Prediction of bubbles in presence of α-stable aggregates moving averages

Gilles de Truchis[1], Sébastien Fries[2], Arthur Thomas[3]

Abstract

Financial markets frequently exhibit dramatic episodes where asset prices undergo rapid growth followed by abrupt collapses, that are incompatible with standard linear time series models. While anticipative heavy-tailed linear processes offer a promising alternative for modeling such phenomena, they impose uniform bubble patterns across different episodes, contradicting empirical evidence. This paper introduces a new model, based on α-stable moving average aggregates, that accommodates heterogeneous bubble dynamics. We establish the theoretical properties of this model, demonstrating that it admits a semi-norm representation on a unit cylinder, thereby enabling the prediction of extreme trajectories with varying growth dynamics. We develop a minimum distance estimation procedure based on the joint characteristic function and establish its asymptotic properties. Monte Carlo simulations confirm the estimator's good finite-sample performance across various specifications, and we implement a subsampling methodology to empirically verify the convergence to asymptotic normality. Our empirical application to the CBOE Crude Oil ETF Volatility Index successfully decomposes observed volatility dynamics into distinct components with different persistence properties, revealing that what appears as a single bubble episode actually consists of multiple superimposed processes with heterogeneous growth rates and crash probabilities.

Keywords: Aggregated processes, Stable random vectors, Spectral representation, Anticipative processes, Financial bubbles

[1]Université d'Orléans, Laboratoire d'Economie d'Orléans (LEO), 45000 Orléans, France, gilles.detruchis@univ-orleans.fr.
[2]Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[3]Corresponding author: Université Paris Dauphine, Université PSL, LEDA, CNRS, IRD, 75016 Paris , France, arthur.thomas@dauphine.psl.eu.

1. Introduction

Financial markets regularly witness dramatic episodes where asset prices undergo rapid growth followed by abrupt collapses. These phenomena, termed rational asset pricing bubbles when they diverge from fundamental values (Blanchard and Watson , 1982; Tirole , 1985), have become increasingly prominent alongside well-documented features such as heavy-tailed distributions and volatility clustering. These bubbles emerge as solutions to linear rational expectation models that admit multiple stationary equilibria through infinite variance innovations (Gouriéroux et al. , 2020). Another theoretical paper that provides justification for bubbles having different growth rates is Lux and Sornette (2002), which demonstrates how agent interactions can create bubble formation, with bubbles growing in seemingly rational ways driven by investor expectations. Their model effectively captures sudden dramatic crashes and replicates the "fat tails" observed in empirical financial market data, suggesting that traditional representative rational agent models inadequately explain these phenomena.

From an empirical perspective, so-called mixed-causal (or anticipative) models appear as good candidates to account for the non-linear dynamics of bubbles and the non-Gaussian environment characterized by Lux and Sornette (2002) and Gouriéroux et al. (2020). Indeed, future-oriented models may generate intermittent periods of explosive growth and relative stability within a stationary linear framework while also admitting a regular time representation involving non-linear dynamics or non-i.i.d. innovations. Among others, we can mention (Andrews et al. , 2009; Lanne and Saikkonen , 2011, 2013; Hecq et al. , 2016, 2017; Cavaliere et al. , 2020; Velasco and Lobato , 2018; Fries and Zakoian , 2019; Hecq et al. , 2020; Gouriéroux and Jasiak , 2016, 2018; Gourieroux and Jasiak , 2023; Hecq and Velasquez-Gaviria , 2025; Gouriéroux et al. , 2025). Most importantly, this framework exhibits intriguing properties, such as a predictive distribution with lighter tails than the marginal distribution. This enables more accurate predictions of higher-order moments (see e.g. Fries , 2022) and forecasts based on pattern recognition (see de Truchis et al. , 2025a), which are critical for informed investment decisions.

However, anticipative models impose a similar increase rate for all bubbles, fully determined by the non-causal autoregressive coefficients (Gouriéroux and Zakoian , 2017). This lack of flexibility might conflicts with empirical evidence on financial markets where the surge of explosive episodes can exhibit very different pattern. Moreover, Gouriéroux et al. (2021) recall that aggregation implies various sources of noise and is hence very different from mixed-causal AR processes and more generally, different from any two-sided moving average. As it incorporates independent unobservable stochastic factors aggregation it is more suitable for financial applications. For instance, if one want to build derivatives to hedge portfolios against the uncertainty associated with the anticipative components and the risk of sudden bubble crashes, the two factor of risk should be priced and accounted for in the derivatives.

In this paper, we make two contributions to the literature on econometric modeling of financial bubbles.

First, we introduce a novel flexible framework that overcomes a key limitation of existing anticipative heavy-tailed models, which impose uniform growth patterns across different bubble episodes. Our approach allows for diverse bubble dynamics by aggregating multiple latent components, each with distinct stochastic properties. We derive the theoretical tail properties of this model and demonstrate that, similarly to non-aggregated processes (de Truchis et al. , 2025a), it admits a semi-norm representation on a unit cylinder, expect if one of the underling component is purely non-anticipative, thereby enabling the prediction of extreme trajectories with heterogeneous growth patterns.

Second, we develop an inference procedure for anticipative stable aggregates, departing from Gouriéroux and Zakoian (2017) and building upon Knight and Yu (2002). While Gouriéroux and Zakoian (2017) focus on continuous support distributions for the aggregation weights in the specific case of anticipative Cauchy AR(1) processes, our approach extends to the general α-stable case with discrete support, a framework more suitable for empirical applications. We propose a deconvolution minimum distance estimator based on the joint characteristic function that effectively identifies the model parameters. Our methodology draws from Knight and Yu (2002) and Xu and Knight (2010), who developed asymptotic theory for minimum distance estimation using the empirical characteristic function in stationary time series, but we extend their approach to handle the heavy-tailed stable distributions. We establish the asymptotic properties of our estimator under suitable regularity conditions, proving consistency and asymptotic normality. To empirically validate the finite-sample convergence toward the limiting Gaussian distribution, we implement a subsampling procedure following Politis and Romano (1994) and Politis, Romano and Wolf (1999), which reveals heterogeneous convergence speeds across parameter dimensions and confirms that while certain parameters approach asymptotic normality relatively quickly, others, particularly the autoregressive coefficients, require substantially larger sample sizes to achieve reliable normal approximations.

As an empirical illustration, we estimate an aggregation of purely anticipative stable AR(1) processes using the CBOE Crude Oil ETF Volatility Index (OVX) data, and we demonstrate that the observed volatility patterns can be effectively decomposed into multiple latent stable components with heterogeneous persistence properties. The empirical analysis reveals that what initially appears as a single explosive episode actually consists of several superimposed processes with distinct autoregressive parameters and crash probabilities.

The remainder of this paper is organized as follows. Section 2 introduces the stable aggregates model and develops a new minimum distance estimator based on the characteristic functions of the unobserved latent components. Section 3 extends the representation theorem of de Truchis et al. (2025a) to stable aggregates and theoretically derives the conditions under which the forecast of a stable aggregate is possible. Section 4 documents the finite sample performance of the minimum distance estimator through Monte Carlo simulations and implements a subsampling methodology to empirically verify the asymptotic normality of the estimator. An application to the CBOE Crude Oil ETF Volatility Index is proposed in Section 5. Section

6 concludes. All proofs are provided in the Appendix A, while subsampling diagnostics and complementary convergence results are collected in the Online Supplement.

2. Estimating stable-aggregate of moving average

Consider X_t an α-stable moving average defined by

$$X_t = \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k}, \qquad \varepsilon_t \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta, \sigma, 0) \tag{2.1}$$

with $d_0 > 0$, (d_k) a real deterministic sequence such that if $\alpha \neq 1$ or $(\alpha, \beta) = (1, 0)$,

$$\sum_{k\in\mathbb{Z}} |d_k|^s < +\infty, \quad \text{for some} \quad s \in (0, \alpha) \cap [0, 1], \tag{2.2}$$

and if $\alpha = 1$ and $\beta \neq 0$,

$$0 < \sum_{k\in\mathbb{Z}} |d_k| \Big| \ln |d_k| \Big| < +\infty. \tag{2.3}$$

For $d_k = \rho^k$, X_t is a simple strictly stationary anticipative AR(1). For X_t the strictly stationary solution of $\Psi(F)\Phi(B)X_t = \Theta(F)H(B)\varepsilon_t$, with F and B the lead and lag operators, the process belongs to the class of mixed-phase ARMA. Furthermore, if $\Theta = H = 1$, X_t is called a mixed-causal or MAR(p, q) process, where $p = \deg(\Phi)$ and $q = \deg(\Psi)$. Adding the $(\alpha, \beta) = (1, 0)$ restrictions (let say $\mathcal{S}1\mathcal{S}$), X_t actually comes down to the so-called anticipative Cauchy AR(1) studied, e.g., in Gouriéroux and Jasiak (2018). As emphasized in the introduction, stable moving averages of the form (2.1) generate trajectories bound to feature the same pattern $t \mapsto c d_{\tau - t}$ (up to a scaling c and a time shift τ) recurrently through time. This can be seen as a strong limitation when it comes to time series modelling as argued by Gouriéroux and Zakoian (2017) in the context of explosive bubbles. They suggest to alleviate this restriction by considering processes resulting from the linear combination of different models.

Definition 2.1. *Let $(X_{1,t}), \dots, (X_{J,t})$ be $J \geq 1$ stable moving averages, each satisfying (2.1)-(2.3), for some distinct coefficients sequences $(d_{j,k})_k$ and mutually independent error sequences $\varepsilon_{j,t} \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta_j, 1, 0)$, $j = 1, \dots, J$. Let also $(\pi_j)_{j=1,\dots,J}$ be positive numbers summing to 1, $\sigma > 0$ be a scale parameter and define*

$$\mathcal{X}_t = \sigma \sum_{j=1}^{J} \pi_j X_{j,t}, \qquad \textit{for} \quad t \in \mathbb{Z}.$$

We will call such process $\mathcal{X}_t$ a stable aggregate, and call $X_{j,t}$, $j = 1, \dots, J$ the latent components of $\mathcal{X}_t$.

The estimator we propose is valid for any strictly stationary stable aggregate satisfying Definition 2.1, but in practice, it requires to formally derive the characteristic function of the latent components which can be tedious. In the rest of this section, we focus on $\beta_j = \beta$ for simplicity and provide the derivation

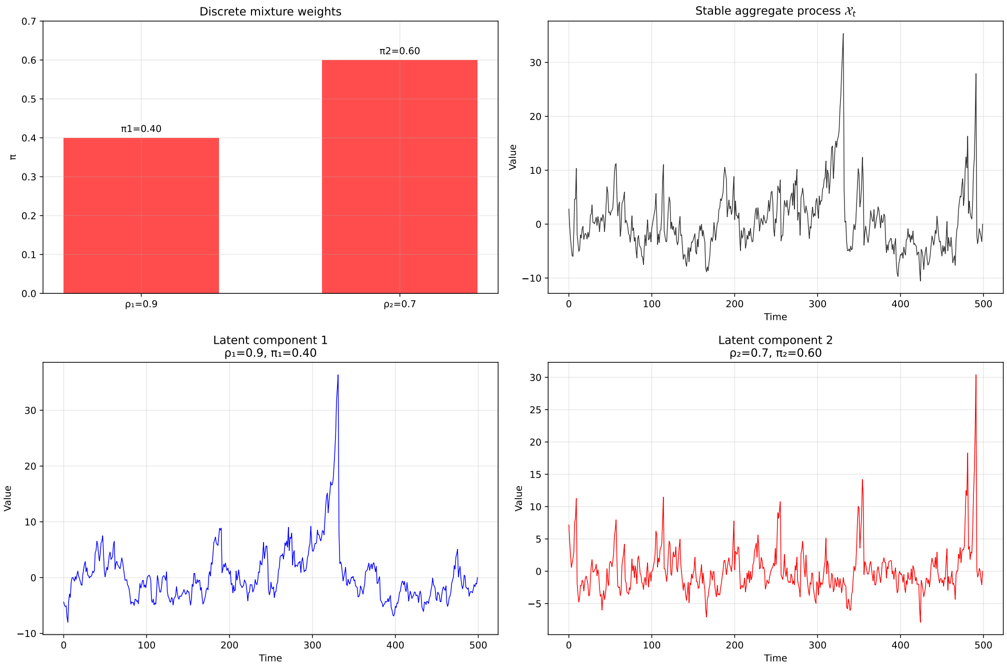

Figure 1: Simulated stable aggregate dynamics with two components. Top left: Distribution of weights for the two components with $\rho_1 = 0.90$, $\pi_1 = 0.40$ for the first component and $\rho_2 = 0.70$, $\pi_2 = 0.60$ for the second component. Top right: The resulting trajectory of the aggregated process $\mathcal{X}_t$. Middle and bottom panels: The individual latent component processes with different persistence parameters.

for two important parametric cases: the aggregation of purely anticipative AR(1) or MAR(0,1) processes and the aggregation of mixed causal-noncausal MAR(1,1) processes. Notice that even in these specific frameworks, these aggregations feature much richer dynamics than single-component stable processes, as illustrated in Figure 1. To disentangle the components of $\mathcal{X}_t$, our method leverages the independence of the latent processes and the resulting structure of the joint characteristic function:

$$\varphi_{\mathcal{X}}(u,v) = \mathbb{E}\Big(\exp\{i(u\mathcal{X}_t + v\mathcal{X}_{t+1})\}\Big) = \prod_{j=1}^{J} \varphi_{X_j}(\sigma\pi_j u, \sigma\pi_j v) \tag{2.4}$$

where φ_{X_j} is the joint characteristic function of a single latent component.

2.1. Case 1: Aggregation of Anticipative AR(1) Processes

We first restrict our attention to the case where each latent component $X_{j,t}$ is a purely anticipative AR(1) process. Its moving average representation is given by $d_{j,k} = \rho_j^k \mathbf{1}_{k\geq 0}$, with $0 < \rho_j < 1$. This restriction ensures that the asymmetry parameter β is preserved through the infinite summation defining each latent component $X_{j,t}$. From an economic perspective, positive autoregressive coefficients correspond to monotonic bubble growth patterns without oscillations, which is the empirically relevant case for financial applications modeling speculative bubbles.[4] The process is thus defined by $X_{j,t} = \sum_{k=0}^{\infty} \rho_j^k \varepsilon_{j,t+k}$. The joint characteristic function of the vector $(X_{j,t}, X_{j,t+1})$ is given by

$$\varphi_{X_j}(u,v) = \mathbb{E}\Big(\exp i(uX_{j,t} + vX_{j,t+1})\Big) = \mathbb{E}\Big(\exp i\big((u\rho_j + v)X_{j,t+1} + u\varepsilon_{j,t}\big)\Big), \tag{2.5}$$

for $(u,v) \in \mathbb{R}^2$. Due to the independence of the innovations, this simplifies to

$$\varphi_{X_j}(u,v) = \mathbb{E}\Big(\exp i(u\rho_j + v)X_{j,t+1}\Big)\mathbb{E}\Big(\exp iu\varepsilon_{j,t}\Big),$$

Assuming for simplicity a common asymmetry parameter $\beta_j = \beta$, we have for $\alpha \neq 1$

$$\log \mathbb{E}\Big(\exp i(u\rho_j + v)X_{j,t+1}\Big) = -\frac{|u\rho_j + v|^\alpha}{1 - |\rho_j|^\alpha}\left(1 - i\beta \operatorname{sign}(u\rho_j + v)\tan\left(\frac{\pi\alpha}{2}\right)\right)$$
$$\log \mathbb{E}(\exp iu\varepsilon_{j,t}) = -\left(1 - i\beta \operatorname{sign}(u)\tan\left(\frac{\pi\alpha}{2}\right)\right)|u|^\alpha.$$

The log-characteristic function of the aggregate is then obtained by substituting these expressions into Equation (2.4)

$$\log \varphi_{\mathcal{X}}(u,v) = -\sigma^\alpha \sum_{j=1}^{J} \pi_j^\alpha \left(\frac{|u\rho_j + v|^\alpha}{1 - |\rho_j|^\alpha}\left(1 - i\beta \operatorname{sign}(u\rho_j + v)\tan\frac{\pi\alpha}{2}\right) + |u|^\alpha\left(1 - i\beta \operatorname{sign}(u)\tan\frac{\pi\alpha}{2}\right)\right).$$

The Cauchy case examined in Gouriéroux and Zakoian (2017) is recovered for $\alpha = 1$, $\beta = 0$, leading to $\log \mathbb{E}(\exp iu\varepsilon_{j,t}) = -|u|$ and

$$\log \varphi_{\mathcal{X}}(u,v) = -\sigma \sum_{j=1}^{J} \pi_j \left(\frac{|u\rho_j + v|}{1 - |\rho_j|} + |u|\right).$$

As each latent component satisfies $|\rho_j| < 1$, the strict stationarity condition for $\mathcal{X}_t$ is given by

$$\sum_{j=1}^{J} \frac{\pi_j^s}{1 - |\rho_j|^s} < \infty \quad \text{for } s \in (0,\alpha) \cap [0,1]. \tag{2.6}$$

2.2. Case 2: Aggregation of Mixed Causal-Noncausal MAR(1,1) Processes

We now consider a richer dynamic structure where each latent component $X_{j,t}$ is a mixed causal-noncausal MAR(1,1) process defined by $(1-\phi_j L)(1-\psi_j L^{-1})X_{j,t} = \varepsilon_{j,t}$, with $|\phi_j| < 1$ and $|\psi_j| < 1$. The corresponding MA(∞) coefficients are given by $\psi_j^k(1-\phi_j\psi_j)^{-1}$ if $k \geq 0$ and $\phi_j^{|k|}(1-\phi_j\psi_j)^{-1}$ for $k < 0$. The log-characteristic function for a single component X_j is derived from the linear combination of innovations

$$uX_{j,t} + vX_{j,t+1} = \sum_{k=-\infty}^{\infty} (ud_{j,k} + vd_{j,k-1})\varepsilon_{j,t+k}.$$

[4] To see why this matters, recall that for a sum $\sum_{k=0}^{\infty} c_k u_k$ with $u_k \overset{i.i.d.}{\sim} \mathcal{S}(\alpha,\beta,\sigma,0)$, the resulting distribution is $\mathcal{S}(\alpha,\beta',\sigma',0)$ where $\beta' = \frac{\sum_{k=0}^{\infty} |c_k|^\alpha \operatorname{sign}(c_k)}{\sum_{k=0}^{\infty} |c_k|^\alpha} \cdot \beta$. When $\rho_j > 0$, all coefficients $c_k = \rho_j^k$ are positive, yielding $\beta' = \beta$. However, when $\rho_j < 0$, the coefficients alternate in sign, leading to $\beta' \neq \beta$. The case $\rho_j < 0$ would thus require a component-specific modified asymmetry parameter β_j' in the characteristic function.

In the symmetric ($\mathcal{S}\alpha\mathcal{S}$) case, the log-characteristic function is

$$\log \varphi_{X_j}(u,v) = -\sum_{k=-\infty}^{\infty} |ud_{j,k} + vd_{j,k-1}|^{\alpha}.$$

We split the sum into its causal ($k \leq 0$) and non-causal ($k \geq 1$) parts. For the causal part ($k \leq 0$), the generic term is $ud_{j,k} + vd_{j,k-1} = (1-\phi_j\psi_j)^{-1}(u\phi_j^{|k|} + v\phi_j^{|k-1|}) = (u+v\phi_j)(1-\phi_j\psi_j)^{-1}\phi_j^{|k|}$. For the non-causal part ($k \geq 1$), the generic term is $ud_{j,k} + vd_{j,k-1} = (1-\phi_j\psi_j)^{-1}(u\psi_j^{k} + v\psi_j^{k-1}) = (u\psi_j + v)(1-\phi_j\psi_j)^{-1}\psi_j^{k-1}$. The sum becomes the sum of two geometric series

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} |ud_{j,k} + vd_{j,k-1}|^{\alpha} &= \frac{1}{|1-\phi_j\psi_j|^{\alpha}} \left(\sum_{k=-\infty}^{0} |(u+v\phi_j)\phi_j^{|k|}|^{\alpha} + \sum_{k=1}^{\infty} |(u\psi_j + v)\psi_j^{k-1}|^{\alpha} \right) \\
&= \frac{1}{|1-\phi_j\psi_j|^{\alpha}} \left(|u+v\phi_j|^{\alpha} \sum_{l=0}^{\infty} (|\phi_j|^{\alpha})^{l} + |u\psi_j + v|^{\alpha} \sum_{l=0}^{\infty} (|\psi_j|^{\alpha})^{l} \right) \\
&= \frac{1}{|1-\phi_j\psi_j|^{\alpha}} \left(\frac{|u+v\phi_j|^{\alpha}}{1-|\phi_j|^{\alpha}} + \frac{|u\psi_j+v|^{\alpha}}{1-|\psi_j|^{\alpha}} \right).
\end{aligned}$$

Finally, substituting this result into the aggregate function from Equation (2.4), we obtain the log-characteristic function for the MAR(1,1) aggregate for the $\mathcal{S}\alpha\mathcal{S}$ case . For the asymmetric case with $\alpha \neq 1$ we impose, for simplicity but without loss of generality, that all components satisfy $\phi_j > 0$ and $\psi_j > 0$ and we obtain

$$\log \varphi_{\mathcal{X}}(u,v) = -\sigma^{\alpha} \sum_{j=1}^{J} \frac{\pi_j^{\alpha}}{|1-\phi_j\psi_j|^{\alpha}} \left(\mathcal{C}_j(u,v) + \mathcal{A}_j(u,v)\right) \tag{2.7}$$

where $\mathcal{C}_j(u,v)$ and $\mathcal{A}_j(u,v)$ represent the complex-valued contributions from the causal and non-causal dynamics of each component j, respectively

$$\begin{aligned}
\mathcal{C}_j(u,v) &= \frac{|u+v\phi_j|^{\alpha}}{1-|\phi_j|^{\alpha}} \left(1 - i\beta\, \mathrm{sign}(u+v\phi_j) \tan\left(\frac{\pi\alpha}{2}\right)\right), \\
\mathcal{A}_j(u,v) &= \frac{|u\psi_j+v|^{\alpha}}{1-|\psi_j|^{\alpha}} \left(1 - i\beta\, \mathrm{sign}(u\psi_j+v) \tan\left(\frac{\pi\alpha}{2}\right)\right).
\end{aligned}$$

The strict stationarity condition for $\mathcal{X}_t$ is now given by

$$\sum_{j=1}^{J} \frac{\pi_j^{s}}{|1-\phi_j\psi_j|^{s}} \left(\frac{1}{1-|\psi_j|^{s}} + \frac{|\phi_j|^{s}}{1-|\phi_j|^{s}} \right) < \infty \quad \text{for } s \in (0,\alpha) \cap [0,1]. \tag{2.8}$$

2.3. The minimum distance estimator

As suggested by Knight and Yu (2002) and Gouriéroux and Zakoian (2017), one can rely on the empirical counterpart of the joint characteristic function (ECF) to build a minimum distance estimator (MDE). The ECF is simply defined as

$$\varphi_n(u,v) = \frac{1}{n-1} \sum_{j=1}^{n-1} \exp(i(u\mathcal{X}_{j+1} + v\mathcal{X}_j)) \tag{2.9}$$

which can be decomposed into real and imaginary parts:

$$\varphi_n(u,v) = \frac{1}{n-1}\sum_{j=1}^{n-1}\cos(u\mathcal{X}_{j+1}+v\mathcal{X}_j) + \frac{1}{n-1}\sum_{j=1}^{n-1} i\sin(u\mathcal{X}_{j+1}+v\mathcal{X}_j) \tag{2.10}$$

By the law of large numbers, $\varphi_n(u,v) \xrightarrow{P} \varphi(u,v;\theta_0)$ as $n \to \infty$, where θ_0 denotes the true parameter values. Then, the identification of the parameters $\theta = (\sigma, \rho_1, \ldots, \rho_J, \pi_1, \ldots, \pi_J, \alpha, \beta)$ relies on distinct asymptotic behaviors of the joint characteristic function for different values of (u, v). For small values of u, the limit behavior of (2.9) is dominated by the α-stable distribution's properties. Specifically, for $u > 0$,

$$\alpha = \lim_{u\to 0} \frac{\log\log|\varphi_n(u,0)|^{-1}}{\log|u|} \tag{2.11}$$

and

$$\beta = -\lim_{u\to 0} \frac{\operatorname{Im}(\log\varphi_n(u,0))}{\operatorname{Re}(\log\varphi_n(u,0))} \cdot \cot\frac{\pi\alpha}{2}. \tag{2.12}$$

For the identification of the remaining parameters, we exploit the behavior of the function

$$g_n(\lambda) = \lim_{u\to 0} \frac{\log|\varphi_n(u,\lambda u)|}{|u|^\alpha} \approx -\sigma^\alpha \sum_{j=1}^{J} \pi_j^\alpha \left(\frac{|1+\rho_k\lambda|^\alpha}{1-|\rho_j|^\alpha} + |\lambda|^\alpha \right) \tag{2.13}$$

for $v = \lambda u$ and $\lambda \in \mathbb{R}$. By evaluating $g_n(\lambda)$ for $2J+1$ different values of λ, we can obtain a system of equations to identify $(\sigma, \rho_1, \ldots, \rho_J, \pi_1, \ldots, \pi_J)$.

Now we can define the MDE estimator as the minimizer of the objective distance measure

$$D_{\mathcal{X}}(\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |\varphi_n(u,v) - \varphi(u,v;\theta)|^2 w(u,v)\,du\,dv \tag{2.14}$$

where $w(u,v)$ is a weighting function ensuring the convergence of the integral. The MDE estimator is then defined as

$$\hat{\theta}_n = \arg\min_\theta D_{\mathcal{X}}(\theta). \tag{2.15}$$

Knight and Yu (2002), show that under the following regularity conditions, the MDE estimator has standard limit theory. They suggest that it could accommodate α-stable models. Actually, some of their assumptions, listed hereafter, does not readily extend to the α-stable case. The characteristic functions of α-stable distributions are likely to exhibit singularities in their derivatives when $\alpha \in (0,2)$, particularly near points where $|\rho_j u + v|^\alpha$ vanishes. Without appropriate regularization through the weight function, these singularities can cause the integrals defining the first and second derivatives of (2.14) to diverge. The following lemma establishes the precise conditions under which their regularity assumptions remains valid for α-stable aggregates.

Lemma 2.1. *Consider the MDE objective function defined by*

$$D_{\mathcal{X}}(\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |\varphi_n(u,v) - \varphi(u,v;\theta)|^2 w(u,v)\, du\, dv \tag{2.16}$$

where $w(u,v) = \exp(-\kappa(u^2+v^2))$ *with* $\kappa > 0$ *a positive constant.*

Then,

(ι) *For any* $\alpha > 0$*, the objective function* $D_{\mathcal{X}}(\theta)$ *belongs to the differentiability class* $C^1(\Theta)$*.*

($\iota\iota$) *For any* $\alpha > 1$*, the objective function* $D_{\mathcal{X}}(\theta)$ *belongs to* $C^2(\Theta)$ *and Assumption 3, 6, 7 and 8 are satisfied.*

Lemma 2.1, shows that we need to reduce the parameter space of α by introducing Assumption 2, in addition to whole set of assumptions of Knight and Yu (2002), to recover their asymptotic theory in presence of α-stable models. It also reveals the critical role of the decaying exponential weights $w(u,v)$. Assumption 4 is satisfied under the condition given by (2.6) or (2.8) and Assumption 5 is satisfied by the global identification conditions exposed in (2.11), (2.12) and (2.13). The proof of Lemma 2.1 is postponed in Section A.

Assumption 1. $\theta \in \Theta$ *where the parameter space* $\Theta \subset \mathbb{R}^{2J+3}$ *is a compact set with* $\theta_0 \in Int(\Theta)$*.*

Assumption 2. *The tail parameter space is such that* $\alpha \in (1,2)$ *and* $w(u,v)$ *is an exponential weight function of form* $\exp(-\kappa(u^2+v^2))$ *with* $\kappa > 0$ *a positive constant.*

Assumption 3. *With probability one,* $D_{\mathcal{X}}(\theta)$ *is twice continuously differentiable under the integral sign with respect to* θ *over* Θ*.*

Assumption 4. *The sequence* $\{\mathcal{X}_t\}$ *is strictly stationary and ergodic.*

Assumption 5. *Let* $D_0(\theta) = \iint |\varphi(u,v;\theta_0) - \varphi(u,v;\theta)|^2 w(u,v) du dv$ *and* $D_0(\theta) = 0$ *only if* $\theta = \theta_0$*.*

Assumption 6. $K(x;\theta)$ *is a measurable function of* x *for all* θ *and bounded, where*

$$\begin{aligned} K(x;\theta) = \iint \Bigg[&(\cos(ux_{j+1} + vx_j) - Re\ \varphi(u,v;\theta))\frac{\partial Re\ \varphi(u,v;\theta)}{\partial\theta} \\ &+(\sin(ux_{j+1} + vx_j) - Im\ \varphi(u,v;\theta))\frac{\partial Im\ \varphi(u,v;\theta)}{\partial\theta}\Bigg] w(u,v) du dv. \end{aligned} \tag{2.17}$$

Assumption 7. *The* $(2J+3)\times(2J+3)$ *matrix*

$$\Sigma(\theta_0) = \iint \left(\frac{\partial\varphi(u,v;\theta_0)}{\partial\theta}\right)\left(\frac{\partial\overline{\varphi}(u,v;\theta_0)}{\partial\theta'}\right) w(u,v) du dv$$

is nonsingular and

$$\frac{\partial^2\varphi(u,v;\theta)}{\partial\theta\partial\theta'}$$

is uniformly bounded by a w*-integrable function over* Θ*.*

Assumption 8. *Let $\mathcal{F}_j$ be a σ-algebra such that $\{K_j, \mathcal{F}_j\}$ is an adapted stochastic sequence, where $K_j = K(x_j;\theta)$. We can think of $\mathcal{F}_j$ as being the σ-algebra generated by the entire current and past history of K_j. Let $\nu_j = \mathbb{E}[K_0|K_j, K_{j-1}, \dots] - \mathbb{E}[K_0|K_{j-1}, K_{j-2}, \dots]$ for $j \geq 0$. Assume that $\mathbb{E}(K_0|\mathcal{F}_{-m})$ converges in mean square to 0 as $m \to \infty$ and $\sum_{j=0}^{\infty} \mathbb{E}[\nu_j' \nu_j]^{1/2} < \infty$.*

Proposition 2.1. *Under Assumptions 1-8*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N\big(0, \Sigma(\theta_0)^{-1}\Omega(\theta_0)\Sigma(\theta_0)^{-1}\big) \tag{2.18}$$

where $\Sigma(\theta_0)$ is defined in Assumption 7, and $\Omega(\theta_0)$ is the long-run variance matrix of the score function $K(x;\theta_0)$ from Assumption 6

$$\Omega(\theta_0) = \mathbb{V}\big(K(x_1;\theta_0)\big) + 2\sum_{j=2}^{\infty} Cov\big(K(x_1;\theta_0), K(x_j;\theta_0)\big)$$

The proof of this theorem is omitted as, by Lemma 2.1, it follows from a straightforward extension of Theorem 2.1 of Knight and Yu (2002). Notice that in our α-stable framework, unlike Xu and Knight (2010), $\Sigma(\theta_0)$ and $\Omega(\theta_0)$ have no closed-form solutions. Moreover, to alleviate the optimization problem from a numerical standpoint, we directly estimate the products $\varsigma_j = \sigma \times \pi_j$ for $j = 1, \dots, J$.

2.4. Case 3: Aggregation of Mixed Stable and Gaussian Processes

Our estimation framework can also be extended to accommodate aggregates mixing α-stable and Gaussian components, an approach explored in Gouriéroux and Zakoian (2017) and Gouriéroux et al. (2021) but only for the Cauchy case. Consider a process $\mathcal{X}_t$ resulting from the aggregation of an α-stable MAR$(p,1)$, $p \in \{0,1\}$ with $\alpha \in (1,2)$ and a Gaussian AR(1) component $X_{\mathcal{N},t}$. As the distinction between causal and non-causal dynamics is unidentifiable when $\alpha = 2$, we adopt the standard causal specification for the Gaussian component. The log-characteristic function of the Gaussian AR(1) component $X_{\mathcal{N},t} = \phi_{\mathcal{N}} X_{\mathcal{N},t-1} + \eta_t$, $\eta_t \sim \mathcal{N}(0,1)$, for the vector $(X_{\mathcal{N},t}, X_{\mathcal{N},t-1})$ is given by

$$\log \varphi_{\mathcal{N}}(u,v) = -\frac{1}{2}\frac{1}{1-\phi_{\mathcal{N}}^2}(u\phi_{\mathcal{N}} + v)^2 - \frac{u^2}{2}$$

The resulting aggregate log-characteristic function, $\log \varphi_{\mathcal{X}}(u,v)$, is the sum of the stable component's characteristic functions $\log \varphi_{X_j}(u,v)$ and $\log \varphi_{\mathcal{N}}(u,v)$, scaled by their respective aggregation weights as in Equation (2.4). This composite function can be directly employed in the MDE objective function (2.14). The estimator $\hat{\theta}_n$ defined in (2.15) remains valid because the stability index $\alpha = 2$ for the Gaussian component is fixed and not estimated. Since $\log \varphi_{\mathcal{N}}(u,v)$ is C^{∞} with respect to its parameters, and $\log \varphi_{\mathcal{X}}(u,v)$ is C^2 for $\alpha \in (1,2)$ (as established in Lemma 2.1), their sum remains C^2. The regularity conditions required for the asymptotic theory of the MDE estimator (Proposition 2.1) are thus satisfied, allowing for the joint identification of the parameters of both the stable and Gaussian latent processes.

3. Forecasting aggregation of moving averages

This section begins by summarizing relevant findings from de Truchis et al. (2025a), DFT henceforth, concerning the description of stable random vectors on the unit cylinder.[5] Let the vector $\boldsymbol{X} = (X_1, \dots, X_d)$ be an α-stable random vector, Γ a finite spectral measure on the Euclidean unit sphere S_d and $\boldsymbol{\mu}^0$ a non-random vector in $\mathbb{R}^d$, such that,

$$\mathbb{E}\Big(e^{i\langle \boldsymbol{u},\boldsymbol{X}\rangle}\Big) = \exp\Bigg\{ - \int_{S_d} |\langle \boldsymbol{u}, \boldsymbol{s}\rangle|^\alpha \Big(1 - i\,\mathrm{sign}(\langle \boldsymbol{u}, \boldsymbol{s}\rangle) w(\alpha, \langle \boldsymbol{u}, \boldsymbol{s}\rangle)\Big)\Gamma(d\boldsymbol{s}) + i\,\langle \boldsymbol{u}, \boldsymbol{\mu}^0\rangle \Bigg\}, \quad \forall \boldsymbol{u} \in \mathbb{R}^d, \tag{3.1}$$

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 $s \in \mathbb{R}$. Drawing on DFT, we explore alternative representations of $\boldsymbol{X}$ where the integration is performed over a unit cylinder $C_d^{\|\cdot\|} := \{\boldsymbol{s} \in \mathbb{R}^d : \ \|\boldsymbol{s}\| = 1\}$, defined by a semi-norm $\|\cdot\|$ on $\mathbb{R}^d$, in presence of stable aggregates. The reason why we are interested in alternative representations is that, in the presence of the Euclidean norm, the spectral measure encodes information in all directions of $\mathbb{R}^d$ and does not allow us to predict future elements of the vector $\boldsymbol{X}$ while ensuring that these future elements are not themselves carriers of information for prediction. By contrast, the semi-norm $\|\cdot\|$ is flexible enough to force some directions $\mathbb{R}^d$ to vanish.

We will say that $\boldsymbol{X}$ is representable on $C_d^{\|\cdot\|}$ if $\boldsymbol{X}$ can be written as in (3.1) with $(S_d, \Gamma, \boldsymbol{\mu}^0)$ replaced by $(C_d^{\|\cdot\|}, \Gamma^{\|\cdot\|}, \boldsymbol{\mu}^0_{\|\cdot\|})$. As demonstrated in DFT for the single-component model, $\boldsymbol{X}$ is representable on $C_d^{\|\cdot\|} \iff \Gamma(K^{\|\cdot\|}) = 0$ when $\alpha \neq 1$ or if $\boldsymbol{X}$ is $\mathcal{S}1\mathcal{S}$. Moreover, $\Gamma^{\|\cdot\|}(d\boldsymbol{s}) = \|\boldsymbol{s}\|_e^{-\alpha}\,\Gamma \circ T^{-1}_{\|\cdot\|}(d\boldsymbol{s})$ with $T_{\|\cdot\|} : S_d \setminus K^{\|\cdot\|} \longrightarrow C_d^{\|\cdot\|}$ defined by $T_{\|\cdot\|}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|$. Importantly, this new representation inherits from the traditional representation the following asymptotic conditional tail property: for any Borel sets $A, B \subset C_d^{\|\cdot\|}$ with $\Gamma^{\|\cdot\|}\Big(\partial(A \cap B)\Big) = \Gamma^{\|\cdot\|}(\partial B) = 0$, and $\Gamma^{\|\cdot\|}(B) > 0$,

$$\mathbb{P}_x^{\|\cdot\|}(\boldsymbol{X}, A|B) \underset{x\to+\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}(A \cap B)}{\Gamma^{\|\cdot\|}(B)}, \tag{3.2}$$

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

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

To build a forecasting strategy upon these theoretical results, DFT considers vectors of the form $\boldsymbol{X_t} = (X_{t-m}, \dots, X_t, X_{t+1}, \dots, X_{t+h})$, $m \geq 0$, $h \geq 1$, derived from a stable moving average process and choose, without loss of generality, semi-norms satisfying

$$\|(x_{-m}, \dots, x_0, x_1, \dots, x_h)\| = 0 \iff x_{-m} = \dots = x_0 = 0, \tag{3.3}$$

[5] We exclude the Gaussian case from further discussion as anticipative dynamics are not identifiable when $\alpha = 2$.

for any $(x_{-m}, \dots, x_h) \in \mathbb{R}^{m+h+1}$. They show that for $\alpha \neq 1$ 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. More precisely, they find that sequences of consecutive zero values in must either be of finite length or extend infinitely to the left :

$$\forall k \in \mathbb{Z}, \quad \Big[(d_{k+m}, \dots, d_k) = \mathbf{0} \quad \Longrightarrow \quad \forall \ell \leq k-1, \quad d_\ell = 0\Big]. \tag{3.4}$$

This result surprisingly establishes that the anticipativeness of a stable moving average is a necessary condition (and sufficient for $\alpha \neq 1$ and $(\alpha, \beta) = (1, 0)$) to make use of (3.2) in order to feasibly predict $\boldsymbol{X_t}$. The more non-anticipative a moving average is (i.e., the larger the gaps of zeros in its forward-looking coefficients), the larger m must be to achieve representability of $(X_{t-m}, \dots, X_t, X_{t+1}, \dots, X_{t+h})$ on the appropriate unit cylinder.

3.1. Extending the representation to stable aggregates

To extend these results to stable aggregates, we first provide the spectral representation of paths of the aggregated process $\mathcal{X}_t$ on the Euclidean unit sphere.

Lemma 3.1. *Let $\mathcal{X}_t$ be an α-stable aggregate with latent moving averages $(X_{1,t}), \dots, (X_{J,t})$ as in Definition 2.1, but now allowing $\beta_j \in [-1, 1]$ to vary across components, and $\boldsymbol{X_t} = (\mathcal{X}_{t-m}, \dots, \mathcal{X}_t, \mathcal{X}_{t+1}, \dots, \mathcal{X}_{t+h})$ for any $m \geq 0$, $h \geq 1$.*

Then, $\boldsymbol{X}_t$ is α-stable and its spectral representation $(\Gamma, \boldsymbol{\mu}^0)$ on the Euclidean unit sphere S_{m+h+1} writes

$$\Gamma = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|_e^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|_e}\right\}}, \tag{3.5}$$

$$\boldsymbol{\mu}^0 = \begin{cases} \mathbf{0}, & \text{if } \alpha \neq 1 \\ -\frac{2\sigma}{\pi} \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} \pi_j \beta_j \boldsymbol{d}_{j,k} \ln \|\sigma \pi_j \boldsymbol{d}_{j,k}\|_e, & \text{if } \alpha = 1 \end{cases}$$

where $\boldsymbol{d}_{j,k} = (d_{j,k+m}, \dots, d_{j,k}, d_{j,k-1}, \dots, d_{j,k-h})$, $w_{j,\vartheta} = (1 + \vartheta\beta_j)/2$, for any $k \in \mathbb{Z}$, $j = 1, \dots, J$, δ is the Dirac mass, $\vartheta \in S_1$ with $S_1 = \{-1, +1\}$, and if $\boldsymbol{d}_{j,k} = \mathbf{0}$, the term vanishes by convention from the sums.

Notice that $\Gamma = \sigma^\alpha \sum_{j=1}^{J} \pi_j^\alpha \Gamma_j$, where Γ_j denotes the spectral measure of the path $\boldsymbol{X}_{j,t}$ from the moving average $(X_{j,t})$, $j = 1, \dots, J$. If all the $\boldsymbol{X}_{j,t}$'s are symmetric ($\beta_j = 0$ for all j), then $\boldsymbol{X}_t$ and Γ are symmetric as well, but the reciprocal however does not hold true. The measure Γ will be symmetric if and only if $\sigma^\alpha \sum_{j=1}^{J} \pi_j^\alpha \Big(\Gamma_j(A) - \Gamma_j(-A)\Big) = 0$ for any Borel set $A \subset S_{m+h+1}$. The latter condition is necessary and sufficient for $\boldsymbol{X}_t$ to be symmetric in the case where $\alpha \neq 1$, whereas for $\alpha = 1$, it guarantees that $\boldsymbol{X}_t$ will be symmetric up to an additive shifting, as $\boldsymbol{\mu}^0$ may be non-zero. The symmetry of paths intervenes in the representability conditions provided in the following lemma.

Lemma 3.2. *Let $\mathcal{X}_t$ be an α-stable aggregate with latent moving averages $(X_{1,t}), \ldots, (X_{J,t})$ as in Definition 2.1, where each component j has asymmetry parameter $\beta_j \in [-1, 1]$. Let $m \geq 0$, $h \geq 1$ and $\|\cdot\|$ be a semi-norm on $\mathbb{R}^{m+h+1}$ satisfying (3.3). When either $\alpha \neq 1$ or $\boldsymbol{X}_t \sim \mathcal{S}1\mathcal{S}$, the vector $\boldsymbol{X}_t$ is representable on $C^{\|\cdot\|}_{m+h+1}$ if and only if condition (3.4) holds with m for all coefficient sequences $(d_{j,k})_k$, $j = 1, \ldots, J$. For $\alpha = 1$ and $\boldsymbol{X}_t$ asymmetric, the vector $\boldsymbol{X}_t$ is representable on $C^{\|\cdot\|}_{m+h+1}$ if and only if (3.4) holds and*

$$\sum_{k \in \mathbb{Z}} \|\boldsymbol{d}_{j,k}\|_e \left| \ln \left(\|\boldsymbol{d}_{j,k}\| / \|\boldsymbol{d}_{j,k}\|_e \right) \right| < +\infty, \quad \forall j \in \{1, \ldots, J\} \tag{3.6}$$

hold with m and h for all sequences $(d_{j,k})_k$, $j = 1, \ldots, J$.

The next proposition extends to stable aggregated processes the notion of past-representability introduced in DFT and helps to understand to what extent anticipativeness is crucial in this more general framework.

Proposition 3.1. *Let $\mathcal{X}_t$ be an α-stable aggregate with latent moving averages $(X_{1,t}), \ldots, (X_{J,t})$ as in Definition 2.1, where $\mathcal{X}_t = \sigma \sum_{j=1}^{J} \pi_j X_{j,t}$ with scale parameter $\sigma > 0$.*

(ι) Define for $j = 1, \ldots, J$ the sets $\mathcal{M}_j = \{m \geq 1 : \ \exists k \in \mathbb{Z}, \ d_{j,k+m} = \ldots = d_{j,k+1} = 0, \ d_{j,k} \neq 0\}$, and

$$m_{0,j} = \begin{cases} \sup \mathcal{M}_j, & \textit{if} \quad \mathcal{M}_j \neq \emptyset, \\ 0, & \textit{if} \quad \mathcal{M}_j = \emptyset. \end{cases} \tag{3.7}$$

(a) For $\alpha \neq 1$, the aggregated process $\mathcal{X}_t$ is past-representable if and only if $(X_{j,t})$ is past-representable for all $j = 1, \ldots, J$, i.e.,

$$\sup_{j=1,\ldots,J} m_{0,j} < +\infty. \tag{3.8}$$

Moreover, letting $m \geq 0$, $h \geq 1$, $\mathcal{X}_t$ is (m,h)-past-representable if and only if (3.8) holds and $m \geq \max_{j=1,\ldots,J} m_{0,j}$.

(b) For $\alpha = 1$, the process $\mathcal{X}_t$ is past-representable if and only if (3.8) holds and there exists a pair (m,h), $m \geq \max_{j=1,\ldots,J} m_{0,j}$, $h \geq 1$ such that either

$$\boldsymbol{X}_t \textit{ is } \mathcal{S}1\mathcal{S}, \qquad \textit{or,} \qquad \boldsymbol{X}_t \textit{ asymmetric and (3.6) holds for all sequences } (d_{j,k})_k.$$

If such a pair exists, then the process $\mathcal{X}_t$ is (m,h)-past-representable.

$(\iota\iota)$ Let $\|\cdot\|$ be a semi-norm satisfying (3.3) and assume that $\mathcal{X}_t$ is (m,h)-past-representable for some $m \geq 0$, $h \geq 1$. The spectral representation $(\Gamma^{\|\cdot\|}, \boldsymbol{\mu}^{\|\cdot\|})$ of the vector $\boldsymbol{X_t} = (\mathcal{X}_{t-m}, \ldots, \mathcal{X}_t, \mathcal{X}_{t+1}, \ldots, \mathcal{X}_{t+h})$ on $C^{\|\cdot\|}_{m+h+1}$ is given by:

$$\Gamma^{\|\cdot\|} = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \right\}}, \tag{3.9}$$

$$\boldsymbol{\mu}^{\|\cdot\|} = \begin{cases} \mathbf{0}, & \textit{if } \alpha \neq 1 \\ -\frac{2\sigma}{\pi} \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} \pi_j \beta_j \boldsymbol{d}_{j,k} \ln \|\sigma \pi_j \boldsymbol{d}_{j,k}\|, & \textit{if } \alpha = 1 \end{cases} \tag{3.10}$$

where $\boldsymbol{d}_{j,k} = (d_{j,k+m}, \ldots, d_{j,k}, d_{j,k-1}, \ldots, d_{j,k-h})$, $w_{j,\vartheta} = (1 + \vartheta\beta_j)/2$, *for any* $k \in \mathbb{Z}$, $j = 1, \ldots, J$, δ *is the Dirac mass,* $\vartheta \in S_1$ *with* $S_1 = \{-1, +1\}$, *and if* $\boldsymbol{d}_{j,k} = \boldsymbol{0}$, *the term vanishes by convention from the sums.*

The necessary condition (3.8) extends what was noticed in the Proposition 3 of DFT, namely, that anticipativeness is a minimal requirement for past-representability. Importantly, notice that a single non-anticipative latent moving average is enough to render the aggregated process not past-representable, regardless of the other latent components. Also, for $\alpha \neq 1$, the past-representability of an aggregated process is equivalent to that of its latent moving averages, but this does not seem to hold in general for $\alpha = 1$. In the latter case however, if all the latent moving averages are symmetric, that is, $\beta_1 = \ldots = \beta_J = 0$, then the paths $\boldsymbol{X}_t$ are $\mathcal{S}1\mathcal{S}$ for any $m \geq 0$, $h \geq 1$ and $(\iota)(b)$ collapses to $(\iota)(a)$.

The representability condition also simplifies in the case of aggregated ARMA processes and requires each latent ARMA process to be anticipative.

Corollary 3.1. *For any* $j = 1, \ldots, J$, *let* $(X_{j,t})$ *be the ARMA strictly stationary solution of* $\Psi_j(F)\Phi_j(B)X_{j,t} = \Theta_j(F)H_j(B)\varepsilon_{j,t}$, *with mutually independent sequences* $\varepsilon_{j,t} \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta_j, 1, 0)$. *Define* $\mathcal{X}_t = \sigma \sum_{j=1}^{J} \pi_j X_{j,t}$ *for any positive weights* π_j *summing to 1 and* $\sigma > 0$. *Then, for any* $\alpha \in (0, 2)$, $(\beta_1, \ldots, \beta_J) \in [-1, 1]^J$, *the following statements are equivalent:*

(ι) $(\mathcal{X}_t)$ *is past-representable,*

$(\iota\iota)$ $\inf_j \deg(\Psi_j) \geq 1$,

$(\iota\iota\iota)$ $\sup_j m_{0,j} < +\infty$,

with the $m_{0,j}$*'s as in* (3.7). *Moreover, letting* $m \geq 0$, $h \geq 1$, *the aggregated process* $(\mathcal{X}_t)$ *is* (m, h)*-past-representable if and only if for any* $j = 1, \ldots, J$, $m_{0,j} < +\infty$, *and* $m \geq \max_j m_{0,j}$.

3.2. Tail conditional distribution of stable aggregates

Now, we derive the tail conditional distribution of linear stable aggregates. The case of a general past-representable stable aggregate is considered. We also pay a particular attention to the anticipative $\mathcal{G}\alpha\mathcal{S}$ AR(1) because to the best of our knowledge, no deconvolution estimation techniques exists for stable aggregates as defined in 2.1, except for the anticipative $\mathcal{G}\alpha\mathcal{S}$ AR(1) discussed in Section 2. To be relevant for the prediction framework, the Borel set B appearing in Equation 3.2 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 $\mathcal{X}_{t+1}, \ldots, \mathcal{X}_{t+h}$. For $\|\cdot\|$ a semi-norm on $\mathbb{R}^{m+h+1}$ satisfying (3.3), denote $S_{m+1}^{\|\cdot\|} = \{(s_{-m}, \ldots, s_0) \in \mathbb{R}^{m+1} : \quad \|(s_{-m}, \ldots, s_0, 0, \ldots, 0)\| = 1\}$.[6] Then, for any Borel set $V \subset S_{m+1}^{\|\cdot\|}$, define the Borel set $B(V) \subset C_{m+h+1}^{\|\cdot\|}$ as

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

[6]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.

Notice in particular that for $V = S^{\|\cdot\|}_{m+1}$, we have $B(V) = C^{\|\cdot\|}_{m+1}$. 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 $S^{\|\cdot\|}_{m+1}$. It will be useful in the following to notice that

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

where f the function defined by

$$f : \begin{array}{ccc} \mathbb{R}^{m+h+1} & \longrightarrow & \mathbb{R}^{m+1} \\ (x_{-m}, \ldots, x_0, x_1, \ldots, x_h) & \longmapsto & (x_{-m}, \ldots, x_0) \end{array}. \tag{3.11}$$

Let $\mathcal{X}_t$ an α-stable aggregate as in Definition 2.1. Assume $\mathcal{X}_t$ is (m,h)-past-representable, for some $m \geq 0$, $h \geq 1$. Also, we know by Proposition 3.1 $(\iota\iota)$, that $\Gamma^{\|\cdot\|}$ is of the form

$$\Gamma^{\|\cdot\|} = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \right\}}. \tag{3.12}$$

Proposition 3.2. *Let $\mathcal{X}_t$ be an α-stable aggregate as in Definition 2.1. Assume $\mathcal{X}_t$ is (m,h)-past-representable, for some $m \geq 0$, $h \geq 1$. Also, we know by Proposition 3.1 $(\iota\iota)$, that $\Gamma^{\|\cdot\|}$ is of the form*

$$\Gamma^{\|\cdot\|} = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \right\}}.$$

Under the above assumptions, we have

$$\mathbb{P}_x^{\|\cdot\|}\left(\boldsymbol{X_t}, A \middle| B(V)\right) \underset{x \to +\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}\left(\left\{ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \in A : \quad \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} \in V \right\}\right)}{\Gamma^{\|\cdot\|}\left(\left\{ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \in C^{\|\cdot\|}_{m+h+1} : \quad \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} \in V \right\}\right)}, \tag{3.13}$$

for any Borel sets $A \subset C^{\|\cdot\|}_{m+h+1}$, $V \subset S^{\|\cdot\|}_{m+1}$ such that $\left\{ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \in C^{\|\cdot\|}_{m+h+1} : \quad \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} \in V \right\} \neq \emptyset$, $\Gamma^{\|\cdot\|}\Big(\partial(A \cap B(V))\Big) = \Gamma^{\|\cdot\|}(\partial B(V)) = 0$, where $B(V) = V \times \mathbb{R}^h$ and f is as in (3.11).

Observe that setting $V = S^{\|\cdot\|}_{m+1}$, and A an arbitrarily small closed neighbourhood of all the points $(\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|)_{\vartheta,j,k}$, as in the single-component case we have $\lim_{x \to +\infty} \mathbb{P}\Big(\boldsymbol{X_t}/\|\boldsymbol{X_t}\| \in A \Big| \|\boldsymbol{X_t}\| > x\Big) = 1$. In other terms, when far from central values, the trajectory of process (X_t) necessarily features patterns of the same shape as some $\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$, which is a finite piece of a moving average's coefficient sequence. The index j indicates from which of the J underlying moving averages the pattern stems from, the index k points to which piece $(d_{j,k+m}, \ldots, d_{j,k}, d_{j,k-1}, \ldots, d_{j,k-h})$ 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 $(\varepsilon_{j,\tau})$). The likelihood of a pattern $\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$ can be evaluated by setting A to be a small neighbourhood of that point. In particular, only one pattern $\boldsymbol{d}_k/\|\boldsymbol{d}_k\|$ can appear through time for $J=1$ (up to a time shift and sign flipping). This is no longer the case in general for $J \geq 2$, where the shape of each extreme event appears as if being drawn from a collection of patterns.

Interestingly, as in DFT in the non-aggregated case, the observed path $(\mathcal{X}_{t-m}, \ldots, \mathcal{X}_{t-1}, \mathcal{X}_t)/\|\boldsymbol{X_t}\|$ will *a fortiori* be of the same shape as some $\vartheta(d_{j,k+m}, \ldots, d_{j,k+1}, d_{j,k})/\|\boldsymbol{d}_{j,k}\|$ 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. In practice, we anticipate that matching an observed path to a particular pattern j among the collection of J patterns will be challenging, even for a small number of latent components.

3.3. Example: Aggregation of Anticipative AR(1) Processes

We now consider the aggregation of stable anticipative AR(1) processes discussed in Section 2. We assume without loss of generality that the ρ_j's are distinct. For each anticipative AR(1) with parameter ρ_j, the moving average coefficients are of the form $(\rho_j^k \mathbb{1}_{\{k\geq 0\}})_k$, and thus, $m_{0,j}=0$ for all j, where the $m_{0,j}$'s are given in (3.7). By Corollary (3.1), we know for any $m \geq 0$, $h \geq 1$, the aggregated process $\mathcal{X}_t$ 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 3.3. *Let $\mathcal{X}_t$ be an aggregation of α-stable anticipative AR(1) processes as in Definition 2.1 with $d_{j,k}=\rho_j^k$ and general scale parameter $\sigma>0$.*

Letting $\boldsymbol{X_t} = (\mathcal{X}_{t-m}, \ldots, \mathcal{X}_t, \mathcal{X}_{t+1}, \ldots, \mathcal{X}_{t+h})$ for $m \geq 0$, $h \geq 1$, its spectral measure on $C_{m+h+1}^{\|\cdot\|}$ for a seminorm satisfying (3.3) is given by

$$\Gamma^{\|\cdot\|} = \sum_{\vartheta \in S_1} \left[w_\vartheta \delta_{\{(\vartheta,0,\ldots,0)\}} + \sum_{j=1}^{J} \sigma^\alpha \pi_j^\alpha \left(w_{j,\vartheta} \sum_{k=-m+1}^{h-1} \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right\}} + \frac{\bar{w}_{j,\vartheta}}{1-|\rho_j|^\alpha} \|\boldsymbol{d}_{j,h}\|^\alpha \delta_{\left\{\frac{\vartheta\, \boldsymbol{d}_{j,h}}{\|\boldsymbol{d}_{j,h}\|}\right\}} \right) \right], \tag{3.14}$$

where for all $\vartheta \in S_1$, $j \in \{1, \ldots, J\}$ and $-m+1 \leq k \leq h$,

$$\boldsymbol{d}_{j,k} = (\rho_j^{k+m} \mathbb{1}_{\{k \geq -m\}}, \ldots, \rho_j^k \mathbb{1}_{\{k\geq 0\}}, \rho_j^{k-1} \mathbb{1}_{\{k \geq 1\}}, \ldots, \rho_j^{k-h} \mathbb{1}_{\{k \geq h\}}),$$
$$w_{j,\vartheta} = (1+\vartheta \beta_j)/2,$$
$$w_\vartheta = \sum_{j=1}^{J} \sigma^\alpha \pi_j^\alpha w_{j,\vartheta},$$
$$\bar{w}_{j,\vartheta} = (1+\vartheta \bar{\beta}_j)/2,$$
$$\bar{\beta}_j = \beta_j \frac{1-\rho_j^{<\alpha>}}{1-|\rho_j|^\alpha},$$

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 the ρ_j's are positive. Let us first introduce useful neighbourhoods of the distinct charged points of $\Gamma^{\|\cdot\|}$. Denote $\boldsymbol{d}_{0,-m} = (\overbrace{1,0,\ldots,0}^{m+h+1})$ so that the charged points of $\Gamma^{\|\cdot\|}$ are all of the form $\vartheta\boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$ with indexes (ϑ,j,k) in the set $\mathcal{I} := S_1 \times \Big(\{1,\ldots,J\}\times\{-m,h\}\cup\{(0,-m)\}\Big)$. With f as in (3.11), define for any $(\vartheta_0,j_0,k_0)\in\mathcal{I}$, the set V_0 as any closed neighbourhood of $\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\|$ such that

$$\forall(\vartheta',j',k')\in\mathcal{I},\qquad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|}\in V_0 \implies \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}, \tag{3.15}$$

In other terms, $V_0\times\mathbb{R}^d$ is a subset of $C_{m+h+1}^{\|\cdot\|}$ in which the only points charged by $\Gamma^{\|\cdot\|}$ all have the first $(m+1)^{\text{th}}$ coinciding with $\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\|$. Define also $A_{\vartheta,j,k}$ for any (ϑ,j,k) as any closed neighbourhood of $\vartheta\boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$ which does not contain any other charged point of $\Gamma^{\|\cdot\|}$, that is,

$$\forall(\vartheta',j',k')\in\mathcal{I},\qquad \frac{\vartheta'\boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|}\in A_{\vartheta,j,k} \implies (\vartheta',j',k')=(\vartheta,j,k). \tag{3.16}$$

Proposition 3.3. *Let $\mathcal{X}_t$ be an aggregation of α-stable anticipative AR(1) processes as in Definition 2.1 with $d_{j,k}=\rho_j^k\in(0,1)$ for all j's.,Let $\boldsymbol{X_t}$, the $\boldsymbol{d}_{j,k}$'s and the spectral measure of $\boldsymbol{X_t}$ be as given in Lemma 3.3, for any $m\geq 0$, $h\geq 1$. Let V_0 be any small closed neighbourhood of $\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\|$ in the sense of (3.15) for some $(\vartheta_0,j_0,k_0)\in\mathcal{I}$ and let $B(V_0)=V_0\times\mathbb{R}^h$. Then, with $A_{\vartheta,j,k}$ an arbitrarily small neighbourhood of some $\vartheta\boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$ as in (3.16), the following hold.*

(ι) ***Case** $m\geq 1$.*

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

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t},A_{\vartheta,j,k}\Big|B(V_0)\Big)\underset{x\to\infty}{\longrightarrow}\begin{cases}|\rho_{j_0}|^{\alpha k}(1-|\rho_{j_0}|^{\alpha})\delta_{\vartheta_0}(\vartheta)\delta_{j_0}(j), & 0\leq k\leq h-1,\\[2ex] |\rho_{j_0}|^{\alpha h}\delta_{\vartheta_0}(\vartheta)\delta_{j_0}(j), & k=h.\end{cases}$$

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

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t},A_{\vartheta,j,k}\Big|B(V_0)\Big)\underset{x\to\infty}{\longrightarrow}\delta_{\vartheta_0}(\vartheta)\delta_{j_0}(j)\delta_{k_0}(k).$$

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

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t},A_{\vartheta,j,k}\Big|B(V_0)\Big)\underset{x\to\infty}{\longrightarrow}\begin{cases}\dfrac{\sum_{i=1}^J\pi_i^{\alpha}w_{i,\vartheta_0}}{\sum_{i=1}^J p_{i,\vartheta_0}}\delta_{\{\vartheta_0\}}(\vartheta), & k=0\\[3ex] \dfrac{p_{j,\vartheta_0}}{\sum_{i=1}^J p_{i,\vartheta_0}}|\rho_j|^{\alpha k}(1-|\rho_j|^{\alpha})\delta_{\{\vartheta_0\}}(\vartheta), & 1\leq k\leq h-1,\\[3ex] \dfrac{p_{j,\vartheta_0}}{\sum_{i=1}^J p_{i,\vartheta_0}}|\rho_j|^{\alpha h}\delta_{\{\vartheta_0\}}(\vartheta), & k=h,\end{cases}$$

with $p_{j,\vartheta_0}=\pi_j^{\alpha}w_{j,\vartheta_0}/(1-|\rho_j|^{\alpha})$.

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 $k_0 \in \{0, \ldots, h\}$ or $k_0 \in \{-m, \ldots, -1\}$. For the latter, the asymptotic probability of the whole path $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ being in an arbitrarily small neighbourhood of $\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$ is 1 if and only if $\vartheta = \vartheta_0$, $j = j_0$, $k = k_0$: 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 $\vartheta = \vartheta_0$ and $j = j_0$, 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 $\boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$ and of their restrictions to the first $m+1$ components $f(\boldsymbol{d}_{j,k})/\|\boldsymbol{d}_{j,k}\|$. On the one hand (omitting ϑ),

$$\frac{\boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} = \begin{cases} \dfrac{(\overbrace{\rho_j^{k+m}, \ldots, \rho_j^k}^{m+1}, \overbrace{\rho_j^{k-1}, \ldots, \rho_j, 1, 0, \ldots, 0}^{h})}{\|(\rho_j^{k+m}, \ldots, \rho_j^k, \rho_j^{k-1}, \ldots, \rho_j, 1, 0, \ldots, 0)\|}, & \text{for} \qquad k \in \{0, \ldots, h\}, \\[4ex] \dfrac{(\rho_j^{k+m}, \ldots, \rho_j, 1, 0, \ldots, 0, 0, \ldots, 0)}{\|(\underbrace{\rho_j^{k+m}, \ldots, \rho_j, 1, 0, \ldots, 0}_{m+1}, \underbrace{0, \ldots, 0}_{h})\|}, & \text{for} \qquad k \in \{-m, \ldots, -1\}. \end{cases}$$

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

$$\frac{f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} = \begin{cases} \dfrac{(\overbrace{\rho_j^{m}, \ldots, \rho_j, 1}^{m+1})}{\|(\underbrace{\rho_j^{m}, \ldots, \rho_j, 1}_{m+1}, \underbrace{0, \ldots, 0, 0, \ldots, 0}_{h})\|}, & \text{for} \qquad k \in \{0, \ldots, h\}, \\[4ex] \dfrac{(\overbrace{\rho_j^{k+m}, \ldots, \rho_j, 1, 0, \ldots, 0}^{m+1})}{\|(\underbrace{\rho_j^{k+m}, \ldots, \rho_j, 1, 0, \ldots, 0}_{m+1}, \underbrace{0, \ldots, 0}_{h})\|}, & \text{for} \qquad k \in \{-m, \ldots, -1\}. \end{cases}$$

Thus, conditioning the trajectory on the event $\{f(\boldsymbol{X_t})/\|\boldsymbol{X_t}\| \approx f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\|\}$ 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 j^{th} 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 the h^{th}. However, the shape of the observed path is that of a piece of exponential with growth rate ρ_j^{-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 $|\rho_j|^{\alpha k}(1-|\rho_j|^{\alpha})$

corresponds to the probability that the bubble will peak in exactly k periods ($0 \leq k < h$), and $|\rho_j|^{\alpha h}$ corresponds to the probability that the bubble will last at least h more periods.

The previous statement confirms the interpretation of the conditional moments proposed in Fries (2022) for the stable anticipative AR(1) case ($J = 1$). It also extends it in two ways:

(ι) by accounting for paths rather than point prediction,

($\iota\iota$) by showing that the aggregation of AR(1) processes also features killed exponential explosive episodes but with various growth rates and crash probabilities.

Proposition 3.3 furthermore shows that asymptotically, as few as two observations are sufficient to identify the growth rate ρ_j^{-1} of an ongoing extreme episode,[7] and the conditional dynamics within this given event will be similar to that of a simple AR(1) with corresponding parameter. An identification of the growth rate in the early developments of the bubble appears possible, allowing to infer in advance the odds of crashes, as long as the latent components parameters are identified.

4. Monte-Carlo Simulation

4.1. Estimation accuracy

This section presents Monte Carlo evidence on the finite-sample performance of the minimum distance estimator introduced in Section 2. The observed process is generated by the aggregation of two independent α-stable AR(1) processes:

$$\mathcal{X}_t = \pi_1 X_{1,t} + \pi_2 X_{2,t}, \tag{4.1}$$

$$X_{j,t} = \rho_j X_{j,t-1} + \varepsilon_{j,t}, \quad \varepsilon_{j,t} \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, \beta, 1, 0), \tag{4.2}$$

where $j \in \{1, 2\}$ and $\rho_j \in (0, 1)$. We fix $\sigma = 1.6$, $\pi_1 = 7/16$ and $\pi_2 = 9/16$, leading to the combined scale parameters $\varsigma_1 = \sigma\pi_1 = 0.7$ and $\varsigma_2 = \sigma\pi_2 = 0.9$. Three distributional settings are considered: (i) Cauchy ($\mathcal{S}1\mathcal{S}$, $\alpha = 1$, $\beta = 0$); (ii) symmetric α-stable ($\mathcal{S}\alpha\mathcal{S}$, $\alpha = 1.5$, $\beta = 0$); and (iii) general α-stable ($\mathcal{G}\alpha\mathcal{S}$, $\alpha = 1.5$, $\beta = 0.3$). For each case, we perform 1,000 replications with sample sizes $T \in \{250, 500, 1{,}000\}$. Estimation uses uniform weights and grids of 10 equally spaced points in $[-0.5, 0.5]$ for u and v.

Table 1 reports the bias, root mean square error (RMSE), and mean relative error (MRE) across all scenarios. Several patterns emerge. First, the dominant autoregressive coefficient ρ_1 is precisely estimated across all settings, with MRE below 5% in the $\mathcal{S}1\mathcal{S}$ case and below 9% even in the $\mathcal{G}\alpha\mathcal{S}$ case for $T = 250$. Second, the smaller coefficient ρ_2 and the combined scale parameters ς_1, ς_2 exhibit substantially higher

[7] This holds asymptotically in the (semi-)norm of the observed path, but in practice it can be expected that the noise surrounding the trajectory will make this identification difficult with only two observations. Longer path lengths (higher m) may provide robustness to the identification, but could also incorporate some bias by taking into account past extreme events, such as now-collapsed bubbles. One can suspect a bias-variance trade-off when searching for an optimal choice of m.

relative errors, reflecting the inherent difficulty in disentangling the individual component contributions from the aggregate signal. Third, the tail index α is recovered with high accuracy (MRE around 8% at $T = 250$, declining to 4–5% at $T = 1{,}000$), confirming the informational content of the empirical characteristic function for identifying heavy-tail behavior. Fourth, in the $\mathcal{G}\alpha\mathcal{S}$ setting, the asymmetry parameter β is the most difficult to estimate (MRE of 63% at $T = 250$), although its identification is not required for the autoregressive and scale parameters. Across all scenarios, the RMSE and MRE decrease consistently with T, confirming the good finite-sample behavior of the estimator.

Table 1: Monte Carlo estimation accuracy for α-stable AR(1) aggregates (1,000 replications)

		$T = 250$			$T = 500$			$T = 1000$		
θ	True	Bias	RMSE	MRE	Bias	RMSE	MRE	Bias	RMSE	MRE
Panel A: Cauchy ($\mathcal{S}1\mathcal{S}$, $\alpha = 1$, $\beta = 0$)										
ρ_1	0.800	−0.008	0.050	0.046	−0.004	0.030	0.029	−0.002	0.020	0.019
ς_1	0.700	−0.023	0.322	0.368	0.002	0.249	0.275	−0.002	0.179	0.196
ρ_2	0.300	−0.035	0.169	0.455	−0.025	0.128	0.333	−0.012	0.089	0.223
ς_2	0.900	−0.035	0.279	0.249	−0.015	0.193	0.171	−0.007	0.134	0.119
Panel B: Symmetric α-stable ($\mathcal{S}\alpha\mathcal{S}$, $\alpha = 1.5$, $\beta = 0$)										
ρ_1	0.800	0.013	0.081	0.073	0.000	0.062	0.057	−0.005	0.046	0.044
ς_1	0.700	−0.101	0.276	0.325	−0.021	0.216	0.247	0.019	0.168	0.195
ρ_2	0.300	−0.127	0.209	0.600	−0.096	0.190	0.529	−0.072	0.148	0.391
ς_2	0.900	−0.124	0.235	0.212	−0.098	0.197	0.175	−0.071	0.153	0.135
α	1.500	−0.037	0.173	0.086	−0.014	0.127	0.065	−0.004	0.088	0.046
Panel C: General α-stable ($\mathcal{G}\alpha\mathcal{S}$, $\alpha = 1.5$, $\beta = 0.3$)										
ρ_1	0.800	−0.005	0.091	0.084	−0.015	0.073	0.069	−0.013	0.057	0.055
ς_1	0.700	−0.054	0.285	0.335	0.013	0.234	0.278	0.026	0.200	0.236
ρ_2	0.300	−0.089	0.208	0.594	−0.068	0.186	0.510	−0.052	0.146	0.388
ς_2	0.900	−0.124	0.234	0.211	−0.115	0.205	0.179	−0.092	0.178	0.154
α	1.500	−0.026	0.164	0.083	−0.009	0.120	0.060	−0.003	0.082	0.042
β	0.300	−0.007	0.262	0.628	−0.003	0.186	0.457	−0.001	0.132	0.328

Notes: True parameter values are $\rho_1 = 0.8$, $\rho_2 = 0.3$, $\varsigma_1 = 0.7$, $\varsigma_2 = 0.9$. MRE denotes the mean relative error $\mathbb{E}[|\hat{\theta} - \theta_0|/|\theta_0|]$. Estimation uses the empirical characteristic function with uniform weights on a 10×10 grid in $[-0.5, 0.5]^2$.

4.2. Subsampling-based verification of asymptotic normality

To empirically verify Proposition 2.1, we implement a subsampling methodology following Politis and Romano (1994) and Politis, Romano and Wolf (1999). A detailed description of the subsampling methodology (Section S.1.2) and comprehensive results for all sample sizes and parameterizations are provided in the Online Supplement.[8] Given a full sample of size n, we construct non-overlapping subsamples of size $b < n$, with $b/n \to 0$:

$$\mathcal{X}_b^{(i)} = \{\mathcal{X}_{(i-1)b+1}, \dots, \mathcal{X}_{ib}\}, \quad i = 1, \dots, N_b = \lfloor n/b \rfloor. \tag{4.3}$$

Since the individual scale parameters ς_1 and ς_2 exhibit slow finite-sample convergence (see Section S.1.2), we apply a post-estimation reparameterization:

$$\vartheta = (\rho_1, \rho_2, \sigma, \pi_1, \alpha), \quad \text{where} \quad \sigma = \varsigma_1 + \varsigma_2, \quad \pi_1 = \frac{\varsigma_1}{\sigma}. \tag{4.4}$$

The scaled subsample deviations are

$$Z_b^{(i)} = \sqrt{b}\left(\hat{\vartheta}_b^{(i)} - \hat{\vartheta}_n\right), \quad i = 1, \dots, N_b, \tag{4.5}$$

where $\hat{\vartheta}_n$ is the full-sample estimator mapped to the reparameterized space. Following Politis, Romano and Wolf (1999), we set $b = \lfloor n^{2/3} \rfloor$. We conduct $M = 200$ Monte Carlo replications for each sample size $n \in \{250, 500, 1{,}000, 10{,}000, 50{,}000, 100{,}000\}$ with true parameter values $\theta_0 = (0.8, 0.3, 0.7, 0.9, 1.5)$, yielding $\vartheta_0 = (0.8, 0.3, 1.6, 0.4375, 1.5)$. [9]

Table 2 confirms the theoretical prediction of Proposition 2.1. The tail index α and the total scale σ exhibit the fastest convergence toward normality, achieving near-nominal CI coverage at $n \geq 10{,}000$ (α: 90.0%/94.0% at the 90%/95% levels; σ: 88.5%/95.5%). The autoregressive coefficients ρ_1 and ρ_2 converge more slowly, with ρ_2 exhibiting persistent positive mean drift. The mixing proportion π_1 is the most challenging parameter: its coverage collapses to 14.5% at the 95% level for $n = 100{,}000$, raising the question of whether it is genuinely identified by the CF-based objective. A criterion difference test for $H_0\colon \pi_1 = 0.5$ (see Section S.1.5) confirms that the MDE objective is essentially flat in the π_1 direction, as the test never rejects H_0 at any conventional level for n up to 50,000. Imposing $\pi_1 = 1/2$ stabilizes inference on σ and α, which achieve near-nominal coverage at moderate sample sizes, but introduces a specification bias on ρ_1 and ρ_2 that becomes detectable at very large samples (see Section S.1.4 for detailed results).

To disentangle the respective contributions of specification bias and slow finite-sample convergence, we repeat the restricted exercise under the true constraint $\pi_1 = \pi_{1,0} = 0.4375$ (Section S.1.5). Table 3

[8] All tables, sections, or figures whose numbering begins with "S" refer to the online supplementary material

[9] Table S.1 summarizes the Monte Carlo design. Comprehensive results for all sample sizes are reported in Tables S.9–S. 14. Visual diagnostics are displayed in Figures S.7–S.12.

Table 2: Subsampling results in the reparameterized space for selected sample sizes

n	Param.	True	Avg. Est.	Std. Est.	Dev. Mean	Skew	Kurt	Quantile CI		Normal CI	
								qCov90	qCov95	nCov90	nCov95
1,000	ρ_1	0.80	0.812	0.086	−0.15	−0.41	−0.62	0.555	0.605	0.690	0.745
	ρ_2	0.30	0.296	0.157	0.57	0.06	−0.96	0.450	0.455	0.595	0.660
	σ	1.60	1.543	0.188	−0.55	0.19	−0.52	0.600	0.635	0.665	0.715
	π_1	0.4375	0.414	0.162	−0.47	−0.23	−0.42	0.285	0.315	0.510	0.610
	α	1.50	1.493	0.093	0.13	−0.15	−0.59	0.800	0.815	0.825	0.885
10,000	ρ_1	0.80	0.786	0.029	0.43	0.18	−0.70	0.530	0.620	0.770	0.885
	ρ_2	0.30	0.260	0.055	1.15	0.09	−0.97	0.505	0.520	0.705	0.825
	σ	1.60	1.599	0.053	−1.38	0.01	−0.69	0.825	0.885	0.885	0.955
	π_1	0.4375	0.473	0.058	−1.55	−0.72	−0.44	0.210	0.230	0.490	0.735
	α	1.50	1.498	0.028	0.00	−0.02	−0.35	0.895	0.925	0.900	0.940
100,000	ρ_1	0.80	0.778	0.013	1.11	0.89	0.85	0.145	0.155	0.155	0.170
	ρ_2	0.30	0.249	0.025	1.75	−0.38	−0.02	0.175	0.200	0.170	0.205
	σ	1.60	1.609	0.017	−1.48	−0.31	−0.12	0.795	0.850	0.905	0.950
	π_1	0.4375	0.491	0.028	−2.59	−1.25	1.43	0.145	0.155	0.145	0.145
	α	1.50	1.501	0.009	−0.05	−0.00	−0.20	0.910	0.960	0.925	0.965

Notes: Avg. Est. and Std. Est. are computed over $M = 200$ replications. Dev. Mean, Skew, and Kurt refer to the mean, skewness, and excess kurtosis of the non-overlapping block scaled deviations $Z_b^{(i)}$. qCov and nCov denote the average coverage of quantile-based and normal-approximation confidence intervals. Complete results for all sample sizes are in the Online Supplement, Section S.1.

reports the results for $n \in \{1{,}000,\ 10{,}000,\ 100{,}000\}$. The improvement relative to both the unrestricted and the misspecified-restricted cases is substantial. At $n = 10{,}000$, ρ_1 is essentially unbiased ($\bar{\hat{\rho}}_1 = 0.799$) with 96.5%/98.0% normal CI coverage at the 90%/95% levels, compared to 27.0%/42.5% under $\pi_1 = 1/2$. For ρ_2, coverage reaches 86.5%/92.0%, up from 62.5%/72.5%. The parameters σ and α retain near-nominal coverage comparable to the $\pi_1 = 1/2$ case. At $n = 100{,}000$, all four parameters achieve nominal or near-nominal coverage (ρ_1: 95.0%/98.0%; ρ_2: 90.0%/98.0%; σ: 89.5%/94.5%; α: 94.0%/96.5%), confirming that the correctly restricted estimator converges to its asymptotic Gaussian limit without residual specification bias. The only notable departure from normality is a persistent left skew in ρ_1 (skewness ≈ -0.77 at $n = 100{,}000$), which inflates the Shapiro–Wilk rejection rate to 44.5% but does not materially affect coverage.

5. Application to financial markets

To illustrate the empirical relevance of our estimator and forecasting theoretical results, we apply them to financial data. In particular, we focus on the CBOE Crude Oil ETF Volatility Index (OVX), which

Table 3: Restricted subsampling results under $\pi_1 = 0.4375$ (true value) for selected sample sizes

								Quantile CI		Normal CI	
n	Param.	True	Avg. Est.	Std. Est.	Dev. Mean	Skew	Kurt	qCov90	qCov95	nCov90	nCov95
1,000	ρ_1	0.80	0.761	0.131	-0.51	-0.49	-0.34	0.750	0.780	0.860	0.885
	ρ_2	0.30	0.321	0.190	0.42	0.05	-1.07	0.370	0.375	0.510	0.595
	σ	1.60	1.594	0.151	0.12	0.09	-0.47	0.680	0.740	0.750	0.810
	α	1.50	1.496	0.090	0.22	-0.11	-0.61	0.800	0.820	0.825	0.880
10,000	ρ_1	0.80	0.799	0.010	-0.60	-0.70	0.19	0.890	0.920	0.965	0.980
	ρ_2	0.30	0.298	0.050	0.37	0.13	-1.04	0.775	0.775	0.865	0.920
	σ	1.60	1.599	0.044	0.09	0.14	-0.44	0.830	0.880	0.860	0.925
	α	1.50	1.498	0.028	0.12	0.01	-0.36	0.880	0.915	0.895	0.925
100,000	ρ_1	0.80	0.800	0.003	-0.25	-0.77	1.43	0.920	0.955	0.950	0.980
	ρ_2	0.30	0.301	0.017	0.04	-0.01	-0.11	0.925	0.975	0.900	0.980
	σ	1.60	1.601	0.014	0.02	0.07	-0.21	0.885	0.935	0.895	0.945
	α	1.50	1.500	0.008	0.06	0.01	-0.19	0.920	0.960	0.940	0.965

Notes: See notes to Table 2. Estimation under the true restriction $\pi_1 = \varsigma_1/\sigma = 0.4375$. Complete results for all sample sizes are in the Online Supplement, Section S.1.4.

reflects by essence the market's anticipation of the volatility of crude oil ETF prices over the next 30 days. VIX-type indexes are often referred to as *fear indices* as they aggregate all sources of investors' expectations. The large body of literature on heterogeneous agent models (e.g. Agliari et al. , 2018) suggests that the fundamentalist/chartist dichotomy is likely to generate distinct dynamics in such an index, particularly during periods of growing market fear. We collect the CBOE OVX index from the FRED website, sampled at weekly frequency over the period 23/05/2015–23/05/2025 ($T = 522$), and linearly detrended to avoid high-frequency noise contamination (see Hecq and Voisin , 2021, for a discussion on the pre-treatment of data). We estimate three specifications as described in Section 4, with initial values obtained from de Truchis et al. (2025b): a general α-stable model ($\mathcal{G}\alpha\mathcal{S}$), a symmetric α-stable model ($\mathcal{S}\alpha\mathcal{S}$), and a Cauchy model ($\mathcal{S}1\mathcal{S}$).

Table 4 reveals several key patterns. The $\mathcal{G}\alpha\mathcal{S}$ specification provides strong evidence of anticipative dynamics: both AR coefficients are highly significant ($\hat{\rho}_1 = 0.80$, $\hat{\rho}_2 = 0.85$), and the two latent components are clearly differentiated, the first captures more abrupt volatility bursts (lower weight $\hat{\pi}_1 = 0.28$), while the second drives more persistent explosive episodes ($\hat{\pi}_2 = 0.72$). The estimated tail index $\hat{\alpha} = 1.47$ confirms the presence of heavy tails well beyond Gaussian accommodation. The asymmetry parameter $\hat{\beta} = -0.13$ is not significant at the 5% level, yet numerically distorts the parameter structure of the $\mathcal{S}\alpha\mathcal{S}$ model considerably ($\hat{\rho}_1 = 0.25$, $\hat{\rho}_2 = 0.99$). The $\mathcal{S}1\mathcal{S}$ Cauchy restriction ($\alpha = 1$) appears overly binding given the estimated $\hat{\alpha}$

Table 4: Estimation results for the OVX index under three specifications.

	$\mathcal{G}\alpha\mathcal{S}$			$\mathcal{S}\alpha\mathcal{S}$			$\mathcal{S}1\mathcal{S}$		
$\widehat{\theta}$	Estimate	Std.	t-stat	Estimate	Std.	t-stat	Estimate	Std.	t-stat
ρ_1	0.7989	0.0673	11.862	0.2507	0.0077	32.477	0.9226	0.0082	112.824
ρ_2	0.8470	0.0668	12.678	0.9865	0.0040	244.560	0.9346	0.0074	126.404
α	1.4686	0.0995	14.764	1.2405	0.0084	147.613	–	–	–
β	−0.1275	0.0684	−1.863	–	–	–	–	–	–
σ	2.0932	0.2400	8.723	0.8964	0.0212	42.226	0.1966	0.0294	6.692
π_1	0.2790	0.0403	6.930	0.8915	0.0052	171.084	0.5029	0.0511	9.833
π_2	0.7210	0.0409	17.622	0.1085	0.0182	5.957	0.4971	0.0545	9.121

values.

Figure 2 presents the deconvolution of the OVX index, obtained via the dual MCMC filtering procedure of de Truchis et al. (2025b). The two components exhibit clearly differentiated roles: the first component captures abrupt, short-lived volatility bursts, while the second tracks more sustained explosive patterns. Periods of extreme oil market stress, most visibly the 2020 disruption, feature a superposition of both dynamics, whose heterogeneous persistence properties are precisely what motivates our aggregate modeling framework.

To demonstrate the forecasting potential of the estimated model, we conduct an in-sample prediction exercise for the 2020 oil market disruption. Setting January 2020 as the cut-off, we apply Proposition 3.3 to forecast crash probabilities and trajectory paths for each latent component (pattern matching length $m = 20$). Our approach exploits the theoretical result that during extreme events, trajectories conform to specific normalized patterns $\vartheta \mathbf{d}_{j,k}/\|\mathbf{d}_{j,k}\|$: we first identify the component j_0 and position k_0 by matching the observed pre-cut-off trajectory, then compute conditional crash probabilities $|\rho_{j_0}|^{\alpha k}(1-|\rho_{j_0}|^{\alpha})$ and survival probabilities $|\rho_{j_0}|^{\alpha h}$ for all future horizons h.

For the first component ($k_0 = 3$), forecasted values escalate rapidly from 11.27 to 208.73 before the predicted crash at horizon $h = 14$ (at the 99% risk threshold). For the second component ($k_0 = 1$), crash probabilities build more gradually but the trajectory reaches higher absolute values (up to 223.7 at $h = 18$) before the predicted collapse. Detailed per-component crash probability profiles and forecast trajectories across three risk thresholds (90%, 95%, 99%) are provided in the Online Supplement, Section S.2 (Figures S.25–S.26 and Table S.28).

Figure 3 presents the combined forecast at the 99% threshold. The aggregate trajectory closely tracks the realized path during the March 2020 spike, reaching approximately 220 before the predicted crash,,

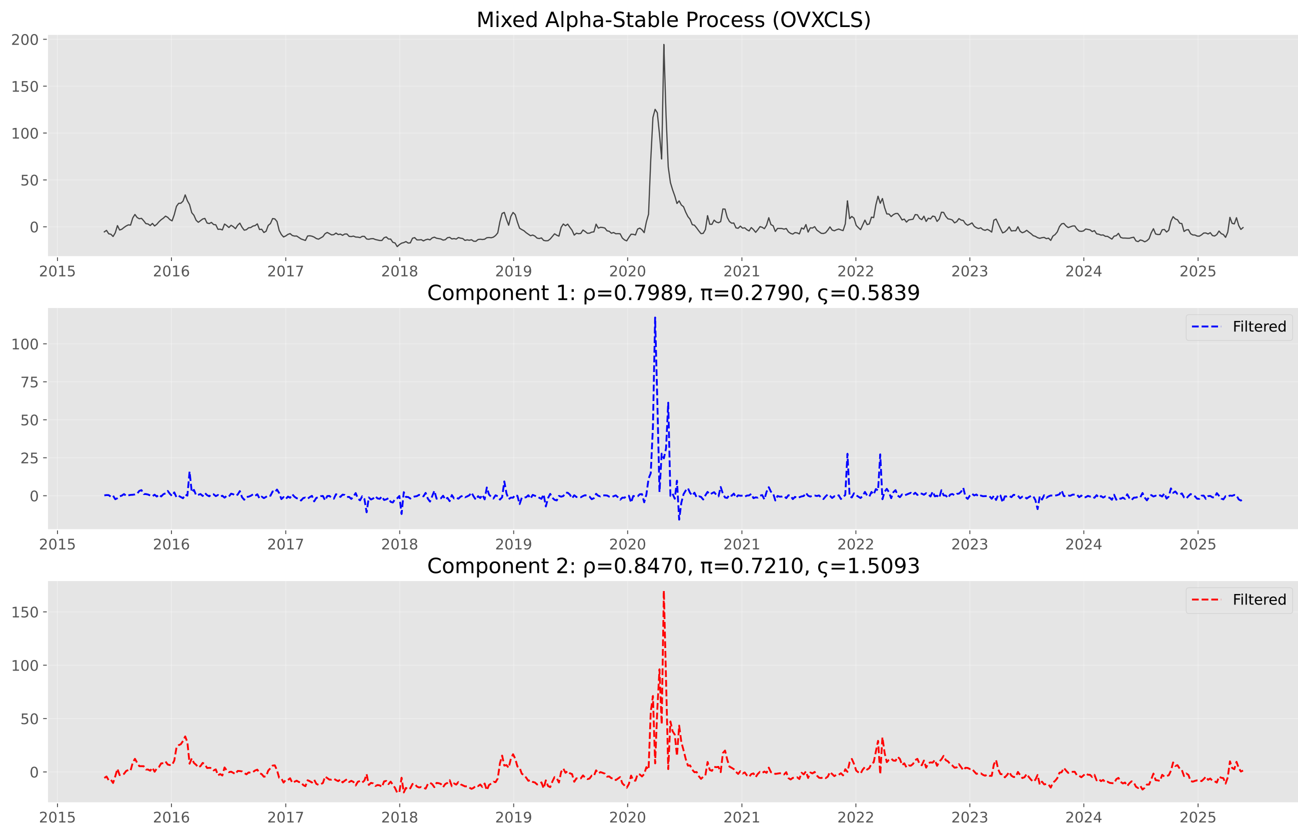

Figure 2: Deconvolution of the OVX index from the $\mathcal{G}\alpha\mathcal{S}$ two-component model, filtered using de Truchis et al. (2025b). Top panel: observed detrended OVX series. Middle panel: filtered first component ($\hat{\rho}_1 = 0.7989$, $\hat{\pi}_1 = 0.2790$). Bottom panel: filtered second component ($\hat{\rho}_2 = 0.8470$, $\hat{\pi}_2 = 0.7210$).

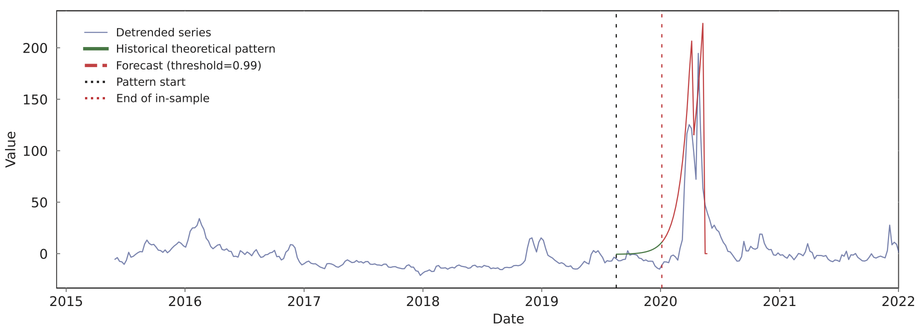

Figure 3: Combined in-sample forecast at the 99% risk threshold for the 2020 OVX bubble. The blue line shows the detrended series, the green segment indicates the matched historical theoretical pattern starting at the black dotted vertical line, and the red curve displays the out-of-sample forecast beyond the January 2020 cut-off (red dotted vertical line).

remarkably close to the observed peak of around 230. This validates our framework's ability to provide early warning signals for extreme volatility events in commodity markets, and illustrates how disentangling heterogeneous bubble components enhances forecast precision relative to single-component models.

6. Conclusion

This paper addresses a fundamental limitation in the empirical modeling of rational asset bubbles in financial markets by introducing a novel framework based on α-stable moving average aggregates. Traditional approaches to bubble modeling based on anticipative heavy-tailed processes impose uniform bubble patterns across different episodes, contradicting the observed heterogeneity in market dynamics. Our contribution is both theoretical and methodological. Theoretically, we develop a flexible model built on α-stable moving average aggregates that accommodates diverse bubble growth patterns and crash dynamics. We establish that this model admits a semi-norm representation on a unit cylinder, similar to non-aggregated moving averages, thereby enabling the forecasting of bubble episodes with heterogeneous growth trajectories. We extend the spectral representation of stable processes to aggregated components and derive conditions under which the tail conditional distribution can be used for prediction, showing that anticipativeness remains a necessary condition for past-representability even in the aggregated case. Methodologically, we develop a minimum distance estimation procedure based on the joint characteristic function that effectively identifies the parameters of stable aggregates. Unlike existing approaches limited to the Cauchy case with continuous support distributions, our framework extends to the general α-stable family with discrete support, making it more suitable for empiricalapplications. Monte Carlo simulations confirm robust finite-sample performance, and a subsampling procedure supports the asymptotic normality of the estimator, while revealing heterogeneous convergence speeds across parameter dimensions and a near-flat objective surface in the mixing proportion direction. An empirical illustration using the CBOE OVX index reveals the presence of multiple anticipative components with distinct persistence properties and asymmetric weights. The deconvolution analysis shows that the 2020 oil market disruption actually comprises multiple superimposed processes with heterogeneous growth rates and crash probabilities, and our forecasting framework successfully anticipates both the timing and magnitude of the March 2020 volatility spike.

References

Agliari, A., Naimzada, A., Pecora, N. (2018). Boom-bust dynamics in a stock market participation model with heterogeneous traders. *Journal of Economic Dynamics and Control*, 91, 458-468.

Andrews, B., Calder, M., and Davis, R. 2009. Maximum likelihood estimation for α-stable autoregressive process. *The Annals of Statistics*. 37, 1946–1982.

Blanchard, O.J., and Watson, M.W. 1982. Bubbles, Rational Expectations and Financial Markets. *NBER Working Papers*, 0945.

Brockwell, P.J., and Davis, R.A. 1991. *Time series: theory and methods*. Springer science & business media.

Cavaliere, G., Nielsen, H.B., and Rahbek, A. 2020. Bootstrapping noncausal autoregressions: with applications to explosive bubble modeling. *Journal of Business & Economic Statistics*, 38(1), 55-67.

Cubadda, G., Hecq, A., Telg, S. (2019). Detecting Co-Movements in Non-Causal Time Series. *Oxford Bulletin of Economics and Statistics*, 81(3), 697-715.

de Truchis, G., Fries, S., Thomas, A. (2025). Forecasting Extreme Trajectories Using Seminorm Representations *Working Paper 2025-06*, Chaire Économie du Climat, Paris.

de Truchis, G., Thomas, A., L. Vaudrée (2025). Deconvolution and Filtering of Non-Causal Alpha-Stable Processes *On going paper.*

Doukhan, P. 1994. Mixing: Properties and Examples. *Lecture Notes in Statistics, Vol. 85. New York: Springer-Verlag.*

Erdős, P. and A. H., Stone. 1970. On the sum of two Borel sets. *Proceedings of the American Mathematical Society*, 25, 304-306.

Fries, S. (2022) Conditional Moments of Noncausal Alpha-Stable Processes and the Prediction of Bubble Crash Odds. *Journal of Business & Economic Statistics*, 40:4, 1596-1616

Fries, S., Zakoian, J. M. (2019). Mixed causal-noncausal ar processes and the modelling of explosive bubbles. *Econometric Theory*, 35(6), 1234-1270.

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

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

Gourieroux, C., Jasiak, J. (2018). Misspecification of noncausal order in autoregressive processes. *Journal of Econometrics*, 205(1), 226-248.

Gourieroux, C. and Jasiak, J. 2023. Generalized Covariance Estimator. *Journal of Business & Economic Statistics.* 41(4), 1315–1327.

Gourieroux, C., Jasiak, J., Monfort, A. (2020). Stationary bubble equilibria in rational expectation models. *Journal of Econometrics*, 218(2), 714-735.

Gourieroux, C., Jasiak, J., & Tong, M. (2021). Convolution-based filtering and forecasting: An application to WTI crude oil prices. *Journal of Forecasting, 40(7), 1230-1244.*

Gouriéroux, C., Lu, Y., and Robert, C.-Y. 2025. The Causal-Noncausal Tail Processes: An Introduction. *arXiv preprint arXiv:2506.04046.*

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

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

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

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

Hecq, A., Issler, J. V., and Telg, S. 2020. Mixed causal–noncausal autoregressions with exogenous regressors. *Journal of Applied Econometrics*. 35(3), 328–343.

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

Hecq, A. and Velasquez-Gaviria, D. 2025. Non-causal and non-invertible ARMA models: Identification, estimation and application in equity portfolios. *Journal of Time Series Analysis*. 46(2), 325–352.

Knight, J. L., Yu, J. (2002). Empirical characteristic function in time series estimation. *Econometric Theory*, 18(3), 691-721.

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

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

Lux, T., and Sornette, D. (2002). On rational bubbles and fat tails. *Journal of Money, Credit, and Banking*, 34(3), 589-610.

Rootzén, H. (1978): Extremes of moving averages of stable processes. *The Annals of Probability*, 847-869.

Tirole, J. 1985. Asset bubbles and overlapping generations. *Econometrica*, 1499-1528.

Politis, D.N. and Romano, J.P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics*, 22(4), 2031–2050.

Politis, D.N., Romano, J.P. and Wolf, M. (1999). *Subsampling*. Springer Series in Statistics, Springer-Verlag, New York.

Velasco, C. and Lobato, I. N. 2018. Frequency domain minimum distance inference for possibly noninvertible and noncausal ARMA models. *The Annals of Statistics*. 46(2), 555–579.

Xu, D. and Knight, J. (2010). Continuous empirical characteristic function estimation of mixtures of normal parameters. *Econometric Reviews*, 30(1) 25–50.

A. Proofs

A.1. Proof of Lemma 2.1

We first establish the $C^k(\Theta)$ regularity of (2.14), the MDE objective function. The proof proceeds by analyzing the theoretical characteristic function structure and establishing precise control over its derivatives under Assumptions 1 and 2. We then show that under the condition obtained to insure that (2.14) belongs to the $C^2(\Theta)$ class, Assumptions 3, 6, 7 and 8 are satisfied. For simplicity, the proof is only developed for the MAR(0,1) case although it also holds for the MAR(1,1) case.

A.1.1. $C^k(\Theta)$ regularity and validation of Assumption 3

Recall that for the α-stable MAR(0,1) component, the variable $uX_{j,t} + vX_{j,t+1}$ decomposes into two independent parts: $(\rho_j u + v)X_{j,t+1}$ and $u\varepsilon_{j,t}$. The joint log-characteristic function is the sum of their log-characteristic functions. Recalling that $\omega(\alpha, x) = \tan(\pi\alpha/2)$ if $\alpha \neq 1$ and $w(1, x) = -\frac{2}{\pi} \ln |x|$ we have

$$\log \varphi_{X_j}(u, v; \theta) = -\frac{\sigma^\alpha}{1 - |\rho_j|^\alpha} |\rho_j u + v|^\alpha \mathcal{A}(\rho_j u + v) - \sigma^\alpha |u|^\alpha \mathcal{A}(u), \tag{A.1}$$

where $\mathcal{A}(x) = 1 - i\beta \operatorname{sign}(x)\omega(\alpha, x)$. Let $K \subset \Theta$ be any compact subset satisfying the uniform bounds: $\inf_{j,\theta\in K}(1 - |\rho_j|) \geq \delta' > 0$, $\inf_{\theta\in K} \alpha \geq \alpha_0 > 0$, $\sup_{\theta\in K} \sigma \leq M < \infty$, and $\sup_{\theta\in K} |\beta| \leq B < \infty$. From these assumptions, we can establish a uniform lower bound for $1 - |\rho_j|^\alpha$. Since $|\rho_j| \leq 1 - \delta'$, it follows that $1 - |\rho_j|^\alpha \geq 1 - (1 - \delta')^\alpha$, which is increasing in α (since $1 - \delta' \in (0, 1)$). Therefore, its minimum value over K is attained at α_0. We can thus define a single constant $\delta = 1 - (1 - \delta')^{\alpha_0} > 0$, which ensures that for all $\theta \in K$, we have $1 - |\rho_j|^\alpha \geq \delta$.

(ι) The derivative with respect to π_k is computed from the decomposition $\log \varphi(u, v; \theta) = \sigma^\alpha \sum_{j=1}^J \pi_j^\alpha \tilde{\varphi}_j(u, v)$, where $\tilde{\varphi}_j$ is the standardized log-characteristic function (i.e., the expression in (A.1) divided by σ^α). We have:

$$\frac{\partial \varphi}{\partial \pi_k}(u, v; \theta) = \varphi(u, v; \theta) \cdot \alpha \pi_k^{\alpha-1} \sigma^\alpha \tilde{\varphi}_k(u, v).$$

Substituting the explicit form $\tilde{\varphi}_k(u, v) = -\left[\frac{|\rho_k u+v|^\alpha \mathcal{A}(\rho_k u+v)}{1-|\rho_k|^\alpha} + |u|^\alpha \mathcal{A}(u)\right]$ and using the uniform bound $|\mathcal{A}(\cdot)| \leq 1 + B|\tan(\pi\alpha_{\max}/2)| := M_{\mathcal{A}}$ on the compact set K, we obtain:

$$\begin{aligned}\left|\frac{\partial \varphi}{\partial \pi_k}(u, v; \theta)\right| &\leq |\varphi(u, v; \theta)| \cdot \alpha \pi_k^{\alpha-1} \sigma^\alpha M_{\mathcal{A}} \left[\frac{|\rho_k u + v|^\alpha}{\delta} + |u|^\alpha\right] \\ &= C_\pi G_\pi(u, v),\end{aligned}$$

where C_π is a constant depending on K. The bounding function $G_\pi(u, v)$ grows polynomially (degree α) and is integrable against $w(u, v)$ for any $\alpha > 0$.

($\iota\iota$) The derivative with respect to α is computed using the decomposition of the log-characteristic function in (A.1). We have

$$\begin{aligned}\frac{\partial \log \varphi_{X_j}}{\partial \alpha} &= -\sigma^\alpha \ln \sigma \left[\frac{|\rho_j u + v|^\alpha \mathcal{A}(\rho_j u + v)}{1 - |\rho_j|^\alpha} + |u|^\alpha \mathcal{A}(u) \right] \\ &\quad - \sigma^\alpha \left[\frac{|\rho_j u + v|^\alpha \ln |\rho_j u + v|}{1 - |\rho_j|^\alpha} \mathcal{A}(\rho_j u + v) + |u|^\alpha \ln |u| \mathcal{A}(u) \right] \\ &\quad + \sigma^\alpha \left[\frac{|\rho_j u + v|^\alpha |\rho_j|^\alpha \ln |\rho_j|}{(1 - |\rho_j|^\alpha)^2} \right] \mathcal{A}(\rho_j u + v) \\ &\quad - \sigma^\alpha \left[\frac{|\rho_j u + v|^\alpha}{1 - |\rho_j|^\alpha} \frac{\partial \mathcal{A}(\rho_j u + v)}{\partial \alpha} + |u|^\alpha \frac{\partial \mathcal{A}(u)}{\partial \alpha} \right].\end{aligned}$$

Using the uniform bounds on the compact set K (specifically $1 - |\rho_j|^\alpha \geq \delta$ and the boundedness of $\mathcal{A}$ and $\partial_\alpha \mathcal{A}$), we obtain the following majoration

$$\begin{aligned}\left| \frac{\partial \log \varphi_{X_j}}{\partial \alpha} \right| &\leq C_1 \left[\frac{|\rho_j u + v|^\alpha}{\delta} + |u|^\alpha \right] \\ &\quad + C_2 \left[\frac{|\rho_j u + v|^\alpha |\ln |\rho_j u + v||}{\delta} + |u|^\alpha |\ln |u|| \right],\end{aligned}$$

where C_1 and C_2 are finite constants depending only on K. This leads to a bounding function of the form

$$\left| \frac{\partial \varphi}{\partial \alpha}(u, v; \theta) \right| \leq C_\alpha |\varphi(u, v; \theta)| \left[\frac{H_\alpha(\rho_j u + v)}{\delta} + H_\alpha(u) \right] = C_\alpha G_\alpha(u, v),$$

where $H_\alpha(x) = |x|^\alpha (1 + |\ln |x||)$. To conclude on the integrability of $G_\alpha(u, v)$ against $w(u, v)$, we use a continuity argument. Since $\alpha \geq \alpha_0 > 0$ on K, the function $x \mapsto H_\alpha(x)$ is continuous on $\mathbb{R}$ (prolonged by 0 at $x = 0$ since $\lim_{x \to 0} |x|^\alpha \ln |x| = 0$). Consequently, $H_\alpha(x)$ is bounded on any compact set and grows polynomially at infinity. At this stage, we need Assumption 2 as it imposes $w(u, v) = \exp(-\kappa(u^2 + v^2))$. As a consequence, the growth is dominated by the exponential decay of $w(u, v)$, ensuring that $\int G_\alpha(u, v) w(u, v) du dv < \infty$.

($\iota\iota\iota$) The derivative with respect to σ is computed by noting that $\log \varphi_{X_j}(u, v; \theta) = \sigma^\alpha \tilde{\varphi}_j(u, v)$. We have

$$\begin{aligned}\frac{\partial \log \varphi_{X_j}}{\partial \sigma} &= \alpha \sigma^{\alpha - 1} \tilde{\varphi}_j(u, v) \\ &= -\alpha \sigma^{\alpha - 1} \left[\frac{|\rho_j u + v|^\alpha \mathcal{A}(\rho_j u + v)}{1 - |\rho_j|^\alpha} + |u|^\alpha \mathcal{A}(u) \right].\end{aligned}$$

Summing over j (weighted by π_j^α) and applying the uniform bounds on the compact set K (specifically $|\mathcal{A}(\cdot)| \leq M_{\mathcal{A}}$), we obtain

$$\begin{aligned}\left| \frac{\partial \varphi}{\partial \sigma}(u, v; \theta) \right| &\leq |\varphi(u, v; \theta)| \sum_{j=1}^J \pi_j^\alpha \left| \frac{\partial \log \varphi_{X_j}}{\partial \sigma} \right| \\ &\leq |\varphi(u, v; \theta)| \cdot \alpha M_\sigma M_{\mathcal{A}} \sum_{j=1}^J \pi_j^\alpha \left[\frac{|\rho_j u + v|^\alpha}{\delta} + |u|^\alpha \right] \\ &= C_\sigma G_\sigma(u, v),\end{aligned}$$

where $M_\sigma = \max(M^{\alpha-1}, \sigma_{\min}^{\alpha-1})$ is the uniform bound for $\sigma^{\alpha-1}$ on K. Note that when $\alpha \geq 1$, we have $M_\sigma = M^{\alpha-1}$, while for $\alpha < 1$, we have $M_\sigma = \sigma_{\min}^{\alpha-1}$ since $\sigma^{\alpha-1}$ is decreasing in σ when $\alpha - 1 < 0$. The bounding function $G_\sigma(u,v)$ is a finite sum of terms with polynomial growth of degree α, which is integrable against $w(u,v)$ for any $\alpha > 0$.

($\iota\nu$) The derivative with respect to β is obtained by differentiating the asymmetry terms $\mathcal{A}(x) = 1 - i\beta\text{sign}(x)\omega(\alpha)$ within the log-characteristic function expression (A.1). We have $\partial_\beta \mathcal{A}(x) = -i\text{sign}(x)\omega(\alpha, x)$ for $\alpha \neq 1$. Thus,

$$\begin{aligned}
\frac{\partial \log \varphi_{X_j}}{\partial \beta} &= -\frac{\sigma^\alpha}{1-|\rho_j|^\alpha}|\rho_j u + v|^\alpha \frac{\partial \mathcal{A}(\rho_j u + v)}{\partial \beta} - \sigma^\alpha |u|^\alpha \frac{\partial \mathcal{A}(u)}{\partial \beta} \\
&= i\sigma^\alpha \omega(\alpha) \left[\frac{|\rho_j u + v|^\alpha \text{sign}(\rho_j u + v)}{1-|\rho_j|^\alpha} + |u|^\alpha \text{sign}(u) \right].
\end{aligned}$$

The derivative of the full characteristic function is $\frac{\partial \varphi}{\partial \beta} = \varphi \sum_{j=1}^J \pi_j^\alpha \frac{\partial \log \varphi_{X_j}}{\partial \beta}$. Taking the modulus and applying the triangle inequality along with the uniform bounds on the compact set K (specifically $|\text{sign}(\cdot)| \leq 1$ and $|\omega(\alpha)| \leq \sup_{\alpha \in K} |\tan(\pi\alpha/2)| := M_\omega$), we obtain

$$\begin{aligned}
\left| \frac{\partial \varphi}{\partial \beta}(u,v;\theta) \right| &\leq |\varphi(u,v;\theta)| \sum_{j=1}^J \pi_j^\alpha \sigma^\alpha M_\omega \left[\frac{|\rho_j u + v|^\alpha}{\delta} + |u|^\alpha \right] \\
&\leq C_\beta |\varphi(u,v;\theta)| \left[\sum_{j=1}^J \frac{|\rho_j u + v|^\alpha}{\delta} + J|u|^\alpha \right].
\end{aligned}$$

The functional form of this bound is identical to that found for the derivative with respect to σ (a polynomial of degree α in u, v multiplied by the characteristic function). Consequently, it is integrable against the exponential weight $w(u,v)$ for any $\alpha > 0$. Notice that when $\alpha = 1$, the simple polynomial bound no longer holds due to the logarithmic term in the characteristic function, but the integrability is preserved by the exponential decay of $w(u,v) = \exp(-\kappa(u^2+v^2))$, utilizing the same argument as for the α derivative in ($\iota\iota$).

(ν) Finally, we turn to the most critical case, the derivative with respect to ρ_k. Using the decomposition in (A.1), the derivative is given by

$$\frac{\partial \log \varphi_{X_k}}{\partial \rho_k} = -\sigma^\alpha \frac{\partial}{\partial \rho_k} \left[\frac{|\rho_k u + v|^\alpha}{1-|\rho_k|^\alpha} \right] \mathcal{A}(\rho_k u + v) - \frac{\sigma^\alpha |\rho_k u + v|^\alpha}{1-|\rho_k|^\alpha} \frac{\partial \mathcal{A}(\rho_k u + v)}{\partial \rho_k}.$$

The derivative of the asymmetry term $\mathcal{A}(x) = 1 - i\beta\text{sign}(x)\omega(\alpha)$ involves the derivative of the sign function, which is zero almost everywhere (the Dirac mass contribution on the line $v = -\rho_k u$ does not affect the L^1 integrability). Thus, the second term vanishes almost everywhere. The dominant behavior comes from the first term

$$\frac{\partial}{\partial \rho_k} \left[\frac{|\rho_k u + v|^\alpha}{1-|\rho_k|^\alpha} \right] = \frac{\alpha u \, \text{sign}(\rho_k u + v)|\rho_k u + v|^{\alpha-1}}{1-|\rho_k|^\alpha} + \frac{|\rho_k u + v|^\alpha \alpha |\rho_k|^{\alpha-1} \text{sign}(\rho_k)}{(1-|\rho_k|^\alpha)^2}.$$

Using the uniform bounds on the compact set K (specifically $|\mathcal{A}(\cdot)| \leq M_{\mathcal{A}}$), we define the bound for the singular part:

$$T_1(u,v) = \frac{\alpha|u||\rho_k u + v|^{\alpha-1}}{\delta} M_{\mathcal{A}}.$$

Again, we need Assumption 2 and impose $w(u,v) = \exp(-\kappa(u^2+v^2))$, for $\kappa > 0$, to prove the convergence. We rely on the polar coordinates of the integral of T_1:

$$\begin{aligned}
I_1 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T_1(u,v)\exp(-\kappa(u^2+v^2))\,du\,dv \\
&= \frac{\alpha M_{\mathcal{A}}}{\delta}\int_0^{2\pi}\int_0^{\infty} r|\cos\theta|\cdot r^{\alpha-1}|\rho_k\cos\theta+\sin\theta|^{\alpha-1}e^{-\kappa r^2}r\,dr\,d\theta \\
&= \frac{\alpha M_{\mathcal{A}}}{\delta}\int_0^{2\pi}|\cos\theta||\rho_k\cos\theta+\sin\theta|^{\alpha-1}\left(\int_0^{\infty} r^{\alpha+1}e^{-\kappa r^2}\,dr\right)d\theta,
\end{aligned}$$

with $u = r\cos\theta$, $v = r\sin\theta$. This decomposition reveals that the radial integral converges for $\alpha > -2$:

$$\int_0^{\infty} r^{\alpha+1}e^{-\kappa r^2}\,dr = \frac{\Gamma((\alpha+2)/2)}{2\kappa^{(\alpha+2)/2}}.$$

The angular integral, near singularities θ_0 where $\rho_k\cos\theta + \sin\theta = 0$, converges for $\alpha > 0$:

$$\int_{\theta_0-\epsilon}^{\theta_0+\epsilon}|\rho_k\cos\theta+\sin\theta|^{\alpha-1}d\theta \sim \int_{-\epsilon}^{\epsilon}|C\tau|^{\alpha-1}d\tau = \frac{2C^{\alpha-1}\epsilon^{\alpha}}{\alpha} < \infty,$$

with $\epsilon > 0$ an arbitrary small constant and $C = \sqrt{1+\rho_k^2}$. Therefore, for any $\alpha \in (0,2)$, the derivative is bounded by an integrable function:

$$\left|\frac{\partial\varphi}{\partial\rho_k}(u,v;\theta)\right| \leq C_\rho|\varphi(u,v;\theta)|\left[\frac{|u||\rho_k u+v|^{\alpha-1}}{\delta} + \frac{|\rho_k u+v|^{\alpha}}{\delta^2} + |u|^{\alpha}\right] = C_\rho G_\rho(u,v), \tag{A.2}$$

where C_ρ is a suitable constant. All terms in $G_\rho(u,v)$ are integrable against $w(u,v)$ for $\alpha > 0$. This completes the first-order derivative analysis with precise bounds, establishing uniform integrability that enables application of the dominated convergence theorem for $C^1(\Theta)$ regularity when $\alpha > 0$.

Finally, we analyze the second derivatives to establish $C^2(\Theta)$ regularity. The most critical terms arise from the second derivative with respect to ρ_k, specifically from the modulus term $|\rho_k u + v|^\alpha$. Using the decomposition in (A.1), we have

$$\frac{\partial^2\log\varphi_{X_k}}{\partial\rho_k^2} = -\frac{\sigma^\alpha\mathcal{A}(\rho_k u+v)}{1-|\rho_k|^\alpha}\frac{\partial^2}{\partial\rho_k^2}|\rho_k u+v|^\alpha + R_k^{(2)}(u,v),$$

where $R_k^{(2)}(u,v)$ collects terms involving first derivatives of the modulus and derivatives of the coefficients, which are less singular. Specifically, $R_k^{(2)}(u,v) = O(|u|^\alpha + |v|^\alpha)$ as $||(u,v)|| \to \infty$ and is locally integrable. The dominant singular term is

$$\frac{\partial^2}{\partial\rho_k^2}|\rho_k u+v|^\alpha = \alpha(\alpha-1)u^2|\rho_k u+v|^{\alpha-2}.$$

The integrability of this term against $w(u,v)$ determines the $C^2(\Theta)$ regularity. Let us bound the integral of the modulus of this second derivative

$$T_2(u,v) = C|\varphi(u,v;\theta)|u^2|\rho_k u + v|^{\alpha-2}.$$

Using polar coordinates ($u = r\cos\theta, v = r\sin\theta$) and the exponential weight $w(u,v) = e^{-\kappa r^2}$, the integral becomes

$$\begin{aligned} I_2 &= \int_0^{2\pi}\int_0^{\infty} r^2\cos^2\theta \cdot r^{\alpha-2}|\rho_k\cos\theta + \sin\theta|^{\alpha-2}e^{-\kappa r^2} r\,dr\,d\theta \\ &= \left(\int_0^{\infty} r^{\alpha+1}e^{-\kappa r^2}\,dr\right)\int_0^{2\pi}\cos^2\theta|\rho_k\cos\theta+\sin\theta|^{\alpha-2}d\theta. \end{aligned}$$

The radial integral converges for $\alpha > -2$. The angular integral $J_\alpha = \int_0^{2\pi}\cos^2\theta|\rho_k\cos\theta+\sin\theta|^{\alpha-2}d\theta$ presents singularities when $\rho_k\cos\theta + \sin\theta = 0$. Let θ_0 be such a singularity. Locally, the integrand behaves like $|\theta-\theta_0|^{\alpha-2}$. Convergence requires

$$\int_{\theta_0-\epsilon}^{\theta_0+\epsilon}|\tau|^{\alpha-2}d\tau < \infty \iff \alpha - 2 > -1 \iff \alpha > 1.$$

For $\alpha \in (1,2)$, the angular integral is finite. The remaining terms in the second derivative of the objective function $D_{\mathcal{X}}(\theta)$ involve products of first derivatives (which are square-integrable for $\alpha > 0$) or the second derivative analyzed above. Thus, by the dominated convergence theorem, the objective function is $C^2(\Theta)$ if and only if $\alpha \in (1,2)$. Assumption 3 is satisfied under this condition.

A.1.2. Validation of Assumption 6

Assumption 6 requires the random sequence $K(x;\theta)$ defined in (2.17) to be measurable and bounded. Since trigonometric functions and the theoretical characteristic function $\varphi(u,v;\theta)$ are continuous (and thus measurable), the entire integrand in (2.17) is a measurable function of x for each fixed (u,v,θ). By the Fubini theorem, the integral of this function with respect to (u,v) is a measurable function of x. Next, we demonstrate that $K(x;\theta)$ is uniformly bounded with respect to x. From the natural bounds of trigonometric functions, $|\cos(ux_{j+1}+vx_j)| \le 1$ and $|\sin(ux_{j+1}+vx_j)| \le 1$, and since $|\varphi(u,v;\theta)| \le 1$, we have

$$\begin{aligned} |K(x;\theta)| &\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Bigg[(|\cos(ux_{j+1}+vx_j)| + |\mathrm{Re}\ \varphi(u,v;\theta)|)\left|\frac{\partial \mathrm{Re}\ \varphi(u,v;\theta)}{\partial\theta}\right| \\ &\qquad +(|\sin(ux_{j+1}+vx_j)| + |\mathrm{Im}\ \varphi(u,v;\theta)|)\left|\frac{\partial \mathrm{Im}\ \varphi(u,v;\theta)}{\partial\theta}\right|\Bigg] w(u,v)\,du\,dv \\ &\le 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\left|\frac{\partial \mathrm{Re}\ \varphi(u,v;\theta)}{\partial\theta}\right| + \left|\frac{\partial \mathrm{Im}\ \varphi(u,v;\theta)}{\partial\theta}\right|\right) w(u,v)\,du\,dv \\ &\le 2\sqrt{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\frac{\partial\varphi(u,v;\theta)}{\partial\theta}\right| w(u,v)\,du\,dv := B(\theta), \end{aligned}$$

where the last inequality follows from the Cauchy-Schwarz inequality: for any complex number $z = a+ib$, we have $|a|+|b| \le \sqrt{2}|z|$ since $(|a|+|b|)^2 \le 2(a^2+b^2) = 2|z|^2$.

The first-order analysis in Lemma 2.1 established that for each parameter component θ_i, the integral of the derivatives is finite, i.e.,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\frac{\partial\varphi(u,v;\theta)}{\partial\theta_i}\right| w(u,v)\,du\,dv < \infty.$$

Furthermore, as established in Lemma 2.1, the objective function is $C^1(\Theta)$ and $C^2(\Theta)$ for $\alpha > 1$, which implies that the gradient $\partial\varphi/\partial\theta$ is continuous in θ. Consequently, the integral function $B(\theta)$ is continuous on the parameter space Θ. Since Θ is compact (Assumption 1), the continuous function $B(\theta)$ is bounded. We thus have

$$\sup_{\theta\in\Theta}\sup_{x}|K(x;\theta)| \le \sup_{\theta\in\Theta} B(\theta) < \infty.$$

This uniform boundedness ensures that Assumption 6 is satisfied. □

A.1.3. Validation of Assumption 7

This assumption requires $\Sigma(\theta_0)$ to be nonsingular and the second derivatives to be uniformly bounded. The boundedness follows directly from the $C^2(\Theta)$ regularity analysis in Lemma 2.1 for $\alpha \in (1,2)$. For the nonsingularity of $\Sigma(\theta_0)$, we interpret $\Sigma(\theta_0)$ as the Gram matrix of the score functions in the Hilbert space $L^2_w(\mathbb{R}^2)$. Its nonsingularity is equivalent to the linear independence of the components of the score vector $\nabla_\theta \log\varphi(u,v;\theta)$. Consistent with our estimation strategy, we consider the identifiable parameter vector $\theta = (\varsigma_1,\dots,\varsigma_J,\rho_1,\dots,\rho_J,\alpha,\beta)'$, where $\varsigma_j = \sigma\pi_j$. Recall that $\tilde{\varphi}_k(u,v) = \log\varphi_{X_k}(u,v)$ denotes the log-characteristic function of the k-th latent component with unit scale. The total log-characteristic function is $\log\varphi = \sum_{j=1}^{J}\varsigma_j^\alpha\tilde{\varphi}_j(u,v)$. The scores exhibit the following exact forms and asymptotic behaviors

$$\begin{aligned} g_{\rho_k}(u,v) &= -\varsigma_k^\alpha\frac{\partial}{\partial\rho_k}\left[\frac{|\rho_k u+v|^\alpha}{1-|\rho_k|^\alpha}\mathcal{A}_k(u,v)\right] \\ &= -\varsigma_k^\alpha\frac{\alpha u\,\mathrm{sign}(\rho_k u+v)|\rho_k u+v|^{\alpha-1}}{1-|\rho_k|^\alpha}\mathcal{A}_k(u,v) + R_k(u,v), \end{aligned}$$

where $\mathcal{A}_k(u,v) = 1 - i\beta\mathrm{sign}(\rho_k u+v)\tan(\pi\alpha/2)$. The remainder term satisfies $R_k(u,v) = O(|u|^\alpha + |v|^\alpha)$ but is dominated locally by the singular term as $\rho_k u + v \to 0$ (since $\alpha - 1 < \alpha$).

$$g_{\varsigma_k}(u,v) = \frac{\partial}{\partial\varsigma_k}(\varsigma_k^\alpha\tilde{\varphi}_k(u,v)) = \alpha\varsigma_k^{\alpha-1}\tilde{\varphi}_k(u,v).$$

$$g_\alpha(u,v) = \sum_{j=1}^{J}\varsigma_j^\alpha\left(-|u|^\alpha\ln|u|(1-i\beta\mathrm{sign}(u)\tan(\tfrac{\pi\alpha}{2})) + O(|u|^\alpha)\right) \quad (\text{as } |u|\to\infty).$$

$$g_\beta(u,v) = i\tan\left(\frac{\pi\alpha}{2}\right)\sum_{j=1}^{J}\varsigma_j^\alpha\left(\frac{|\rho_j u+v|^\alpha\mathrm{sign}(\rho_j u+v)}{1-|\rho_j|^\alpha} + |u|^\alpha\mathrm{sign}(u)\right).$$

Consider a linear combination $\sum c_i g_i(u,v) = 0$ holding almost everywhere for any constant coefficients c_i. Then,

(ι) The scores g_{ρ_k} exhibit lines of non-analyticity (singular first derivatives of the modulus) along the lines $v = -\rho_k u$. Since the ρ_k are distinct (see Definition 2.1), these lines do not overlap. The singularity of $|\rho_k u + v|^{\alpha-1}$ in g_{ρ_k} cannot be cancelled by any linear combination of the other scores, which are either smooth or singular on different lines. This implies $c_{\rho_k} = 0$ for all k.

($\iota\iota$) With $c_{\rho_k} = 0$, we are left with the scale scores g_{ς_k}. Each g_{ς_k} is proportional to the log-characteristic function of the k-th component. Due to the distinct ρ_k values, the functions $\tilde{\varphi}_k(u, v)$ exhibit different scalings in the (u, v) plane and are thus linearly independent. Therefore, $c_{\varsigma_k} = 0$.

($\iota\iota\iota$) The term g_α is the only one exhibiting a specific $|u|^\alpha \ln |u|$ asymptotic growth rate as $|u| \to \infty$ (others grow at rate $|u|^\alpha$), ensuring $c_\alpha = 0$.

(ιv) With $c_{\rho_k} = c_{\varsigma_k} = c_\alpha = 0$ established, the linear combination reduces to $c_\beta g_\beta = 0$ almost everywhere. Since g_β is not identically zero—indeed, it has the explicit form

$$g_\beta(u, v) = i \tan \frac{\pi\alpha}{2} \sum_{j=1}^{J} \varsigma_j^\alpha \left[\frac{|\rho_j u + v|^\alpha \operatorname{sign}(\rho_j u + v)}{1 - |\rho_j|^\alpha} + |u|^\alpha \operatorname{sign}(u) \right],$$

which is non-zero for generic (u, v)—we must have $c_\beta = 0$.

Since the weight function $w(u, v)$ is strictly positive, this pointwise linear independence implies the non-singularity of the Gram matrix $\Sigma(\theta_0)$. □

A.1.4. Validation of Assumption 8

Validation of Assumption 8, which is required to apply a central limit theorem for dependent processes, rests on demonstrating that the temporal dependence of the sequence $\{K_j\}$ decays sufficiently fast. We establish this result by showing that the aggregated process $(\mathcal{X}_t)$ is strongly mixing with geometric decay rates, a property that is inherited by the sequence $\{K_j\}$. Each latent process $(X_{j,t})$—whether it is a purely anticipative AR(1) or a mixed MAR(1,1)—admits a two-sided infinite moving average representation:

$$X_{j,t} = \sum_{k=-\infty}^{+\infty} d_{j,k} \varepsilon_{j,t-k},$$

where the coefficients $d_{j,k}$ decay geometrically as $|k| \to \infty$. Specifically, for the MAR(1,1) case, $|d_{j,k}| \leq D \max(|\phi_j|, |\psi_j|)^{|k|}$ for $D > 0$ a finite constant. Since the innovations $\varepsilon_{j,t}$ are i.i.d. with an absolutely continuous distribution (stable distributions with $\alpha \in (1, 2)$ possess smooth densities), the linear process $(X_{j,t})$ is strongly mixing (α-mixing) with mixing coefficients decaying geometrically (see e.g., Doukhan , 1994, Section 2.3). Since the latent processes are mutually independent, the σ-algebra generated by the aggregate $\mathcal{X}_t$ is contained in the σ-algebra generated by the vector of components. Consequently, the mixing coefficient of the aggregated process $\mathcal{X}_t$ satisfies an inequality of the form:

$$\alpha_{\mathcal{X}}(h) \leq \sum_{j=1}^{J} \alpha_{X_j}(h).$$

Given that each component has geometrically decaying mixing coefficients, there exist constants $C > 0$ and $\lambda \in (0, 1)$ such that $\alpha_{\mathcal{X}}(h) \le C\lambda^h$. Finally, the score term K_j defined in (2.17) is a measurable function of a finite number of lagged and lead values of the process $\mathcal{X}_t$ (specifically $\mathcal{X}_j$ and $\mathcal{X}_{j+1}$). Any measurable function of a finite segment of a strongly mixing process is itself strongly mixing with the same decay rate. This ensures the convergence of $\mathbb{E}(K_0|\mathcal{F}_{-m})$ to $\mathbb{E}(K_0) = 0$ in mean square as $m \to \infty$. Furthermore, for a bounded and geometrically mixing sequence, the norms of the projection differences ν_j are summable, satisfying the condition $\sum_{j=0}^{\infty} \mathbb{E}[\nu_j' \nu_j]^{1/2} < \infty$. Assumption 8 is thus satisfied. □

A.2. Proof of Lemma 3.1

Denote $\boldsymbol{X}_{j,t} = (X_{j,t-m}, \ldots, X_{j,t}, X_{j,t+1}, \ldots, X_{j,t+h})$ the paths of the moving averages $(X_{j,t})$, for $j = 1, \ldots, J$. The $\boldsymbol{X}_{j,t}$'s are independent α-stable random vectors with spectral representations $(\Gamma_j, \boldsymbol{\mu}_j^0)$. We consider only the more delicate case $\alpha = 1$ and $\beta_j \in [-1, 1]$ for $j = 1, \ldots, J$. Because of the independence between $\boldsymbol{X}_{1,t}, \ldots, \boldsymbol{X}_{J,t}$, we have with $a = 2/\pi$

$$\begin{aligned}
\mathbb{E}\Big[e^{i\langle \boldsymbol{u}, \boldsymbol{X_t}\rangle}\Big] &= \mathbb{E}\Big[e^{i\langle \boldsymbol{u}, \sigma \sum_{j=1}^J \pi_j \boldsymbol{X}_{j,t}\rangle}\Big] = \prod_{j=1}^J \mathbb{E}\Big[e^{i\langle \sigma\pi_j \boldsymbol{u}, \boldsymbol{X}_{j,t}\rangle}\Big] \\
&= \prod_{j=1}^J \exp\Bigg\{ - \int_{S_{m+h+1}} \Big(|\langle \sigma\pi_j \boldsymbol{u}, \boldsymbol{s}\rangle| + ia\langle \sigma\pi_j \boldsymbol{u}, \boldsymbol{s}\rangle \ln|\langle \sigma\pi_j \boldsymbol{u}, \boldsymbol{s}\rangle|\Big)\Gamma_j(d\boldsymbol{s}) + i\langle \sigma\pi_j \boldsymbol{u}, \boldsymbol{\mu_j}^0\rangle \Bigg\} \\
&= \exp\Bigg\{ - \int_{S_{m+h+1}} \Big(|\langle \boldsymbol{u}, \boldsymbol{s}\rangle| + ia\langle \boldsymbol{u}, \boldsymbol{s}\rangle \ln|\langle \boldsymbol{u}, \boldsymbol{s}\rangle|\Big) \sum_{j=1}^J \sigma^\alpha \pi_j^\alpha \Gamma_j(d\boldsymbol{s}) \\
&\qquad\qquad + i \sum_{j=1}^J \Big(\langle \boldsymbol{u}, \sigma\pi_j \boldsymbol{\mu_j}^0\rangle - a\sigma\pi_j \ln(\sigma\pi_j) \int_{S_{m+h+1}} \langle \boldsymbol{u}, \boldsymbol{s}\rangle \Gamma_j(d\boldsymbol{s}) \Big) \Bigg\}.
\end{aligned}$$

Focusing on the shift vector, we have

$$\sum_{j=1}^J \Big(\langle \boldsymbol{u}, \sigma\pi_j \boldsymbol{\mu_j}^0\rangle - a\sigma\pi_j \ln(\sigma\pi_j) \int_{S_{m+h+1}} \langle \boldsymbol{u}, \boldsymbol{s}\rangle \Gamma_j(d\boldsymbol{s}) \Big) = \langle \boldsymbol{u}, \sum_{j=1}^J \sigma\pi_j (\boldsymbol{\mu}_j^0 - a\ln(\sigma\pi_j)\tilde{\boldsymbol{\mu}}_j)\rangle,$$

with $\tilde{\boldsymbol{\mu}}_j = (\tilde{\mu}_{j,\ell})$ and $\tilde{\mu}_{j,\ell} = \int_{S_{m+h+1}} s_\ell \Gamma_j(d\boldsymbol{s})$, $\ell = -m, \ldots, 0, 1, \ldots, h$. Using the form of Γ_j, i.e., $\Gamma_j = \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \|\boldsymbol{d}_{j,k}\|_e \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|_e}\right\}}$, we get

$$\tilde{\mu}_{j,\ell} = \int_{S_{m+h+1}} s_\ell \Gamma_j(d\boldsymbol{s}) = \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \|\boldsymbol{d}_{j,k}\|_e \frac{\vartheta d_{j,k+\ell}}{\|\boldsymbol{d}_{j,k}\|_e} = \beta_j \sum_{k \in \mathbb{Z}} d_{j,k+\ell}, \quad \ell = -m, \ldots, h.$$

Hence, $\tilde{\boldsymbol{\mu}}_j = \beta_j \sum_{k \in \mathbb{Z}} \boldsymbol{d}_{j,k}$, and using the form of $\boldsymbol{\mu}_j^0$ as given in (3.5),

$$\begin{aligned}
\sum_{j=1}^J \sigma\pi_j (\boldsymbol{\mu}_j^0 - a\ln(\sigma\pi_j)\tilde{\boldsymbol{\mu}}_j) &= \sum_{j=1}^J \sigma\pi_j \Big(- a\beta_j \sum_{k \in \mathbb{Z}} \boldsymbol{d}_{j,k} \ln\|\boldsymbol{d}_{j,k}\|_e - a\ln(\sigma\pi_j)\beta_j \sum_{k \in \mathbb{Z}} \boldsymbol{d}_{j,k} \Big) \\
&= -a \sum_{j=1}^J \sum_{k \in \mathbb{Z}} \sigma\pi_j \beta_j \boldsymbol{d}_{j,k} \ln\|\sigma\pi_j \boldsymbol{d}_{j,k}\|_e \\
&:= \boldsymbol{\mu}^0.
\end{aligned}$$

Therefore,

$$\mathbb{E}\Big[e^{i\langle \boldsymbol{u},\boldsymbol{X}_t\rangle}\Big] = \exp\Bigg\{-\int_{S_{m+h+1}}\Big(|\langle \boldsymbol{u},\boldsymbol{s}\rangle| + ia\langle \boldsymbol{u},\boldsymbol{s}\rangle \ln|\langle \boldsymbol{u},\boldsymbol{s}\rangle|\Big)\sum_{j=1}^{J}\sigma^{\alpha}\pi_j^{\alpha}\Gamma_j(d\boldsymbol{s}) + i\langle \boldsymbol{u},\boldsymbol{\mu}^0\rangle\Bigg\},$$

and the random vector $\boldsymbol{X}_t$ is 1-stable with spectral measure

$$\sum_{j=1}^{J}\sigma^{\alpha}\pi_j^{\alpha}\Gamma_j = \sigma^{\alpha}\sum_{j=1}^{J}\sum_{\vartheta\in S_1}\sum_{k\in\mathbb{Z}} w_{j,\vartheta}\pi_j^{\alpha}\|\boldsymbol{d}_{j,k}\|_e^{\alpha}\delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|_e}\right\}},$$

and shift vector as announced in the lemma.

A.3. Proof of Lemma 3.2

With the usual notations, let the $\boldsymbol{X}_{j,t}$'s be the paths of the moving averages $(X_{j,t})$'s and let Γ_j, $j = 1,\dots,J$, their spectral measures on the Euclidean unit sphere. Let Γ be the spectral measure of $\boldsymbol{X}_t$. By Lemma 3.1, we have:

$$\Gamma = \sigma^{\alpha}\sum_{j=1}^{J}\pi_j^{\alpha}\Gamma_j.$$

Thus, by Proposition 1 of DFT, in the cases where either $\alpha \neq 1$ or $\boldsymbol{X}_t$ is symmetric, the vector $\boldsymbol{X}_t$ is representable on $C_{m+h+1}^{\|\cdot\|}$ if and only if

$$\begin{aligned}\Gamma(K^{\|\cdot\|}) = 0 \quad &\iff \quad \sigma^{\alpha}\sum_{j=1}^{J}\pi_j^{\alpha}\Gamma_j(K^{\|\cdot\|}) = 0\\ &\iff \quad \Gamma_j(K^{\|\cdot\|}) = 0, \ \ \forall\, j = 1,\dots,J,\end{aligned}$$

where the last equivalence follows from the fact that $\sigma^{\alpha} > 0$ and $\pi_j^{\alpha} > 0$ for all $j = 1,\dots,J$. Given that the Γ_j's are the spectral measures of paths of non-aggregated moving averages, we can apply the arguments from the proof of Theorem 1 in DFT. Specifically, for each j, the condition $\Gamma_j(K^{\|\cdot\|}) = 0$ is equivalent to the representability condition (3.4) holding for the sequence $(d_{j,k})_k$ with parameter m. Therefore, $\boldsymbol{X}_t$ is representable on $C_{m+h+1}^{\|\cdot\|}$ if and only if (3.4) holds with m for all sequences $(d_{j,k})_k$, $j = 1,\dots,J$. For the case $\alpha = 1$ and $\boldsymbol{X}_t$ asymmetric, we need to consider the additional condition involving the shift vector $\boldsymbol{\mu}^0$. From Lemma 3.1, we have:

$$\boldsymbol{\mu}^0 = -\mathbb{1}_{\{\alpha=1\}}\frac{2\sigma}{\pi}\sum_{j=1}^{J}\sum_{k\in\mathbb{Z}}\pi_j\beta_j \boldsymbol{d}_{j,k}\ln\|\sigma\pi_j\boldsymbol{d}_{j,k}\|_e.$$

By Proposition 1 of DFT, when $\alpha = 1$ and $\boldsymbol{X}_t$ is asymmetric, representability on $C_{m+h+1}^{\|\cdot\|}$ requires both:

1. $\Gamma(K^{\|\cdot\|}) = 0$, which as shown above is equivalent to (3.4) holding for all sequences $(d_{j,k})_k$;
2. The additional condition (3.6) must hold.

To verify condition (3.6), we need to show:

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

However, in the context of stable aggregates, this condition must be interpreted in terms of the aggregated coefficients. Since $\boldsymbol{X}_t = \sigma \sum_{j=1}^{J} \pi_j \boldsymbol{X}_{j,t}$, the effective coefficients are combinations of the individual sequences $(d_{j,k})_k$. The condition (3.6) in the aggregated case becomes:

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

where $\boldsymbol{d}_k$ now refers to the k-th vector in the aggregated representation. Given the linearity of the aggregation and the fact that the condition must hold for each component individually (as each $\boldsymbol{X}_{j,t}$ must satisfy the representability conditions), the condition (3.6) for the aggregate is satisfied if and only if it holds for all sequences $(d_{j,k})_k$, $j = 1, \dots, J$, with the same parameters m and h.

A.4. Proof of Proposition 3.1

If $\alpha \neq 1$, we have by Theorem 1 and the proof of Proposition 3 of DFT,

$$\begin{aligned}
(\mathcal{X}_t) \text{ past-representable} &\iff \exists\, m \geq 0,\ (3.4) \text{ holds with } m \text{ for all sequences } (d_{j,k})_k \\
&\iff \forall\, j = 1, \dots, J,\ m_{0,j} < +\infty \\
&\iff \forall\, j = 1, \dots, J,\ (X_{j,t}) \text{ past-representable.}
\end{aligned}$$

For a given series $(d_{j,k})_k$, (3.4) holds with $m \geq m_{0,j}$ and does not hold with $m < m_{0,j}$. Regarding the last statement, we know that for $(\mathcal{X}_t)$ (m,h)-past-representable, (3.4) holds with the same m for all the sequences $(d_{j,k})_k$, $j = 1, \dots, J$. This holds if $m \geq \max_j m_{0,j}$ and cannot hold if $m < \max_j m_{0,j}$. In the case where $\alpha = 1$, again by Theorem 1 of DFT and denoting generically by $\boldsymbol{X}_t$ a vector $(\mathcal{X}_{t-m}, \dots, \mathcal{X}_t, \mathcal{X}_{t+1}, \dots, \mathcal{X}_{t+h})$ of size $m+h+1$,

$$\begin{aligned}
&\mathcal{X}_t \text{ past-representable} \\
&\iff \exists\, m \geq 0, h \geq 1, \begin{cases} \boldsymbol{X}_t\ \mathcal{S}1\mathcal{S} \text{ and } (3.4) \text{ holds with } m \text{ for all sequences } (d_{j,k})_k \\ \text{or} \\ \boldsymbol{X}_t \text{ asymmetric and } (3.4)\text{-}(3.6) \text{ hold with } m, h \text{ for all sequences } (d_{j,k})_k \end{cases} \\
&\iff \forall\, j = 1, \dots, J,\ m_{0,j} < +\infty, \text{ and } \exists\, m \geq 0, h \geq 1, \begin{cases} \boldsymbol{X}_t\ \mathcal{S}1\mathcal{S} \\ \text{or} \\ \boldsymbol{X}_t \text{ asymmetric and } (3.6) \text{ hold} \\ \text{with } m, h \text{ for all sequences } (d_{j,k})_k \end{cases}
\end{aligned}$$

We conclude again by noting that the necessary condition (3.4) holds for $m \geq \max_j m_{0,j}$ and is violated for $m < \max_j m_{0,j}$. Now, for part $(\iota\iota)$, let $\|\cdot\|$ be a semi-norm satisfying (3.3) and assume that $\mathcal{X}_t$ is (m,h)-past-representable for some $m \geq 0$, $h \geq 1$. We need to establish the spectral representation of the vector $\boldsymbol{X_t} = (\mathcal{X}_{t-m}, \ldots, \mathcal{X}_t, \mathcal{X}_{t+1}, \ldots, \mathcal{X}_{t+h})$ on $C^{\|\cdot\|}_{m+h+1}$. From Lemma 3.1, we know that the spectral representation $(\Gamma, \boldsymbol{\mu}^0)$ of $\boldsymbol{X}_t$ on the Euclidean unit sphere S_{m+h+1} is given by:

$$\Gamma = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|_e^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|_e}\right\}} \tag{A.3}$$

$$\boldsymbol{\mu}^0 = \begin{cases} \boldsymbol{0}, & \text{if } \alpha \neq 1 \\ -\frac{2\sigma}{\pi} \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} \pi_j \beta_j \boldsymbol{d}_{j,k} \ln \|\sigma \pi_j \boldsymbol{d}_{j,k}\|_e, & \text{if } \alpha = 1 \end{cases}$$

To obtain the spectral representation on $C^{\|\cdot\|}_{m+h+1}$, we apply the transformation established in DFT for changing from Euclidean to semi-norm representations. By Lemma 3.2, since $\mathcal{X}_t$ is (m,h)-past-representable, the vector $\boldsymbol{X}_t$ is representable on $C^{\|\cdot\|}_{m+h+1}$. The transformation from the Euclidean representation to the semi-norm representation proceeds as follows. Let $K^{\|\cdot\|} := \{\boldsymbol{s} \in S_{m+h+1} : \|\boldsymbol{s}\| = 0\}$ be the kernel of the semi-norm on the Euclidean unit sphere. Since $\boldsymbol{X}_t$ is representable on $C^{\|\cdot\|}_{m+h+1}$, we have $\Gamma(K^{\|\cdot\|}) = 0$. Define the projection mapping $T_{\|\cdot\|} : S_{m+h+1} \setminus K^{\|\cdot\|} \to C^{\|\cdot\|}_{m+h+1}$ by:

$$T_{\|\cdot\|}(\boldsymbol{s}) = \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|} \tag{A.4}$$

By Proposition 2 of DFT, the spectral measure on the semi-norm unit cylinder is given by:

$$\Gamma^{\|\cdot\|}(A) = \int_{T^{-1}_{\|\cdot\|}(A)} \|\boldsymbol{s}\|_e^{-\alpha} \Gamma(d\boldsymbol{s}) \tag{A.5}$$

for any Borel set $A \subset C^{\|\cdot\|}_{m+h+1}$. Since the original spectral measure Γ from (A.3) is concentrated on atoms of the form $\{\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|_e\}$, and since $\|\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|_e\|_e = 1$, the transformation yields:

$$\Gamma^{\|\cdot\|}(A) = \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \sigma^\alpha \|\boldsymbol{d}_{j,k}\|_e^\alpha \cdot 1^{-\alpha} \cdot \mathbb{1}_A \left(\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right) \tag{A.6}$$

where we use the fact that $\|\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|_e\|_e = 1$ and $T_{\|\cdot\|}(\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|_e) = \vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$. Applying this transformation to (A.3), we obtain:

$$\Gamma^{\|\cdot\|} = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right\}} \tag{A.7}$$

For the shift vector in the case $\alpha = 1$, the transformation yields:

$$\boldsymbol{\mu}^{\|\cdot\|} = -\frac{2\sigma}{\pi} \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} \pi_j \beta_j \boldsymbol{d}_{j,k} \ln \|\sigma \pi_j \boldsymbol{d}_{j,k}\| \tag{A.8}$$

This completes the proof that the spectral representation $(\Gamma^{\|\cdot\|}, \boldsymbol{\mu}^{\|\cdot\|})$ of $\boldsymbol{X}_t$ on $C^{\|\cdot\|}_{m+h+1}$ is given by (3.5) with the Euclidean norm $\|\cdot\|_e$ replaced by the semi-norm $\|\cdot\|$, and with the scale parameter σ explicitly included in all relevant terms.

A.5. Proof of Corollary 3.1

The equivalence between $(\iota\iota)$ and $(\iota\iota\iota)$ follows from Corollary 2 of DFT. From the proof of the Corollary in DFT, we also know that, for any j, if $m_{0,j} < +\infty$, then (3.6) holds for the sequence $(d_{j,k})_k$ for any $m \geq m_{0,j}$. For the aggregated process $\mathcal{X}_t = \sigma \sum_{j=1}^{J} \pi_j X_{j,t}$ with $\sigma > 0$, the effective moving average coefficients for each component j become $\sigma\pi_j d_{j,k}$ rather than $d_{j,k}$. However, the past-representability conditions depend only on the pattern of zeros and non-zeros in the coefficient sequences, not on their scaling. Specifically, for condition (3.4), we require:

$$\forall k \in \mathbb{Z}, \quad \Big[(\sigma\pi_j d_{j,k+m}, \ldots, \sigma\pi_j d_{j,k}) = \mathbf{0} \quad \Longrightarrow \quad \forall \ell \leq k-1, \quad \sigma\pi_j d_{j,\ell} = 0\Big].$$

Since $\sigma > 0$ and $\pi_j > 0$ for all j, this is equivalent to:

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

Thus, the past-representability condition for the aggregated process is unchanged by the scaling factor σ. For the additional condition (3.6) when $\alpha = 1$ and the process is asymmetric, we need:

$$\sum_{k\in\mathbb{Z}} \|\sigma\pi_j \boldsymbol{d}_{j,k}\|_e \left| \ln\left(\|\sigma\pi_j \boldsymbol{d}_{j,k}\| / \|\sigma\pi_j \boldsymbol{d}_{j,k}\|_e\right)\right| < +\infty.$$

Since $\|\sigma\pi_j \boldsymbol{d}_{j,k}\|_e = \sigma\pi_j \|\boldsymbol{d}_{j,k}\|_e$ and the norm scales homogeneously, this becomes:

$$\sum_{k\in\mathbb{Z}} \sigma\pi_j \|\boldsymbol{d}_{j,k}\|_e \left| \ln\left(\|\boldsymbol{d}_{j,k}\| / \|\boldsymbol{d}_{j,k}\|_e\right)\right| < +\infty.$$

Since $\sigma\pi_j > 0$ is a finite constant, this condition is equivalent to:

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

which is precisely condition (3.6) for the unscaled sequences. Therefore:

$$\begin{aligned}\sup_j m_{0,j} < +\infty &\Longrightarrow \text{(3.6) holds for any sequence } (d_{j,k})_k \text{ for any } m \geq m_{0,j} \\ &\Longrightarrow \text{(3.6) holds for any sequence } (\sigma\pi_j d_{j,k})_k \text{ for any } m \geq \max_j m_{0,j}.\end{aligned}$$

Thus, $(\iota\iota\iota)$ implies (ι). The reciprocal is clear. Regarding the last statement, notice that if $\mathcal{X}_t$ is (m,h)-past-representable for some $m < \max_j m_{0,j}$, there would then exist some j such that $m < m_{0,j}$. Hence, (3.4) would not hold with m for the particular sequence $(\sigma\pi_j d_{j,k})_k$, which is impossible by Lemma 3.2, since the past-representability depends only on the zero pattern, not the scaling.

A.6. Proof of Proposition 3.2

By Proposition 2 of DFT, the asymptotic conditional tail property states that for any Borel sets $A, B \subset C^{\|\cdot\|}_{m+h+1}$ with $\Gamma^{\|\cdot\|}\Big(\partial(A\cap B)\Big) = \Gamma^{\|\cdot\|}\big(\partial B\big) = 0$, and $\Gamma^{\|\cdot\|}(B) > 0$,

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

Setting $B = B(V) = V \times \mathbb{R}^h$, we have

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

From Proposition 3.1 ($\iota\iota$), the spectral representation $(\Gamma^{\|\cdot\|}, \boldsymbol{\mu}^{\|\cdot\|})$ of the vector $\boldsymbol{X_t} = (\mathcal{X}_{t-m},\dots,\mathcal{X}_t,\mathcal{X}_{t+1},\dots,\mathcal{X}_{t+h})$ on $C^{\|\cdot\|}_{m+h+1}$ is given by equation (3.5) with the Euclidean norm $\|\cdot\|_e$ replaced by the semi-norm $\|\cdot\|$. From Lemma 3.1, the spectral measure can be written as:

$$\Gamma^{\|\cdot\|} = \sigma^\alpha \sum_{j=1}^{J}\sum_{\vartheta\in S_1}\sum_{k\in\mathbb{Z}} w_{j,\vartheta}\pi_j^\alpha \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right\}},$$

where $\boldsymbol{d}_{j,k} = (d_{j,k+m},\dots,d_{j,k},d_{j,k-1},\dots,d_{j,k-h})$, $w_{j,\vartheta} = (1+\vartheta\beta_j)/2$, and if $\boldsymbol{d}_{j,k} = \boldsymbol{0}$, the term vanishes by convention from the sums.

Now, we compute the numerator and denominator separately, we start by the numerator: $\Gamma^{\|\cdot\|}(A\cap B(V))$ Since $B(V) = V\times\mathbb{R}^h = \Big\{\boldsymbol{s}\in C^{\|\cdot\|}_{m+h+1}: \quad f(\boldsymbol{s})\in V\Big\}$, we have:

$$A\cap B(V) = \{\boldsymbol{s}\in A: \quad f(\boldsymbol{s})\in V\}.$$

The spectral measure $\Gamma^{\|\cdot\|}$ charges only the points of the form $\dfrac{\vartheta\boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}$ for $(\vartheta,j,k)\in S_1\times\{1,\dots,J\}\times\mathbb{Z}$. Therefore:

$$\begin{aligned}
\Gamma^{\|\cdot\|}(A\cap B(V)) &= \sigma^\alpha \sum_{j=1}^{J}\sum_{\vartheta\in S_1}\sum_{k\in\mathbb{Z}} w_{j,\vartheta}\pi_j^\alpha \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right\}}(A\cap B(V)) \\
&= \sigma^\alpha \sum_{\substack{(\vartheta,j,k):\\ \frac{\vartheta\boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\in A\cap B(V)}} w_{j,\vartheta}\pi_j^\alpha\|\boldsymbol{d}_{j,k}\|^\alpha \\
&= \sigma^\alpha \sum_{\substack{(\vartheta,j,k):\\ \frac{\vartheta\boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\in A \text{ and } \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|}\in V}} w_{j,\vartheta}\pi_j^\alpha\|\boldsymbol{d}_{j,k}\|^\alpha.
\end{aligned}$$

This can be written as:

$$\Gamma^{\|\cdot\|}(A\cap B(V)) = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta\boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\in A: \quad \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|}\in V\right\}\right).$$

For the denominator $\Gamma^{\|\cdot\|}(B(V))$, we proceed as follows:

$$\begin{aligned}\Gamma^{\|\cdot\|}(B(V)) &= \sigma^{\alpha} \sum_{\substack{(\vartheta,j,k):\\ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\in B(V)}} w_{j,\vartheta}\pi_j^{\alpha}\|\boldsymbol{d}_{j,k}\|^{\alpha}\\ &= \sigma^{\alpha} \sum_{\substack{(\vartheta,j,k):\\ \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|}\in V}} w_{j,\vartheta}\pi_j^{\alpha}\|\boldsymbol{d}_{j,k}\|^{\alpha}.\end{aligned}$$

This can be written as:

$$\Gamma^{\|\cdot\|}(B(V)) = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \in C^{\|\cdot\|}_{m+h+1} : \ \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} \in V\right\}\right).$$

Note that the factor σ^{α} appears in both the numerator and denominator, and therefore cancels out in the ratio:

$$\begin{aligned}\frac{\Gamma^{\|\cdot\|}(A\cap B(V))}{\Gamma^{\|\cdot\|}(B(V))} &= \frac{\sigma^{\alpha}\sum_{\substack{(\vartheta,j,k):\\ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\in A \text{ and } \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|}\in V}} w_{j,\vartheta}\pi_j^{\alpha}\|\boldsymbol{d}_{j,k}\|^{\alpha}}{\sigma^{\alpha}\sum_{\substack{(\vartheta,j,k):\\ \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|}\in V}} w_{j,\vartheta}\pi_j^{\alpha}\|\boldsymbol{d}_{j,k}\|^{\alpha}}\\ &= \frac{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \in A : \ \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} \in V\right\}\right)}{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \in C^{\|\cdot\|}_{m+h+1} : \ \frac{\vartheta f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} \in V\right\}\right)}.\end{aligned}$$

This establishes the desired result. The conclusion follows by considering the points of $B(V)$ and $A \cap B(V)$ that are charged by the spectral measure $\Gamma^{\|\cdot\|}$ given in equation (3.12). The presence of the scale parameter σ^{α} does not affect the asymptotic conditional probabilities as it appears multiplicatively in both the numerator and denominator of the ratio, thus canceling out in the final expression.

A.7. Proof of Lemma 3.3

By Proposition 3.1 and setting general scale parameter $\sigma > 0$, we have

$$\Gamma^{\|\cdot\|} = \sum_{j=1}^{J}\sum_{\vartheta\in S_1}\sum_{k\in\mathbb{Z}} w_{j,\vartheta}\sigma^{\alpha}\pi_j^{\alpha}\|\boldsymbol{d}_{j,k}\|^{\alpha}\delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right\}},$$

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

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

Therefore,

$$\Gamma^{\|\cdot\|} = \sum_{j=1}^{J} \sum_{\vartheta \in S_1} w_{j,\vartheta} \sigma^\alpha \pi_j^\alpha \left[\sum_{k=-m}^{h-1} \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{ \frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \right\}} + \sum_{k=h}^{+\infty} |\rho_j|^{\alpha(k-h)} \|\boldsymbol{d}_{j,h}\|^\alpha \delta_{\left\{ \frac{\vartheta \rho_j^{k-h} \boldsymbol{d}_{j,h}}{|\rho_j|^{k-h} \|\boldsymbol{d}_{j,h}\|} \right\}} \right].$$

Moreover,

$$\begin{aligned} \sum_{j=1}^{J} \sum_{\vartheta \in S_1} w_{j,\vartheta} \sigma^\alpha \pi_j^\alpha \sum_{k=h}^{+\infty} |\rho_j|^{\alpha(k-h)} \|\boldsymbol{d}_{j,h}\|^\alpha \delta_{\left\{ \mathrm{sign}(\rho_j)^{k-h} \frac{\vartheta \boldsymbol{d}_{j,h}}{\|\boldsymbol{d}_{j,h}\|} \right\}} \\ = \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sigma^\alpha \pi_j^\alpha \|\boldsymbol{d}_{j,h}\|^\alpha \frac{1}{2} \left[\sum_{k=h}^{+\infty} |\rho_j|^{\alpha(k-h)} + \vartheta \beta_j \sum_{k=h}^{+\infty} (\rho_j^{<\alpha>})^{k-h} \right] \delta_{\left\{ \frac{\vartheta \boldsymbol{d}_{j,h}}{\|\boldsymbol{d}_{j,h}\|} \right\}} \\ = \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sigma^\alpha \pi_j^\alpha \frac{1}{1-|\rho_j|^\alpha} \|\boldsymbol{d}_{j,h}\|^\alpha \bar{w}_{j,\vartheta} \delta_{\left\{ \frac{\vartheta\, \boldsymbol{d}_{j,h}}{\|\boldsymbol{d}_{j,h}\|} \right\}}. \end{aligned}$$

Finally, noticing that for $k=-m$ and any $j \in \{1,\dots,J\}$, $\boldsymbol{d}_{j,k} = (1,0,\dots,0)$,

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

where we have used the definition $w_\vartheta = \sum_{j=1}^{J} \sigma^\alpha \pi_j^\alpha w_{j,\vartheta}$.

A.8. Proof of Proposition 3.3

Lemma A.1. *Let $\Gamma^{\|\cdot\|}$ be the spectral measure given in Lemma 3.3 with $\sigma > 0$ and assume that the ρ_j's are all positive. Letting $(\vartheta_0, j_0, k_0) \in \mathcal{I}$, consider*

$$I_0 := \left\{ \frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} : \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|} \;\; \text{for} \;\; (\vartheta', j', k') \in \mathcal{I} \right\}.$$

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

$$I_0 = \left\{ \frac{\vartheta_0 \boldsymbol{d}_{j_0,k'}}{\|\boldsymbol{d}_{j_0,k'}\|} : \quad 0 \leq k' \leq h \right\}.$$

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

$$I_0 = \begin{cases} \left\{ \frac{\vartheta_0 \boldsymbol{d}_{j_0,k_0}}{\|\boldsymbol{d}_{j_0,k_0}\|} \right\}, & \text{if} \quad -m+1 \leq k_0 \leq -1 \\ \left\{ \frac{\vartheta_0 \boldsymbol{d}_{0,k_0}}{\|\boldsymbol{d}_{0,k_0}\|} \right\} = \{(\vartheta_0, 0, \dots, 0)\}, & \text{if} \quad k_0 = -m. \end{cases}$$

For $m=0$, then

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

Proof. The key observation is that the parameter $\sigma>0$ appears as a multiplicative factor in the spectral measure $\Gamma^{\|\cdot\|}$ but does **not** affect the normalized directions $\vartheta' \boldsymbol{d}_{j',k'}/\|\boldsymbol{d}_{j',k'}\|$ or their projections $\vartheta' f(\boldsymbol{d}_{j',k'})/\|\boldsymbol{d}_{j',k'}\|$. This is because σ only scales the overall magnitude of the spectral measure but does not change the geometric structure of the charged points on the unit cylinder. More precisely, from Lemma 3.3, the spectral measure takes the form:

$$\Gamma^{\|\cdot\|}=\sigma^\alpha\sum_{\vartheta\in S_1}\left[w_\vartheta\delta_{\{(\vartheta,0,\dots,0)\}}+\sum_{j=1}^J\pi_j^\alpha\left(w_{j,\vartheta}\sum_{k=-m+1}^{h-1}\|\boldsymbol{d}_{j,k}\|^\alpha\delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right\}}+\frac{\bar{w}_{j,\vartheta}}{1-|\rho_j|^\alpha}\|\boldsymbol{d}_{j,h}\|^\alpha\delta_{\left\{\frac{\vartheta\,\boldsymbol{d}_{j,h}}{\|\boldsymbol{d}_{j,h}\|}\right\}}\right)\right],$$

The factor σ^α multiplies the entire spectral measure uniformly, but the support of $\Gamma^{\|\cdot\|}$ (i.e., the set of points where $\Gamma^{\|\cdot\|}$ assigns positive mass) consists exactly of the normalized directions:

$$\operatorname{supp}(\Gamma^{\|\cdot\|})=\left\{(\vartheta,0,\dots,0),\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}:\vartheta\in S_1, j\in\{1,\dots,J\}, k\in\{-m+1,\dots,h\}\right\}$$

Since the condition defining I_0 involves only the equality of normalized projections:

$$\frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|}=\frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}$$

and since these normalized directions are independent of σ, the analysis proceeds exactly as in the case $\sigma=1$.

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

If $k'\in\{-m,\dots,-1\}$, the $(m+1)$-th component of $f(\boldsymbol{d}_{j',k'})$ is zero, whereas the $(m+1)$-th component of $f(\boldsymbol{d}_{j_0,k_0})$ is $\rho_{j_0}^{k_0}\neq 0$. This geometric relationship is unaffected by σ.

Necessarily, $\vartheta' f(\boldsymbol{d}_{j',k'})/\|\boldsymbol{d}_{j',k'}\|\neq\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\|$ and

$$I_0=\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|}: \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|}=\frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}\ \text{ for }\ (\vartheta',j',k')\in\{-1,+1\}\times\{1,\dots,J\}\times\{0,\dots,h\}\right\}.$$

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

$$f(\boldsymbol{d}_{j',k'})=(\rho_{j'}^{k'+m},\dots,\rho_{j'}^{k'+1},\rho_{j'}^{k'}),$$
$$f(\boldsymbol{d}_{j_0,k_0})=(\rho_{j_0}^{k_0+m},\dots,\rho_{j_0}^{k_0+1},\rho_{j_0}^{k_0}),$$

and by (3.3) we also have that

$$\|\boldsymbol{d}_{j',k'}\|=\|(\rho_{j'}^{k'+m},\dots,\rho_{j'}^{k'+1},\rho_{j'}^{k'},\overbrace{0,\dots,0}^{h})\|,$$
$$\|\boldsymbol{d}_{j_0,k_0}\|=\|(\rho_{j_0}^{k_0+m},\dots,\rho_{j_0}^{k_0+1},\rho_{j_0}^{k_0},\underbrace{0,\dots,0}_{h})\|.$$

The key observation is that these norms and the resulting normalized directions are independent of σ. Thus,

$$\begin{aligned}
\frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|} \\
&\iff \frac{\vartheta' \rho_{j'}^{k'} f(\boldsymbol{d}_{j',0})}{|\rho_{j'}|^{k'}\|\boldsymbol{d}_{j',0}\|} = \frac{\vartheta_0 \rho_{j_0}^{k_0} f(\boldsymbol{d}_{j_0,0})}{|\rho_{j_0}|^{k_0}\|\boldsymbol{d}_{j_0,0}\|} \\
&\iff \frac{\vartheta' \rho_{j'}^{\ell}}{\|\boldsymbol{d}_{j',0}\|} = \frac{\vartheta_0 \rho_{j_0}^{\ell}}{\|\boldsymbol{d}_{j_0,0}\|}, \qquad \ell = 0,\dots,m \\
&\iff \vartheta'\vartheta_0 \frac{\|\boldsymbol{d}_{j_0,0}\|}{\|\boldsymbol{d}_{j',0}\|} = \left(\frac{\rho_{j_0}}{\rho_{j'}}\right)^{\ell}, \quad \ell = 0,\dots,m \\
&\iff \rho_{j'} = \rho_{j_0} \quad \text{and} \quad \vartheta'\vartheta_0 = 1 \\
&\iff j' = j_0 \quad \text{and} \quad \vartheta' = \vartheta_0,
\end{aligned}$$

because the ρ_j's are assumed to be non-zero and distinct.

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

$$\frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|} \Longrightarrow k' = k_0.$$

$$\begin{aligned}
f(\boldsymbol{d}_{j',k'}) &= (\overbrace{\rho_{j'}^{k'+m}, \dots, \rho_{j'}, 1, 0, \dots, 0}^{m+1}), \\
f(\boldsymbol{d}_{j_0,k_0}) &= (\underbrace{\rho_{j_0}^{k_0+m}, \dots, \rho_{j_0}, 1, 0, \dots, 0}_{m+1}),
\end{aligned}$$

and we also have that

$$\begin{aligned}
\|\boldsymbol{d}_{j',k'}\| &= \|(\overbrace{\rho_{j'}^{k'+m}, \dots, \rho_{j'}, 1, 0, \dots, 0}^{m+1}, \overbrace{0, \dots, 0}^{h})\|, \\
\|\boldsymbol{d}_{j_0,k_0}\| &= \|(\underbrace{\rho_{j_0}^{k_0+m}, \dots, \rho_{j_0}, 1, 0, \dots, 0}_{m+1}, \underbrace{0, \dots, 0}_{h})\|.
\end{aligned}$$

As $k' = k_0 \leq -1$, the condition becomes:

$$\begin{aligned}
\frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|} \\
&\iff \frac{\vartheta' \rho_{j'}^{\ell}}{\|\boldsymbol{d}_{j',k_0}\|} = \frac{\vartheta_0 \rho_{j_0}^{\ell}}{\|\boldsymbol{d}_{j_0,k_0}\|}, \quad \ell = 0,\dots,m+k_0, \text{ and } k' = k_0 \\
&\iff \vartheta'\vartheta_0 \frac{\|\boldsymbol{d}_{j_0,k_0}\|}{\|\boldsymbol{d}_{j',k_0}\|} = \left(\frac{\rho_{j_0}}{\rho_{j'}}\right)^{\ell}, \quad \ell = 0,\dots,m+k_0, \text{ and } k' = k_0.
\end{aligned}$$

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

$$\vartheta' \vartheta_0 \frac{\|\boldsymbol{d}_{j_0,k_0}\|}{\|\boldsymbol{d}_{j',k_0}\|} = \left(\frac{\rho_{j_0}}{\rho_{j'}}\right)^{\ell}, \quad \ell = 0, 1, \ldots, m+k_0, \text{ and } k' = k_0$$
$$\iff \vartheta' = \vartheta_0 \text{ and } j' = j_0 \text{ and } k' = k_0.$$

If $k_0 = -m$, given that $(\vartheta_0, j_0, k_0) \in \mathcal{I} = S_1 \times \Big(\{1, \ldots, J\} \times \{-m, \ldots, -1, 0, 1, \ldots, h\} \cup \{(0, -m)\}\Big)$, then necessarily $j_0 = 0$. Furthermore, as $k' = k_0 = -m$, we similarly have that $j' = j_0 = 0$ and thus $\boldsymbol{d}_{j',k_0} = \boldsymbol{d}_{j_0,k_0} = \boldsymbol{d}_{0,-m} = (1, 0, \ldots, 0)$.

Hence

$$\vartheta' \vartheta_0 \frac{\|\boldsymbol{d}_{j_0,k_0}\|}{\|\boldsymbol{d}_{j',k_0}\|} = \left(\frac{\rho_{j_0}}{\rho_{j'}}\right)^{\ell}, \quad \ell = 0, \text{ and } k' = k_0 = -m \text{ and } j' = j_0 = 0,$$
$$\iff \vartheta' = \vartheta_0 \text{ and } k' = k_0 = -m \text{ and } j' = j_0 = 0$$

Case $m = 0$

If $k_0 \in \{1, \ldots, h\}$ then $f(\boldsymbol{d}_{j_0,k_0}) = \rho_{j_0}^{k_0}$ and by (3.3), $\|\boldsymbol{d}_{j_0,k_0}\| = |\rho_{j_0}|^{k_0}$. Thus, $\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\| = \vartheta_0$.
If $k_0 = -m = 0$, then $j_0 = 0$ and $f(\boldsymbol{d}_{j_0,k_0}) = 1$ and $\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\| = \vartheta_0$.
The same holds for $(\vartheta', j', k') \in \mathcal{I}$ and we obtain that

$$\frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|} \iff \vartheta' = \vartheta_0.$$

Proof. By Proposition 3.2,

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A_{\vartheta,j,k} \Big| B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in A_{\vartheta,j,k} : \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right)}{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C_{m+h+1}^{\|\cdot\|} : \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right)}. \tag{A.9}$$

Focusing on the denominator, we have by (3.15)

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C_{m+h+1}^{\|\cdot\|} : \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right) = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C_{m+h+1}^{\|\cdot\|} : \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right)$$

We will now distinguish the cases arising from the application of Lemma A.1. Recall that we assume for this proposition that the ρ_j's are positive. Thus, $\text{sign}(\rho_j) = 1$ and $\bar{\beta}_j = \beta_j \dfrac{1 - |\rho_j|^{\alpha}}{1 - \rho_j^{<\alpha>}} = \beta_j$ and $\bar{w}_{j,\vartheta} = w_{j,\vartheta}$ in (3.14) for all j's and $\vartheta \in \{-1, +1\}$.

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

By Lemma A.1,

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C^{\|\cdot\|}_{m+h+1}: \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right)$$
$$= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k'}}{\|\boldsymbol{d}_{j_0,k'}\|}: \quad 0 \leq k' \leq h\right\}\right)$$
$$= \sigma^\alpha \pi_{j_0}^\alpha \left[w_{j_0,\vartheta_0} \sum_{k'=0}^{h-1} \|\boldsymbol{d}_{j_0,k'}\|^\alpha + \frac{\bar{w}_{j_0,\vartheta_0}}{1-|\rho_{j_0}|^\alpha} \|\boldsymbol{d}_{j_0,h}\|^\alpha\right]$$

By (3.3), for $k' \in \{0, 1, \ldots, h\}$

$$\|\boldsymbol{d}_{j_0,k'}\| = \|(\rho_{j_0}^{k'+m}, \ldots, \rho_{j_0}^{k'+1}, \rho_{j_0}^{k'}, \underbrace{0, \ldots, 0}_{h})\|$$
$$= |\rho_{j_0}|^{k'-h} \|(\rho_{j_0}^{m+h}, \ldots, \rho_{j_0}^{h+1}, \rho_{j_0}^{h}, \underbrace{0, \ldots, 0}_{h})\|$$
$$= |\rho_{j_0}|^{k'-h} \|\boldsymbol{d}_{j_0,h}\|.$$

Thus,

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C^{\|\cdot\|}_{m+h+1}: \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right)$$
$$= \sigma^\alpha \pi_{j_0}^\alpha w_{j_0,\vartheta_0} \|\boldsymbol{d}_{j_0,h}\|^\alpha \left[\sum_{k'=0}^{h-1} |\rho_{j_0}|^{\alpha(k'-h)} + \frac{1}{1-|\rho_{j_0}|^\alpha}\right]$$
$$= \sigma^\alpha \pi_{j_0}^\alpha w_{j_0,\vartheta_0} \|\boldsymbol{d}_{j_0,h}\|^\alpha \frac{|\rho_{j_0}|^{-\alpha h}}{1-|\rho_{j_0}|^\alpha}.$$

Similarly for the numerator in (A.9), by (3.16),

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in A_{\vartheta,j,k}: \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right)$$
$$= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k'}}{\|\boldsymbol{d}_{j_0,k'}\|} \in A_{\vartheta,j,k}: \quad 0 \leq k' \leq h\right\}\right)$$

$$= \begin{cases} \Gamma^{\|\cdot\|}\left(\left\{\dfrac{\vartheta_0 \boldsymbol{d}_{j_0,k}}{\|\boldsymbol{d}_{j_0,k}\|}\right\}\right), & \text{if} \quad j = j_0 \quad \text{and} \quad \vartheta = \vartheta_0, \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if} \quad j \neq j_0 \quad \text{or} \quad \vartheta \neq \vartheta_0, \end{cases}$$

$$= \begin{cases} \sigma^\alpha \pi_{j_0}^\alpha w_{j_0,\vartheta_0} \|\boldsymbol{d}_{j_0,h}\|^\alpha |\rho_{j_0}|^{\alpha(k-h)} \delta_{\{\vartheta_0\}}(\vartheta) \delta_{\{j_0\}}(j), & \text{if} \quad 0 \leq k \leq h-1, \\ \sigma^\alpha \pi_{j_0}^\alpha w_{j_0,\vartheta_0} \|\boldsymbol{d}_{j_0,h}\|^\alpha \dfrac{1}{1-|\rho_{j_0}|^\alpha} \delta_{\{\vartheta_0\}}(\vartheta) \delta_{\{j_0\}}(j), & \text{if} \quad k = h. \end{cases}$$

The σ^α terms cancel out in the ratio.

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

We have by Lemma A.1

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C^{\|\cdot\|}_{m+h+1} : \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right) = \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k_0}}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right).$$

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

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k_0}}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right) = \sigma^\alpha \pi^\alpha_{j_0} w_{j_0,\vartheta_0} \|\boldsymbol{d}_{j_0,k_0}\|^\alpha,$$

and

$$\begin{aligned}
\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in A_{\vartheta,j,k} : \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right)
&= \Gamma^{\|\cdot\|}\left(A_{\vartheta,j,k} \cap \left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k_0}}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right) \\
&= \begin{cases} \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k_0}}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right), & \text{if } j = j_0 \text{ and } \vartheta = \vartheta_0, \text{ and } k = k_0, \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if } j \neq j_0 \text{ or } \vartheta \neq \vartheta_0 \text{ or } k \neq k_0, \end{cases} \\
&= \sigma^\alpha \pi^\alpha_{j_0} w_{j_0,\vartheta_0} \|\boldsymbol{d}_{j_0,k_0}\|^\alpha \delta_{\{\vartheta_0\}}(\vartheta) \delta_{\{j_0\}}(j) \delta_{\{k_0\}}(k).
\end{aligned}$$

If $k_0 = -m$, then $\boldsymbol{d}_{j_0,k_0} = \boldsymbol{d}_{0,-m} = (1, 0, \ldots, 0)$, and

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k_0}}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right) = \Gamma^{\|\cdot\|}\Big(\{\vartheta_0(1, 0, \ldots, 0)\}\Big) = \sigma^\alpha w_{\vartheta_0},$$

and

$$\begin{aligned}
&\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in A_{\vartheta,j,k} : \quad \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right) \\
&\qquad = \Gamma^{\|\cdot\|}\left(A_{\vartheta,j,k} \cap \left\{\frac{\vartheta_0 \boldsymbol{d}_{j_0,k_0}}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right) \\
&\qquad = \begin{cases} \Gamma^{\|\cdot\|}\Big(A_{\vartheta,j,k} \cap \{\vartheta_0(1, 0, \ldots, 0)\}\Big), & \text{if } \vartheta = \vartheta_0, \text{ and } k = k_0 = -m, \text{ and } j = j_0 = 0 \\ \Gamma^{\|\cdot\|}(\emptyset), & \text{if } \vartheta \neq \vartheta_0 \text{ or } k \neq k_0, \text{ or } j \neq j_0 \end{cases} \\
&\qquad = \sigma^\alpha w_{\vartheta_0} \delta_{\{\vartheta_0\}}(\vartheta) \delta_{\{j_0\}}(j) \delta_{\{k_0\}}(k).
\end{aligned}$$

Again, the σ^α terms cancel out in the ratio.

Case $m = 0$

By Lemma A.1, as the ρ_j's are positive

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C_{m+h+1}^{\|\cdot\|} : \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right)$$
$$= \Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta_0 \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C_{m+h+1}^{\|\cdot\|} : (j',k') \in \{1,\dots,J\} \times \{0,\dots,h\} \cup \{(0,0)\}\right\}\right)$$

Given that $w_{\vartheta_0} = \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0}$ and $\|\boldsymbol{d}_{j',k'}\| = |\rho_{j'}|^{k'}$, for any $1 \le j' \le J$, $1 \le k' \le h$,

$$\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C_{m+h+1}^{\|\cdot\|} : \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} = \frac{\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})}{\|\boldsymbol{d}_{j_0,k_0}\|}\right\}\right)$$
$$= \sigma^{\alpha} w_{\vartheta_0} + \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \left[\sum_{k'=1}^{h-1} \|\boldsymbol{d}_{j',k'}\|^{\alpha} + \frac{\|\boldsymbol{d}_{j',h}\|^{\alpha}}{1 - |\rho_{j'}|^{\alpha}}\right]$$
$$= \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \left[1 + \sum_{k'=1}^{h-1} |\rho_{j'}|^{\alpha k'} + \frac{|\rho_{j'}|^{\alpha h}}{1 - |\rho_{j'}|^{\alpha}}\right]$$
$$= \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \left[\frac{1 - |\rho_{j'}|^{\alpha h}}{1 - |\rho_{j'}|^{\alpha}} + \frac{|\rho_{j'}|^{\alpha h}}{1 - |\rho_{j'}|^{\alpha}}\right]$$
$$= \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \frac{1}{1 - |\rho_{j'}|^{\alpha}}.$$

Similarly, by (3.16),

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

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

$$= \begin{cases} \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \delta_{\{\vartheta_0\}}(\vartheta), & \text{if } \quad k = 0, \\ \sigma^{\alpha} \pi_j^{\alpha} w_{j,\vartheta_0} |\rho_j|^{\alpha k} \delta_{\{\vartheta_0\}}(\vartheta), & \text{if } \quad 1 \le k \le h-1, \\ \sigma^{\alpha} \pi_j^{\alpha} w_{j,\vartheta_0} \dfrac{|\rho_j|^{\alpha h}}{1 - |\rho_j|^{\alpha}} \delta_{\{\vartheta_0\}}(\vartheta), & \text{if } \quad k = h. \end{cases}$$

The conclusion follows.