where $\|\cdot\|$ is such a semi-norm. Contrary to representations involving norms (see Theorem 2.3.8 in Samorodnitsky and Taqqu (1994)), not all stable random vectors admit representations on unit cylinders and a characterisation is provided. It is shown that only if $(X_t)$ is "anticipative enough" will $\boldsymbol{X_t}$ admit a representation by a measure $\Gamma^{\|\cdot\|}$ on $C^{\|\cdot\|}_{m+h+1}$. The property (1.1) is then shown to hold with an adequate semi-norm and with $\Gamma$ (resp. $S_{m+h+1}$) replaced by $\Gamma^{\|\cdot\|}$ (resp. $C^{\|\cdot\|}_{m+h+1}$). The problem finally boils down to choosing the

[2] That is, we do not presume anything on the zeros of $(d_k)$, e.g., $d_k = 0$ for $k > 0$ (purely non-anticipative case) or $k < 0$ (purely anticipative case).

appropriate Borels $B$ in (1.1) reflecting that only the past "shape" $(X_{t-m},\dots,X_t)/\|\boldsymbol{X_t}\|$ is observed.

The use of (1.1) to infer about the future paths of $(X_t)$ has connections with 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 of a multivariate regularly varying time series conditionnally on the first observation being large. The approach followed here differs in that it operates at the representation level of $\alpha$-stable vectors, establishing a link between the spectral representation and the tail conditional distribution of stable linear processes and shedding light on the (un)predictability of their extremes. A natural interpretation of path prediction in terms of pattern identification emerges from Property (1.1) applied to stable linear processes, similar to what Janßen (2019) pointed out in a framework close to that of Basrak and Segers (2009).

Section 2 characterises the representation of general $\alpha$-stable vectors on semi-norm unit cylinders and shows that Property (1.1) can be restated under this new representation. Focusing on $\alpha$-stable moving averages Section 3 studies under which condition on the process $(X_t)$ the vector $(X_{t-m},\dots,X_{t+h})$ admits a representation on the unit cylinder $C^{\|\cdot\|}_{m+h+1}$. The anticipativeness of $(X_t)$ surprisingly arises as a necessary condition for such a representation to exist. Section 4 then exploits Property (1.1) to analyse the tail conditional distribution of some particular processes: the anticipative AR(1), AR(2) and the anticipative fractionally integrated process. Section 5, provides a set of Monte-Carlo simulations that illustrate how, in practice, this pattern identification can be used to predict very accurately the future path. In particular, we suggest simple procedures to forecast crash probabilities and forecast crash dates. In Section 6 we demonstrate the empirical relevance of our theoretical results in a climate forecasting exercise. More precisely, we predict El Niño and La Niña occurrences and their reversal date using the Southern Oscillation Index

(SOI) data.[3] To replicate the numerical and empirical results of the paper and to illustrate the generality of our approach, we develop a web application that allows to replicate our results, find examples of simulated trajectories, and apply our procedures to other times series, mainly macroeconomic, financial, and climate time series. Finally, one can upload any type of time series that exhibits extreme explosive trajectories and use our forecasting approach to predict the crash probability or reversal date.[4] Section 7 concludes. Proofs are collected in Section 8.

## 2 Stable random vectors representation on unit cylinders

This section starts by recalling the characterisation of stable random vectors on the Euclidean unit sphere before exploring the case of unit cylinders relative to semi-norms and reformulating the regularly varying tails property.

**Definition 2.1** *A random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ is said 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 non-random vector $\boldsymbol{D} \in \mathbb{R}^d$ such that*

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

*where $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ are independent copies of $\boldsymbol{X}$. Moreover, if $\boldsymbol{X}$ is stable, then there exists a constant $\alpha \in (0, 2]$ such that the above holds with $C = (A^\alpha + B^\alpha)^{1/\alpha}$, and $\boldsymbol{X}$ is then called $\alpha$-stable.*

The Gaussian case ($\alpha = 2$) is henceforth excluded. For $0 < \alpha < 2$, the vector $\boldsymbol{X} = (X_1, \dots, X_d)$ is an $\alpha$-stable random vector if and only if there exists a unique pair $(\Gamma, \boldsymbol{\mu}^0)$,

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

[4] The web application is under development and will be released soon

$\Gamma$ a finite measure on $S_d$ and $\boldsymbol{\mu}^0$ a non-random vector in $\mathbb{R}^d$, such that,

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

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 $\boldsymbol{X}$, $\Gamma$ 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

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

In the univariate case, (2.1) boils down to

$$\mathbb{E}\Big[e^{iuX}\Big] = \exp\left\{ -\sigma^\alpha|u|^\alpha\Big(1 - i\beta\,\mathrm{sign}(u)w(\alpha,u)\Big) + iu\mu\right\}, \quad \forall u\in\mathbb{R},$$

for some $\sigma > 0$, $\beta\in[-1,1]$ and $\mu\in\mathbb{R}$. The representations (2.1) and (2.2) of a stable random vector involves integration over all directions of $\mathbb{R}^d$,[5] here parameterised by the unit sphere relative to the Euclidean norm. Proposition 2.3.8 in Samorodnitsky and Taqqu (1994) shows that the unit sphere relative to any norm can be used instead, provided a change of spectral measure and shift vector. We study alternative representations where integration is performed over a unit cylinder relative to a semi-norm. For a given semi-norm, not all stable vectors admit such a representation, which motivates the following definition.

**Definition 2.2** 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 $\alpha$-stable random vector.

[5] By *direction* of $\mathbb{R}^d$, it is meant 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 $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^d$.

*(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 $\boldsymbol{\mu}_{\|\cdot\|}^0 \in \mathbb{R}^d$ 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 $\alpha = 1$,

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

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 $(C_d^{\|\cdot\|}, \Gamma^{\|\cdot\|}, \boldsymbol{\mu}_{\|\cdot\|}^0)$.
*(Symmetric case)* In the case where $\boldsymbol{X}$ is symmetric $\alpha$-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 $(S_d, \Gamma)$ replaced by $(C_d^{\|\cdot\|}, \Gamma^{\|\cdot\|})$.

**Remark 2.1** As unit cylinders are unbounded sets, the integrability conditions (2.3)-(2.4) ensure the sanity of the above definition.

We start by characterising stable random vectors that are representable on a given semi-norm unit cylinder.

**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 $\alpha$-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*

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

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

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

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

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

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

**Remark 2.2** The representability condition in the case $\alpha = 1$ and $\boldsymbol{X}$ not symmetric, is slightly stronger than that in the other cases. Indeed, $\int_{K^{\|\cdot\|}} \Big|\ln \|\boldsymbol{s}\|\Big| \Gamma(d\boldsymbol{s}) \leq \int_{S_d} \Big|\ln \|\boldsymbol{s}\|\Big| \Gamma(d\boldsymbol{s}) < +\infty$ necessarily implies that $\Gamma(K^{\|\cdot\|}) = 0$ since $\Big|\ln \|\boldsymbol{s}\|\Big| = +\infty$ for $\boldsymbol{s} \in K^{\|\cdot\|}$.

**Remark 2.3** The case $d = 2$ is insightful. In view of (1.1), the spectral measure of the $\alpha$-stable vector $(X_1, X_2)$ describes its likelihood of being in any particular direction of $\mathbb{R}^2$ when it is large in norm. As unit spheres relative to norms span all the directions of $\mathbb{R}^2$, spectral measures on such spheres can describe any potential tail dependence of $(X_1, X_2)$. Unit cylinders however do not span all directions of $\mathbb{R}^2$ and spectral measures thereon necessarily 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 semi-norm such that $\|(x_1, x_2)\| = |x_1|$ for all $(x_1, x_2) \in \mathbb{R}^2$. It is easy to see that $C_2^{\|\cdot\|}$ spans all directions of $\mathbb{R}^2$ but the ones of $(0, -1)$ and $(0, +1)$. A stable vector $(X_1, X_2)$ will admit a representation on $C_2^{\|\cdot\|}$ provided these directions are irrelevant to characterise its distribution, that is, if $\Gamma\Big(\{(0, -1), (0, +1)\}\Big) = 0$. In terms of tail dependence, the latter condition intuitively means that realisations $(X_1, X_2)$ where $X_2$ is extreme and $X_1$ is not almost never occur (i.e., occur with probability zero).[6]

[6] The conditions $\Gamma\big(\{(0,-1),(0,+1)\}\big) = 0$ and $\int_{S_2} \Big|\ln \|\boldsymbol{s}\|\Big| \Gamma(d\boldsymbol{s}) < +\infty$ can also be related to the stronger condition ensuring the existence of conditional moments of $X_2$ given $X_1$ obtained in Cioczek-

Provided the adequate representation exists, Property (1.1) then holds with semi-norms instead of norms, providing the cornerstone for studying the tail conditional distribution of stable processes.

**Proposition 2.2** *Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ be an $\alpha$-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 $\partial B$ (resp. $\partial(A\cap B)$) denotes the boundary of $B$ (resp. $A\cap B$), and*

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

# 3 Unit cylinder representation for paths of stable linear processes

Given a semi-norm, 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 $(X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$ admit such representations. A characterisation is proposed and is then extended to linear combination of stable moving averages. Any semi-norm satisfying (1.2) could be relevant for the prediction framework mentioned in introduction. However to fix ideas and avoid numerous cases with respect to all the possible kernels, we restrict to semi-norms such that

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

for any $(x_{-m}, \ldots, x_h) \in \mathbb{R}^{m+h+1}$, which in particular satisfy (1.2).

Georges and Taqqu (1994, 1998) (see also Theorem 5.1.3 in Samorodnitsky and Taqqu (1994)) and which requires $\Gamma$ not to be too concentrated around the points $(0, \pm 1)$. Namely, 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 $\gamma < \min(\alpha+\nu, 2\alpha+1)$, despite the fact that $\mathbb{E}[|X_2|^\alpha] = +\infty$. If the previous holds for some $\nu > 0$, then necessarily both of the aforementioned conditions are satisfied.

**Example 3.1** Semi-norms on $\mathbb{R}^{m+h+1}$ satisfying (3.1) can be naturally obtained from norms on the $m+1$ first components of vectors. For any $p \in [1,+\infty]$, one can consider for instance semi-norms $\|\cdot\|$ 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 $p=+\infty$.

### 3.1 The case of moving averages

Consider $(X_t)$ the $\alpha$-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 $m \geq 0$, $h \geq 1$,

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

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. The random vector $\boldsymbol{X_t}$ is multivariate $\alpha$-stable: denoting $\boldsymbol{d}_k := (d_{k+m},\dots,d_k,d_{k-1},\dots,d_{k-h})$ for $k \in \mathbb{Z}$, the spectral representation of $\boldsymbol{X_t}$ on the Euclidean sphere reads $(\Gamma, \boldsymbol{\mu}^0)$ 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 $w_\vartheta = (1+\vartheta\beta)/2$, $S_1 = \{-1,+1\}$, $\delta$ 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 $k$th 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 $\Gamma$ and the random vector $\boldsymbol{X}_t$ are symmetric. The next result characterises the representability of $\boldsymbol{X_t}$ on a unit cylinder for fixed $m$ and $h$.

**Lemma 3.1** *Let $\boldsymbol{X_t}$ satisfy (3.2)-(3.5) and let $\|\cdot\|$ be a semi-norm 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 $C_{m+h+1}^{\|\cdot\|}$ if and only if*

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

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

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

In the cases $\alpha\neq 1$ and $(\alpha,\beta)=(1,0)$, the representability of $\boldsymbol{X_t}$ on a semi-norm unit cylinder depends on the number of observation $m+1$ but not on the prediction horizon $h$. Moreover, it is easy to see that if (3.7) is true for some $m\geq 0$, it then holds for any $m'\geq m$. The case $\alpha=1$, $\beta\neq 0$ is more intricate, the roles of $m$ and $h$ in the validity of the additional requirement (3.8) not being as clear-cut.

A key distinction appears between moving averages according to whether finite length paths admit semi-norm representations. This distinction especially matters for the applicability of Proposition 2.5 when studying the conditional dynamics of a given process. The following definition thus introduces the notion of *past-representability* of a stable moving average.

**Definition 3.1** Let $(X_t)$ be an $\alpha$-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 $C_{m+h+1}^{\|\cdot\|}$ for some semi-norm satisfying (3.1). For any such pair $(m, h)$, we will say that $(X_t)$ is $(m, h)$-past-representable.

**Remark 3.1** It can be noticed that if $\boldsymbol{X_t} = (X_{t-m}, \ldots, X_t, X_{t+1}, \ldots, X_{t+h})$ is representable on $C_{m+h+1}^{\|\cdot\|}$ for some semi-norm satisfying (3.1), then it is representable on unit cylinders relative to any other semi-norms satisfying (3.1). This holds because (3.1) ensures that all these semi-norms have the same kernel. The notion of past-representability can thus be defined independently of the particular choice of a semi-norm.[7]

The following proposition provides a characterisation of past-representability.

**Proposition 3.1** *Let $(X_t)$ be an $\alpha$-stable moving average satisfying (3.2)-(3.4).*
$(\iota)$ *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*

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

$(a)$ *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 $m \geq m_0$.*

[7] This will not be true in general under the weaker assumption (1.2) and different notions of representability of a process could emerge depending on the kernels of the semi-norms.

(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* $C_{m+h+1}^{\|\cdot\|}$ *is then given by* (3.6) *with the Euclidean norm* $\|\cdot\|_e$ *replaced by the semi-norm* $\|\cdot\|$*.*

**Remark 3.2** Note in particular 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 $k < k_0$.

**Remark 3.3** Proposition 3.1 shows that for an $\alpha$-stable moving average to be past-representable, sequences of consecutive zero values in the coefficients $(d_k)$ have to be either of finite lengths, or infinite to the left. This surprisingly places the anticipativeness of a stable moving average as a necessary –and sufficient for $\alpha \neq 1$ and $(\alpha, \beta) = (1, 0)$– condition for its past-representability. The less anticipative a moving average is, in the sense of the larger the gaps of zeros in its forward-looking side, then the higher $m$ has to be chosen so as to have 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.

**Corollary 3.1** Let $(X_t)$ an $\alpha$-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.

**Remark 3.4** This fault line between anticipativeness and non-anticipativeness sheds light on the predictability of extreme events in linear processes. Consider for illustration the two following $\alpha$-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 $(\varepsilon_t)$, $(\eta_t)$ 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 stems from a large realisation of an underlying error $\varepsilon_\tau$ or $\eta_\tau$, at some time $\tau$. On the one hand for the non-anticipative AR(1) (3.12), a jump does not manifest any early visible sign 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) (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 more informative than their plain unconditional likelihood of occurrence. Building on the "information encoding" interpretation of spectral measures given in Remark 2.3, 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 condition for past-representability simplifies for ARMA processes and is equivalent to the autoregressive polynomial having at least one root located inside the unit circle.

**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 $\psi$, $\phi$, $\Theta$, $H$ 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 $\psi$ and $\Theta$ (resp. $\phi$ and $H$) have no common roots. Then, for any $\alpha \in (0, 2)$ and $\beta \in [-1, 1]$, the following statements are equivalent:

$(\iota)$ $(X_t)$ 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 $m_0 < +\infty$.

**Remark 3.5** For ARMA processes, we can notice in particular that the discrepancy between the cases [$\alpha \neq 1$ or $(\alpha,\beta) = (1,0)$] and [$\alpha = 1$, $\beta \neq 0$] vanishes. Also, only the roots of the AR polynomial matter for past-representability, the MA part having no role.

# 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 as well as particular examples.

To be relevant for the prediction framework, the Borel set $B$ appearing in Proposition 2.2 has to 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 realisations $X_{t+1}, \ldots, X_{t+h}$. For $\|\cdot\|$ a semi-norm on $\mathbb{R}^{m+h+1}$ satisfying (3.1), denote $S_{m+1}^{\|\cdot\|} = \{(s_{-m}, \ldots, s_0) \in \mathbb{R}^{m+1} : \quad \|(s_{-m}, \ldots, s_0, 0, \ldots, 0)\| = 1\}$.[8] 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 $B(V) = C_{m+1}^{\|\cdot\|}$. In the following, we will use Borel sets of the above form to condition the distribution of the complete 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 some small neighbourhood on $S_{m+1}^{\|\cdot\|}$. It will be useful in the following to notice that

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

[8] 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.

where $f$ 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 $\alpha$-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 $C^{\|\cdot\|}_{m+h+1}$ for some semi-norm satisfying (3.1), we know by Proposition 3.1 ($\iota\iota$), that $\Gamma^{\|\cdot\|}$ is of the form

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

**Proposition 4.1** *Under the above assumptions, we have*

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

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

**Remark 4.1** ($\iota$) Setting $V = S^{\|\cdot\|}_{m+1}$, and $A$ an arbitrarily small closed neighbourhood 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 terms, when far from central values, the trajectory of 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 $k$ points to which piece $(d_{k+m}, \ldots, d_k, d_{k-1}, \ldots, d_{k-h})$ of this moving average it corresponds, 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 $(\varepsilon_\tau)$). The likelihood of a pattern $\vartheta \boldsymbol{d}_k / \|\boldsymbol{d}_k\|$ can be evaluated by setting $A$ to be a small neighbourhood of that point. ($\iota\iota$) In view of point ($\iota$), the observed path $(X_{t-m}, \dots, X_{t-1}, X_t)/\|\boldsymbol{X_t}\|$ will *a fortiori* be of the same shape as some $\vartheta(d_{k+m}, \dots, d_{k+1}, d_k)/\|\boldsymbol{d}_k\|$ when an extreme event will approach in time. Observing the initial part of the pattern can give information about the remaining unobserved piece: the conditional likelihood of the latter can be assessed by setting $V$ to be a small neighbourhood of the observed pattern.

**Remark 4.2** The tail conditional distribution given in (4.3) highlights three types of uncertainty/approximation for prediction:[9]

($\iota$) 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 realised trajectory. The choice of an adequate conditioning neighbourhood $V$ in (4.3) given a piece of trajectory will thus have to rely on a statistical approach. One could envision tests of hypotheses to determine whether a piece of realised (noisy) trajectory "is more similar" to a certain pattern 1 or to an other pattern 2.

($\iota\iota$) Even for an arbitrarily small neighbourhood $V$ –that is, 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 by the last $h$. The stable anticipative AR(1) is a typical example of this phenomenon that will be studied in the next section. Interestingly, the stable anticipative AR(2) eludes to this as discussed in hereafter.

($\iota\iota\iota$) The tail conditional distribution (4.3) is an asymptotic behaviour as the (semi-)norm of $\boldsymbol{X_t}$ grows infinitely large. It is thus only an approximation of the true dynamics

[9] 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 $(d_{j,k})$ has to be inferred from data.

during extreme events. It would be interesting to obtain a finer asymptotic development in $x$ of the above convergence to gauge the approximation error of the true conditional distribution. It would be especially useful to quantify how far from/how variable around the predicted patterns the future path can be.

### 4.2 The anticipative AR(1)

We now consider $(X_t)$ the stable anticipative AR(1) processes 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 next lemma.

**Lemma 4.1** *Let $(X_t)$ be an $\alpha$-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*

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

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

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

*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 $\rho$ is positive. Let us first introduce useful neighbourhoods 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 $(\vartheta, k)$ 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 neighbourhood 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 $C^{\|\cdot\|}_{m+h+1}$ in which the only points charged by $\Gamma^{\|\cdot\|}$ all have the first $(m+1)^{\text{th}}$ coinciding with $\vartheta_0 f(\boldsymbol{d}_{k_0})/\|\boldsymbol{d}_{k_0}\|$. Define also $A_{\vartheta,k}$ for any $(\vartheta, k)$ as any closed neighbourhood of $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ which does not contain any other charged point of $\Gamma^{\|\cdot\|}$, that is,

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

**Proposition 4.2** *Let $(X_t)$ be an $\alpha$-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 neighbourhood 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 neighbourhood of some $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ as in (4.7), the following hold.*

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

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A_{\vartheta,k}\Big|B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \begin{cases} \dfrac{w_{\vartheta_0}}{p_{\vartheta_0}}\delta_{\{\vartheta_0\}}(\vartheta), & k=0 \\[2em] |\rho|^{\alpha k}(1-|\rho|^\alpha)\delta_{\{\vartheta_0\}}(\vartheta), & 1\le k\le h-1, \\[2em] |\rho|^{\alpha h}\delta_{\{\vartheta_0\}}(\vartheta), & k=h, \end{cases}$$

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

**Remark 4.3** For $m \geq 1$, that is, if the observed path is assumed to be of length at least 2, there is a significant difference between whether $k_0 \in \{0, \dots, h\}$ or $k_0 \in \{-m, \dots, -1\}$. For the latter, the asymptotic probability of the whole path $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ being in an arbitrarily small neighbourhood of $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ is 1 if and only if $\vartheta = \vartheta_0$, $k = k_0$: given the observed path, the shape of the future trajectory is fully determined. For the former, this probability is strictly positive if and only if $\vartheta = \vartheta_0$, but the observed pattern is compatible with several distinct future paths. One can see why this is the case from the form of the sequences $\boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ and of their restrictions to the first $m+1$ components $f(\boldsymbol{d}_k)/\|\boldsymbol{d}_k\|$. On the one hand (omitting $\vartheta$),

$$\frac{\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|} = \begin{cases} \dfrac{(\overbrace{\rho^{k+m},\dots,\rho^k}^{m+1},\overbrace{\rho^{k-1},\dots,\rho,1,0,\dots,0}^{h})}{\|(\rho^{k+m},\dots,\rho^k,\rho^{k-1},\dots,\rho,1,0,\dots,0)\|}, & \text{for} \qquad k\in\{0,\dots,h\}, \\[3em] \dfrac{(\rho^{k+m},\dots,\rho,1,0,\dots,0,0,\dots,0)}{\|(\underbrace{\rho^{k+m},\dots,\rho,1,0,\dots,0}_{m+1},\underbrace{0,\dots,0}_{h})\|}, & \text{for} \qquad k\in\{-m,\dots,-1\}. \end{cases}$$

We can notice that all the above sequences are pieces of explosive exponentials, terminated at some coordinate. For $k \in \{0, \dots, h\}$, the first zero component –the “crash of the bubble”–, is situated at or after the $(m+2)^{\text{th}}$ component, whereas for $k \in \{-m, \dots, -1\}$, it is situated at or before the $(m+1)^{\text{th}}$. Using the homogeneity of the semi-norm and (1.2), we have on the other hand that

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

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 $k_0 \in \{-m, \ldots, -1\}$ amounts to condition on the burst of a bubble being observed in the past trajectory with no new bubble forming yet, which allows to identify exactly the position of the pattern on the moving average's coefficient sequence.

When conditioning with $k_0 \in \{0, \ldots, h\}$ however, the crash date is not observed and can happen either in the next $h-1$ periods, or after $h$. However, the shape of the observed path is that of a piece of exponential with growth rate $\rho^{-1}$ regardless of the remaining time before the burst, which leaves several future paths possible. One can quantify the likelihood of each potential scenario: the quantity $|\rho|^{\alpha k}(1-|\rho|^{\alpha})$ corresponds to the probability that the bubble will peak in exactly $k$ periods ($0 \leq k < h$), and $|\rho|^{\alpha h}$ corresponds to the probability that the bubble will last at least $h$ more periods.

**Remark 4.4** ($\iota$) The previous remark confirms the interpretation of the conditional moments proposed in Fries (2021). It also extends it by accounting for paths rather than point prediction. ($\iota\iota$) Notice that for $m = 0$ (only the present value is assumed to be observed), no pattern can be observed but only the sign of the shock. Hence, the growth rate $\rho^{-1}$ of the ongoing event is unidentifiable, which is reflected in the fact that the asymptotic probabilities of paths with growth rates $\rho^{-1}$, are positive (case ($\iota\iota$) of Proposition 4.2).

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

We focus here on two processes which both share the peculiar property of having a 0-1 tail conditional distribution whenever the observed path is of length at least 2 (i.e., $m \geq 1$): the anticipative AR(2) and the anticipative fractionally integrated white noise (FWN). For an adequate choice of the parameters, the former can generate bubble-like trajectories with accelerating or decelerating growth rate and the latter can accommodate hyperbolic bubbles. In contrast with the anticipative AR(1), these bubbles do not display an exponential profile but still feature an inflation-peak-collapse behaviour. Any extension of those two minimal specifications should preserve the following statements.

**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

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

where $\Gamma(\,\cdot\,)$ denotes –here only– 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 is of length at least 2 is exhibited in the next proposition.

**Proposition 4.3** *Let $(X_t)$ be the $\alpha$-stable anticipative AR(2) (resp. fractionally integrated AR) as in (4.8)-(4.9) (resp. (4.10)-(4.11)). 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) (resp. (4.11)). Let $V_0$ a small neighbourhood of $\vartheta_0 \boldsymbol{d}_{k_0}/\|\boldsymbol{d}_{k_0}\|$ as in (4.6) –where we drop the indexes $j$– for some $\vartheta_0 \in S_1$, $k_0 \geq -m$, and let $B(V_0) = V_0 \times \mathbb{R}^h$. Then,*

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

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

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

**Remark 4.5** 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,[10] 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.

[10] See point $(\iota)$ of Remark 4.2.

# 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

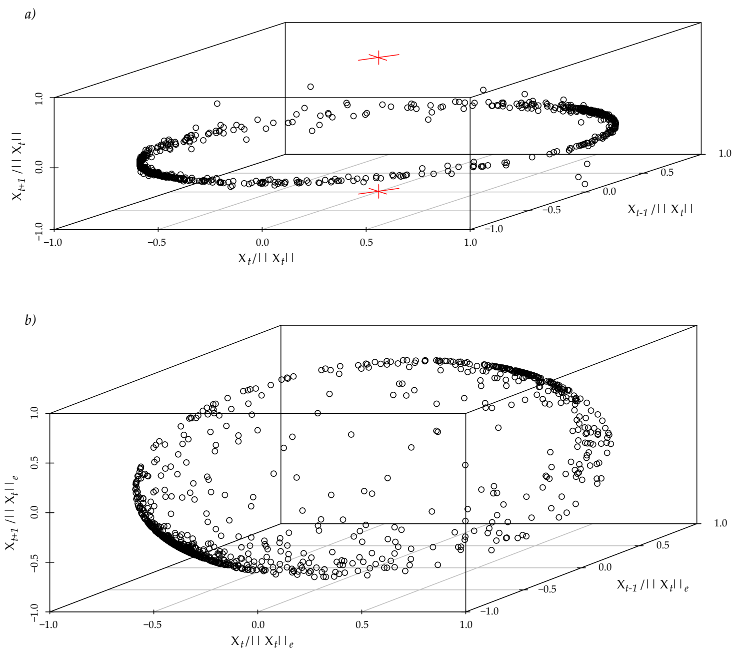
In the spirit of the Remark 2.3, we consider an $\alpha$-stable vector $\boldsymbol{X_t} = (X_{t-1}, X_t, X_{t+1})$ where $X_t$ is an anticipative AR(2) specified as in 5.2. $X_t$ being past-representable, it admits a representation on the unit-cylinder. Furthermore, as discussed in 4.3, its spectral 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 semi-norm representability reflect the fact that extreme realizations of $X_{t+1}$ never occur conditionally to small realisations 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

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




in Figure 1.b) situation where the realisation of $X_{t+1}$ is extreme whereas immediate past realisations are not.

## 5.2 Forecasting crash probabilities

To illustrate through simulations that the probability on the left-hand side of Proposition 4.3 converges to the right-hand side when the conditioning value $\|\boldsymbol{X_t}\|$ is large, we generate 1000 trajectories of $N = 10^6$ observations from the anticipative AR(2) process $X_t = 0.7X_{t+1} + 0.1X_{t+2} + \varepsilon_t$ where $\varepsilon_t \overset{i.i.d.}{\sim} S(1.5, 1, 0.5, 0)$. We focus here on the case $m = 1$ and evaluate the crash probabilities at different forecasting horizons $h = (1, 5, 10)$. 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 $\sqrt{X_t^2 + X_{t-1}^2} \geq 2q$ with $q$ is a theoretical quantile of the marginal distribution of $X_t$. Second we define the conditioning of the small neighbourhood

$B(V_0)$ and we set for $m = 1$ and

$$B(V_0) = \left[\frac{\vartheta_0 \boldsymbol{d}_{k_0-1}}{\|\boldsymbol{d}_{k_0}\|} - 0.1, \frac{\vartheta_0 \boldsymbol{d}_{k_0-1}}{\|\boldsymbol{d}_{k_0}\|} + 0.1\right] \times \left[\frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|} - 0.1, \frac{\vartheta_0 \boldsymbol{d}_{k_0}}{\|\boldsymbol{d}_{k_0}\|} + 0.1\right].$$

We also set $A = B(V_0) \times [-\delta, \delta]$, with $\delta = 0.3$. This is equivalent to estimating the probability of a crash at horizon $h$. 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

$$\hat{p}_q = \frac{\sum_{t=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-1}, X_t)}{\|\boldsymbol{X_t}\|} \in B(V_0)\right\} \cap \{\frac{X_{t+h}}{\|\boldsymbol{X_t}\|} \leq \delta\} \cap \{\|\boldsymbol{X_t}\| > 2q\}\right)}}{\sum_{t=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-1}, X_t)}{\|\boldsymbol{X_t}\|} \in B(V_0)\right\} \cap \{\|\boldsymbol{X_t}\| > 2q\}\right)}} \tag{5.1}$$

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

$$p_q = \frac{\sum_{t=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-1}, X_t)}{\|\boldsymbol{X_t}\|} \in B(V_0)\right\} \cap \{\frac{d_{k_0+h}}{\|\boldsymbol{d}_{k_0}\|} \leq \delta\} \cap \{\|\boldsymbol{X_t}\| > 2q\}\right)}}{\sum_{i=1}^{N-h} \mathbb{1}_{\left(\left\{\frac{(X_{t-1}, X_t)}{\|\boldsymbol{X_t}\|} \in B(V_0)\right\} \cap \{\|\boldsymbol{X_t}\| > 2q\}\right)}} \tag{5.2}$$

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 $\boldsymbol{d}_{k_0}/\|\boldsymbol{d}_{k_0}\| \in A$. 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 $m = 1$. Second, we evaluate $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ for $k \in (\underline{k}, \bar{k})$ where $\underline{k} = 30$ and $\bar{k} = 0$. Finally we check whether some $\vartheta \boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ belongs to a small neighborhood of $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$.

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.

### 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 document that in

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

| | $h = 1$ | $h = 5$ | $h = 10$ |
|---|---|---|---|
| $p_{0.9} \backslash \hat{p}_{0.9}$ | 84.09\22.75 (22.39-23.14) | 93.00\39.70 (39.27-40.14) | 98.45\46.76 (46.32-47.21) |
| $p_{0.99} \backslash \hat{p}_{0.99}$ | 91.56\89.11 (87.90-90.36) | 95.55\94.63 (93.67-95.56) | 95.71\96.85 (96.11-97.56) |
| $p_{0.999} \backslash \hat{p}_{0.999}$ | 99.50\98.75 (97.04-100) | 99.40\99.40 (98.21-100) | 99.72\99.67 (98.63-100) |
| $p_{0.9999} \backslash \hat{p}_{0.9999}$ | 99.96\99.86 (96.42-100) | 99.91\99.92 (99.90-100) | 99.97\99.98 (100-100) |

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 = 1000 simulated trajectories of $N = 10^6$ observations, for $q = q_a$ several $a$-quantiles of the marginal distribution of $X_t$.

practice, for large values of $x$, the approximation

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

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 4.2 affect the performance of our procedure in presence of finitely large realisations. As for a range of realisations, 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 $m$.

Our forecasting procedure proceeds in 4 steps. First, we compute $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ for a given $m$. 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 $k$ satisfied this condition, we reduce the neighbourhood until a unique $k = k_0$ remains. 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. At this stage, we structure $\boldsymbol{X_t}$ as in (3.5) and $\boldsymbol{d}_{k_0}$ as $(d_{k+m}, \dots, d_k, d_{k-1}, \dots, d_{k-h})$. 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

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

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

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 $\alpha=\{0.9,1.2,1.5,1.8\}$. Each simulated path is governed by a S$\alpha$S anticipative AR(2) of the following form: $X_t = 0.7X_{t+1}+0.1X_{t+2}+\varepsilon_t$ where $\varepsilon_t \overset{i.i.d.}{\sim} S(\alpha,0,0.1,0)$. 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 2 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 $\alpha=1.5$.

Table 2: Quantile of the last in-sample observation

| $N/\alpha$ | 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 |

In such a configuration, the realisations of $X_t$ are likely to be only moderately large compared to the asymptotic requirements ($x\to\infty$). 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.

Table 3: Bias for the crash date predictor

| | $m = 1$ | | | | $m = 3$ | | | |
|---|---|---|---|---|---|---|---|---|
| $q_{X_t}/\alpha$ | 0.9 | 1.2 | 1.5 | 1.8 | 0.9 | 1.2 | 1.5 | 1.8 |
| High | -0.9785 | -0.3985 | 0.0262 | 0.2199 | -0.7320 | -0.2420 | -0.0073 | 0.2815 |
| Quite High | 0.7174 | 1.3771 | 1.9292 | 2.2544 | 0.9421 | 1.6189 | 2.0938 | 2.2914 |
| Moderately High | 5.8112 | 6.6166 | 7.1317 | 7.4263 | 6.0680 | 6.8229 | 7.1698 | 7.3565 |

| | $m = 5$ | | | | $m = 7$ | | | |
|---|---|---|---|---|---|---|---|---|
| $q_{X_t}/\alpha$ | 0.9 | 1.2 | 1.5 | 1.8 | 0.9 | 1.2 | 1.5 | 1.8 |
| High | -0.5457 | -0.2076 | 0.0483 | 0.2300 | -0.5043 | -0.1256 | 0.1099 | 0.2715 |
| Quite | 1.2378 | 1.7442 | 2.1075 | 2.3412 | 1.2978 | 1.8118 | 2.0987 | 2.2571 |
| Moderately High | 6.2749 | 6.9284 | 7.2065 | 7.3582 | 6.3193 | 6.9760 | 7.2655 | 7.3763 |

| | $m = 9$ | | | | $m = 11$ | | | |
|---|---|---|---|---|---|---|---|---|
| $q_{X_t}/\alpha$ | 0.9 | 1.2 | 1.5 | 1.8 | 0.9 | 1.2 | 1.5 | 1.8 |
| High | -0.4079 | -0.0811 | 0.1556 | 0.2976 | -0.4200 | -0.0480 | 0.1891 | 0.3300 |
| Quite High | 1.3407 | 1.8471 | 2.1537 | 2.2857 | 1.3633 | 1.8568 | 2.1417 | 2.3565 |
| Moderately High | 6.3805 | 7.0095 | 7.2599 | 7.4021 | 6.4407 | 7.0253 | 7.3097 | 7.4745 |

The results are reported in Table 3. No matter the tail index considered, our procedure can predict the crash date with a bias less than one period as long as $X_t$ is sufficiently large and $m$ is carefully chosen. However, the results shed light on the crucial role of the 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. In such a case, the selection of $m$ is also very important as a large $m$ introduces more noise from observations that presumably belongs to lower quantiles (as we focus on positive shocks here). For a given $m$, the performance deteriorates when $\alpha$ 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 more noise. This confirms the Remark 3.3.

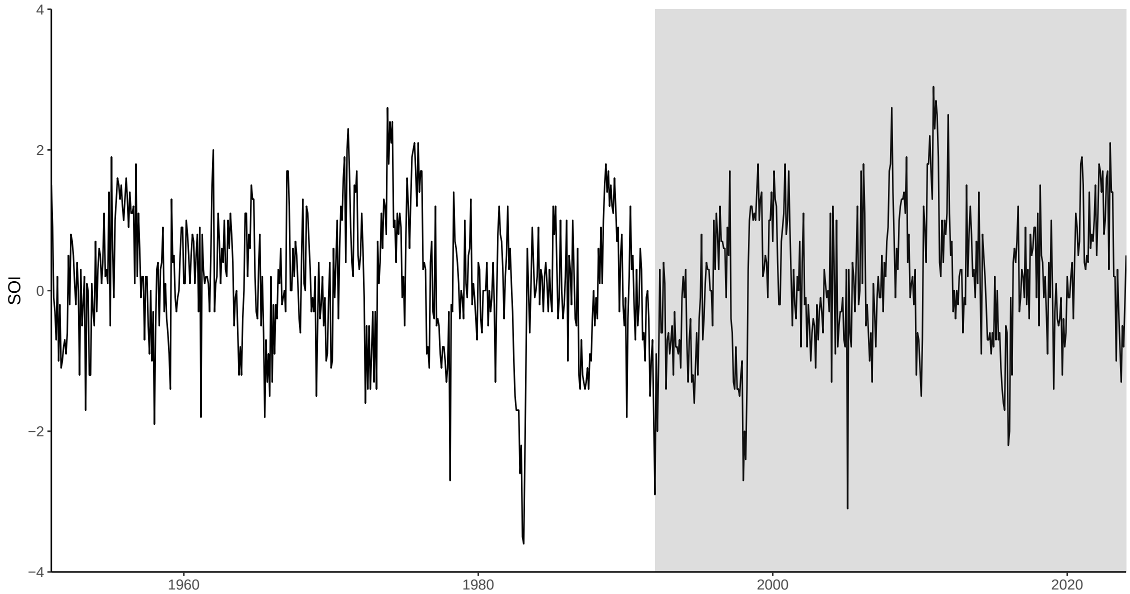
Our theory states that when $x \to \infty$, $m = 1$ can be sufficient. But in practice, the simulation study reveals that the optimal selection of $m$ is not obvious as it interact in a complex manner with the tail index $\alpha$. For instance, when $X_t$ is very high and the tail index is close to 1, a larger $m$ improves the performance of the forecasting procedure. This is not very surprising as $X_t$ is far from central values, $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ is very collinear to the moving average coefficients and the $m$ past observations are not too noisy and help to identify the pattern. At the opposite, when the tail index get closer to the "light" tail case, larger values of $m$ become detrimental and small $m$ are preferable. The same analysis holds for $X_t$ large and moderately large.

## 6 Forecasting climate anomalies

A growing literature highlights the impact of climate variables on economic performance (Dell et al., 2014), a key variable to identify this impact is the El Niño (resp. La Niña) weather shocks. It is known that these shocks have an impact among others on growth, inflation, energy and agricultural commodity returns (Brenner, 2002; Cashin et al., 2017; Makkonen et al., 2021). Providing a forecast of El Niño weather shocks, is of primary interest, as it provides numerous societal benefits, from extreme weather warnings to agricultural planning (Alley et al., 2019). El Niño intensity is defined as a value construct from the Southern Oscillation Index (SOI)[11] This section discusses the performance of the proposed approach in detecting the peak of an El Niño (resp. La Niña) shock and assesses the probability of staying in these episode $h$ period ahead. We split the data in an in-sample

[11] Data and methodology to construct the SOI are available here https://www.ncei.noaa.gov/access/monitoring/enso/soi. SOI is a monthly variable based on air-pressure differentials in the South Pacific, between Tahiti and Darwin.

Figure 2: Southern Oscillation Index (SOI)




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

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 2 display the data sample where the shaded area corresponds to the out-of-sample data. The alternance of boom and burst which seem to be an identifiable deterministic pattern, is distinguishable in Figure 2.

We then rely on the procedure of Lanne and Saikkonen (2011) and the MARX package (see Hecq et al., 2017b, 2020) to select the best specification for SOI. Table 4, summarise the results of this procedure: BIC determine the $p_{pseudo} = r + s$, here equal to 2, then using likelihood criterion, we select the best specification for causal-non-causal models, which is the anticipative AR(2)

$$(1 - \varphi_1 F - \varphi_2 F^2)X_t = \varepsilon_t$$

. Table 5 reports the estimation results for the anticipative AR(2) parameters that drive the dynamics of SOI. The parameters of the retained model are subsequently estimated using a modified version of the MARX package suitable for $\alpha$-stable laws. Standard deviations are estimated using finite differences gradient and Hessian for the parameters in the right space. From Table 5, Ljoung-Box (LB-Test) show that all the autocorrelation in the residuals of AR(2) are removed and the Jarque-Bera test (JB-Test) that the residuals are indeed non-gaussian ($\alpha$ <2).

Table 4: Identification of non causal processes for SOI

| BIC | AR(2,0) | AR(1,1) | AR(0,2) |
|---|---|---|---|
| $p_{pseudo}$ | | Likelihood | |
| 2 | -513.710 | -510.764 | -507.253 |

Table 5: AR(2) estimation for SOI

| $\varphi_1$ | $\varphi_2$ | $\alpha$ | $\beta$ | $\sigma$ | $\mu$ |
|---|---|---|---|---|---|
| 0.44*** | 0.28*** | 1.90*** | 1.00* | 0.46*** | -0.06* |
| (5.22E-13) | (2.11E-12) | (5.03E-7) | (5.54E-4) | (1.94E-07) | (1.78E-07) |

| Specification test | Stats | p-value |
|---|---|---|
| LB-Test (lag=5) | 7.81 | 0.16 |
| JB-Test | 20.76 | $< 10^{-5}$ |

*Notes:* Estimated parameters of $\alpha$-stable anticipative AR(2) process associated with the SOI series for the period 01/1951 - 12/1991. Standard deviations are in parentheses. Asterisks *, **, and *** indicate significance at the 90%, 95% and 99% level, respectively.

An El Niño (resp. La Niña) shock, is defined as SOI below -1 (resp. over 1) during at least the 3 periods. We hence estimate the probability of SOI to go back toward central values after $h$ periods with $h = 3, 5$ by using the procedure detailed in Section 5.2. We set the same neighbourhood $B(V_0)$ of $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ and we consider that $\|\boldsymbol{X_t}\|$ is large when $\sqrt{X_t^2 + X_{t-1}^2} \geq 2q$ with $q = q_{0.95}$, the quantile at 95% of the marginal distribution of $X_t$.[12] As we rely on 492 in-sample (877 whole-sample) observations, higher quantiles are sparse

[12] We apply our procedure to the absolute value of SOI.

and cannot be considered here. We choose $\delta = 0.3$. The results are reported in Table 6 and exhibit a very high empirical (average) probability of returning to central values 3 periods ahead. However, this probability is less than unity revealing that El Niño episodes can occur (or persist if the SOI was already above 1 or below -1 for several periods). When $h = 5$, $\hat{p}_{0.95} = 1$ meaning that very persistent El Niño (resp. La Niña) occurrences are unlikely to appear.

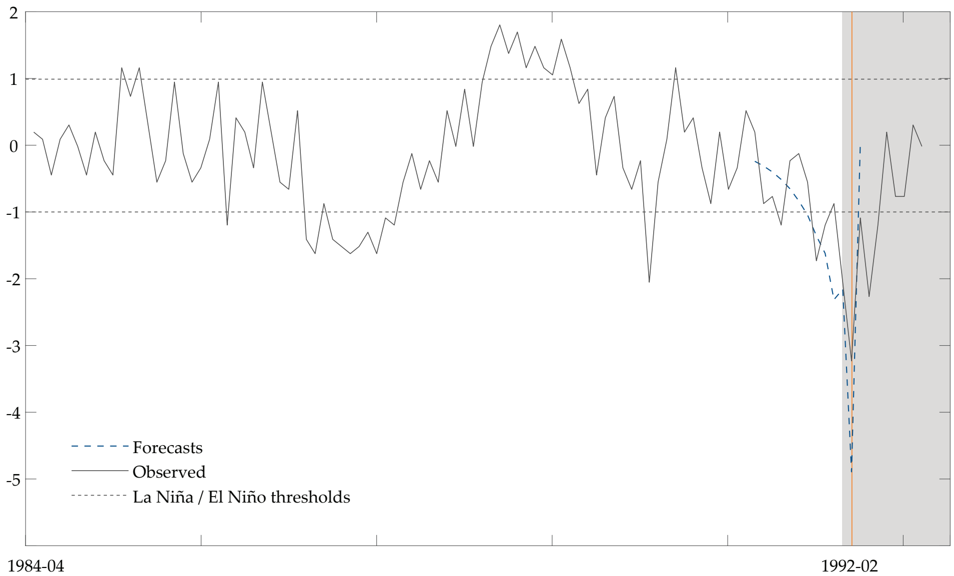
Table 6: Comparison of theoretical and empirical SOI reversal probabilities

| | | $h=1$ | $h=3$ | $h=5$ |
|---|---|---|---|---|
| In-sample | $p_{0.95}\backslash\hat{p}_{0.95}$ | 60.00\65.00 | 100\65.00 | 100.00 \100.00 |
| Out-of-sample | $p_{0.95}\backslash\hat{p}_{0.95}$ | 58.00\53.00 | 76.00 \100.00 | 100.00\100.00 |

Notes: The theoretical reversal probabilities $p_q$ are computed using (5.2). Empirical average probabilities are computed using (5.1).

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 $\vartheta_0 \boldsymbol{d}_{k_0}$. The selected piece of the

Figure 3: El Niño reversal forecast




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

trajectory is represented in Figure 3. 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 reversal date in Figure 3, 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 7. 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 7; 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 7 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 7 the selected $k_0$ and $m$ from our procedure. For SOI, the choice of these two parameters is robust, mainly $m = 10$ and $k_0 = 2$ are chosen.

Table 7: 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 $\alpha$-stable infinite moving averages, the conditional distribution of future paths given the observed past trajectory during extreme events is obtained based on a new spectral representation of stable random vectors on unit cylinders relative to semi-norms. In contrast to usual norm representations, this yields a multivariate regularly varying tails property that is appropriate for prediction purposes, however not all stable random vectors can be represented on semi-norm unit cylinders. A characterisation is provided and reveals that predictions are possible if and only if the process is "anticipative enough". Finite length paths of $\alpha$-stable moving averages, which are themselves multivariate $\alpha$-stable, are embedded into this framework.

Our approach also shows that instead of their attractive "causal" interpretation, non-anticipative processes appear to rather presume, by construction, the unpredictability of extreme events. Anticipative processes, however, instead of "depending on the future", rather assume that future events feature early visible signs betraying their incoming occurrences. These early signs take the form of emerging trends and patterns that an observer can identify and use to infer about future potential outcomes. In some particular cases, we demonstrate that the trajectory does not leave room for indeterminacy and can be deduced, in theory with certainty, and 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 forecasting crash dates. We also discuss some sources of uncertainty that are likely to arise in finite sample and non-asymptotic frameworks. The numerical analysis confirms that both the two procedures we implement are easy to use and perform well in a wide range of situations. To give more insights regarding the empirical relevance of the semi-norm representation of $\alpha$-stable moving averages we show how climate anomalies can be predicted accurately. In particular, the probabilities of occurrence of the so-called La Niña and El Niño episodes are estimated. For a specific El Niño episode, we also detect very precisely, out-of-sample, the reversal date.

## 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 $\alpha$-stable autoregressive process. Annals of Statistics 37, 1946–1982.

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.

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 .

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.

Cubadda, G., Hecq, A., Telg, S., 2019. Detecting co-movements in non-causal time series. Oxford Bulletin of Economics and Statistics 81, 697–715.

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.

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

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., 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.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.

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 .

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

Saikkonen, P., Sandberg, R., 2016. Testing for a unit root in noncausal autoregressive models. Journal of Time Series Analysis 37, 99–125.

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

# 8 Proofs

## 8.1 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 semi-norm $\|\cdot\|$. 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 $\boldsymbol{s}' = T_{\|\cdot\|}^{-1}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|_e$,

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

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

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

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 $\Gamma$ 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 $a = 2/\pi$,

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

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

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

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

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

First, we can see that

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

We will later show the following result:

**Lemma 8.1** *Let $\gamma^{\|\cdot\|}$ a Borel measure on $C_d^{\|\cdot\|}$ satisfying (2.4). Then,*

$$\int_{C_d^{\|\cdot\|}} \|\boldsymbol{s}\|_e \Big| \ln \|\boldsymbol{s}\|_e \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty. \tag{8.1}$$

Assuming Lemma 8.1 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

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

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 $\Gamma$ 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 8.1) 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 8.1**

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

$$\int_{C_d^{\|\cdot\|}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \ln |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) = \int_{C_d^{\|\cdot\|} \cap \{b \leq \|\boldsymbol{s}\|_e \leq M\}} + \int_{C_d^{\|\cdot\|} \cap \{\|\boldsymbol{s}\|_e > M\}} < +\infty. \tag{8.2}$$

and thus, in particular

$$\begin{aligned}
&\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|} \colon \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \ln |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\
&\qquad = \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|} \colon \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \ln \|\boldsymbol{s}\|_e + \ln |\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e}\rangle| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty.
\end{aligned} \tag{8.3}$$

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

$$\begin{aligned}
&\int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|} \colon \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \ln \|\boldsymbol{s}\|_e + \ln |\langle \boldsymbol{u}, \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|_e}\rangle| \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\
&\qquad = \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|} \colon \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \ln \|\boldsymbol{s}\|_e \Big| \Big| 1 + \frac{\ln |\langle \boldsymbol{u}, \boldsymbol{s}/\|\boldsymbol{s}\|_e\rangle|}{\ln \|\boldsymbol{s}\|_e} \Big| \gamma^{\|\cdot\|}(d\boldsymbol{s}) \\
&\qquad \geq \int_{\{\boldsymbol{s}' \in C_d^{\|\cdot\|} \colon \|\boldsymbol{s}'\|_e > M\}} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle| \Big| \ln \|\boldsymbol{s}\|_e \Big| \Bigg| 1 - \Big| \frac{\ln |\langle \boldsymbol{u}, \boldsymbol{s}/\|\boldsymbol{s}\|_e\rangle|}{\ln \|\boldsymbol{s}\|_e} \Big| \Bigg| \gamma^{\|\cdot\|}(d\boldsymbol{s})
\end{aligned} \tag{8.4}$$

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 $\boldsymbol{s} = (s_1, \ldots, s_d) \in R_i$, $\sup_j |s_j| = |s_i|$.[13] We have by (8.3)-(8.4) that for any

[13] 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.

$i = 1, \dots, d$, any $\boldsymbol{u} \in \mathbb{R}^d$,

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

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

$$\int_{\{\boldsymbol{s'} \in C_d^{\|\cdot\|} :\ \|\boldsymbol{s'}\|_e > M\} \cap R_i} |\langle \boldsymbol{e}_i, \boldsymbol{s} \rangle| \Big| \ln \|\boldsymbol{s}\|_e \Big| \left| 1 - \left| \frac{\ln |\langle \boldsymbol{e}_i, \boldsymbol{s}/\|\boldsymbol{s}\|_e \rangle|}{\ln \|\boldsymbol{s}\|_e} \right| \right| \gamma^{\|\cdot\|}(d\boldsymbol{s}) < +\infty. \tag{8.5}$$

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 $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_d)$,

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

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 $j = 1, \dots, d$

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

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 (8.5)

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

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

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

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 $\bigcup_{i=1,\ldots,d} R_i = \mathbb{R}^d$,

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

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 $i = 1, \ldots, d$, 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 $i = 1, \ldots, d$, $\boldsymbol{s} \in \{\boldsymbol{s}' \in C_d^{\|\cdot\|} : b \leq \|\boldsymbol{s}'\|_e \leq M\} \cap R_i$. From (8.2), 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 $i = 1, \ldots, d$. Evaluating the above in particular at $\boldsymbol{u} = \lambda \boldsymbol{e}_i$, for any $\lambda > 2b^{-1} e^A$, we get

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

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

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

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

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

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

### 8.2 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 8.1** *Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ be an $\alpha$-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 $\Gamma^{\|\cdot\|}(\partial A) = 0$,*

$$\lim_{x\to+\infty} x^{\alpha}\mathbb{P}\Big(\|\boldsymbol{X}\| > x, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big) = C_\alpha \Gamma^{\|\cdot\|}(A), \tag{8.6}$$

*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 semi-norm, $K^{\|\cdot\|} = \{\boldsymbol{s} \in S_d : \|\boldsymbol{s}\| = 0\}$, and $\Gamma^{\|\cdot\|}(K^{\|\cdot\|}) = 0$, we have the series representation of $\boldsymbol{X}$, $(X_1, \ldots, X_d) \overset{d}{=} (Z_1, \ldots, Z_d)$ where

$$Z_k = (C_\alpha \Gamma^{\|\cdot\|}(C_d^{\|\cdot\|}))^{1/\alpha} \sum_{i=1}^{\infty} [\Gamma_i^{-1/\alpha} S_i^{(k)} - b_{i,k}(\alpha)], \quad k = 1, \ldots, d, \tag{8.7}$$

with $\boldsymbol{S}_i = (S_i^{(1)}, \ldots, S_i^{(d)})$, $i \geq 1$, 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 semi-norm 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 $\alpha$-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

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

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 (8.7). 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 semi-norms and the fact that any norm is finer than any semi-norm in finite dimension.[14] $\square$

## 8.3 Proof of Lemma 3.1

From Proposition 2.1, we know that a necessary condition for the representability of $\boldsymbol{X_t}$ on $C^{\|\cdot\|}_{m+h+1}$ 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 $\Gamma$ only charges discrete atoms on $C^{\|\cdot\|}_{m+h+1}$,

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

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

$$\forall k \in \mathbb{Z}, \quad \Big[(d_{k+m}, \ldots, d_k) = \mathbf{0} \Longrightarrow (d_{k+m}, \ldots, d_{k-h}) = \mathbf{0}\Big]. \tag{8.8}$$

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

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

It implies that

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

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

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

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

Invoking (8.8), we deduce by recurrence that for any $n \geq 0$,

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

Therefore, (8.8) 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 $\Gamma$ 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

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

### 8.4 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 Lemma 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 Lemma 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

$$d_{k+m_0} = \ldots = d_{k+1} = 0, \quad \text{and} \quad d_k \neq 0. \tag{8.9}$$

Because $(X_t)$ is $(m,h)$-past-representable, we have by Lemma 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 Lemma 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

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

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,

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

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.

### 8.5 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 $m_0 = +\infty$.

### 8.6 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

$$\begin{aligned}
& & \sum_{i=-q}^{p} \varphi_i X_{t+i} &= \sum_{k=-r}^{s} \theta_k \varepsilon_{t+k} \\
&\iff & \sum_{i=-q}^{p} \varphi_i \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k+i} &= \sum_{k=-r}^{s} \theta_k \varepsilon_{t+k} \\
&\iff & \sum_{k\in\mathbb{Z}} \Big( \sum_{i=-q}^{p} \varphi_i d_{k-i} \Big) \varepsilon_{t+k} &= \sum_{k=-r}^{s} \theta_k \varepsilon_{t+k}. \qquad (8.10)
\end{aligned}$$

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

$$\sum_{i=-q}^{p} \varphi_i d_{k-i} = 0. \qquad (8.11)$$

*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 (8.11) 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 subsequence $(k_{g(n)})$. 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

$$X_t = \sum_{k=-\infty}^{\bar{k}} d_k \varepsilon_{t+k}, \qquad t \in \mathbb{Z}. \tag{8.12}$$

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 $\ell_1$ is the degree of the largest order monomial in $B$ of $\Theta(F)H(B)b_1(B)$ (recall that $F = B^{-1}$) and $\ell_2$ is the degree of the largest monomial in $F$ of $B^p\Theta(F)H(B)b_2(B)$. By (8.12), 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

$$\Theta(z^{-1})H(z)b_1(z) = \psi(z^{-1})(P(z^{-1}) + Q(z)), \qquad z \in \mathbb{C}. \tag{8.13}$$

As $\deg(\psi) = p$ and $\psi(z) = 0$ if and only if $|z| > 1$, we know that there are $p$ complex numbers $z_1, \dots, z_p$ such that $0 < |z_i| < 1$ and $\psi(z_i^{-1}) = 0$ for $i = 1, \dots, p$. Evaluating (8.13) at the $z_i$'s, we get that

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

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 $\psi$ and $\Theta$ 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 $\boldsymbol{d}_k = (d_{k+m}, \dots, d_k, d_{k+1}, \dots, d_{k-h})$). 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 $(\iota)$ 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 Lemma 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.

## 8.7 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).

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

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

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,

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

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

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

## 8.9 Proof of Proposition 4.2

**Lemma 8.2** *Let $\Gamma^{\|\cdot\|}$ be the spectral measure given in Lemma 4.1 and assume that the $\rho$ is positive.*

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

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

*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*

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

*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 $\boldsymbol{m\geq 1}$ and $\boldsymbol{k_0\in\{0,\dots,h\}}$**

If $k' \in \{-m, \dots, -1\}$, 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

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

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

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

and by (3.1) we also have that

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

Thus,

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

because $\rho \neq 0$ is assumed.

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

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

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

$$f(\boldsymbol{d}_{k'}) = (\overbrace{\rho^{k'+m}, \ldots, \rho, 1, 0, \ldots, 0}^{m+1}, \overbrace{0, \ldots, 0}^{h}),$$

$$f(\boldsymbol{d}_{k_0}) = (\underbrace{\rho^{k_0+m}, \ldots, \rho, 1, 0, \ldots, 0}_{m+1}, \underbrace{0, \ldots, 0}_{h}),$$

and we also have that

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

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

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

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

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

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

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

$$\mathbb{P}_x^{\|\cdot\|}\left(\boldsymbol{X}_{\boldsymbol{t}}, A_{\vartheta,k} \middle| B(V_0)\right) \underset{x\to\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \quad \frac{\vartheta' f(\boldsymbol{d}_{,k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right)}{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right)}. \tag{8.14}$$

Focusing on the denominator, we have by (4.6)

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

We will now distinguish the cases arising from the application of Lemma 8.2. Recall that we assume for this proposition that the $\rho$ is positive. Thus, $\text{sign}(\rho) = 1$ and $\bar{\beta} = \beta \frac{1-|\rho|^\alpha}{1-\rho^{<\alpha>}} = \beta$ 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 8.2,

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

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,

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

Similarly for the numerator in (8.14), by (4.7),

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

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

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

The conclusion follows.

**Case $m \ge 1$ and $-m \le k_0 \le -1$**

We have by Lemma 8.2

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

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

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

and

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

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

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

The conclusion follows as previously.

**Case $\boldsymbol{m = 0}$**

By Lemma 8.2, as the $\rho$ is positive

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

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

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

Similarly, by (4.7),

$$
\begin{aligned}
&\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in A_{\vartheta,k} : \quad \frac{\vartheta' f(\boldsymbol{d}_{k'})}{\|\boldsymbol{d}_{k'}\|} \in V_0\right\}\right) \\
&\qquad = \Gamma^{\|\cdot\|}\left(A_{\vartheta,k} \cap \left\{\frac{\vartheta_0 \boldsymbol{d}_{k'}}{\|\boldsymbol{d}_{k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \quad k' \in \{0,\dots,h\} \cup \{(0,0)\}\right\}\right) \\
&\qquad = \begin{cases} \Gamma^{\|\cdot\|}\left(\left\{\dfrac{\vartheta_0 \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}\right\}\right), & \text{if } \vartheta = \vartheta_0, \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if } \vartheta \neq \vartheta_0, \end{cases} \\
&\qquad = \begin{cases} w_{j\vartheta_0} \delta_{\{\vartheta_0\}}(\vartheta), & \text{if} \quad k = 0, \\ w_{\vartheta_0} |\rho|^{\alpha k} \delta_{\{\vartheta_0\}}(\vartheta), & \text{if} \quad 1 \leq k \leq h-1, \\ w_{\vartheta_0} \dfrac{|\rho|^{\alpha h}}{1-|\rho|^{\alpha}} \delta_{\{\vartheta_0\}}(\vartheta), & \text{if} \quad k = h. \end{cases}
\end{aligned}
$$

The conclusion follows.

## 8.10 Proof of Proposition 4.3

**Lemma 8.3** *Let $X_t$ be the $\alpha$-stable anticipative AR(2) (resp. fractionally integrated AR) as in (4.8) (resp. (4.10)). With $f$ as in (4.1), and for any $m \geq 1$, $h \geq 0$,*

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

*Proof.*

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

$$
\begin{aligned}
\frac{f(\vartheta_1 \boldsymbol{d}_k)}{\|\boldsymbol{d}_k\|} = \frac{f(\vartheta_2 \boldsymbol{d}_\ell)}{\|\boldsymbol{d}_\ell\|} &\iff \left[\forall\, i = 0,\dots,m, \quad \frac{\vartheta_1 d_{k+i}}{\|\boldsymbol{d}_k\|} = \frac{\vartheta_2 d_{\ell+i}}{\|\boldsymbol{d}_\ell\|}\right] \\
&\iff \frac{d_k}{d_\ell} = \frac{d_{k+1}}{d_{\ell+1}} = \dots = \vartheta_1 \vartheta_2 \frac{\|\boldsymbol{d}_k\|}{\|\boldsymbol{d}_\ell\|}.
\end{aligned} \tag{8.15}
$$

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,

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

□

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 8.3,

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

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.