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

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

Abstract

This paper introduces a novel framework based on α-stable moving average aggregates to model financial bubbles with heterogeneous growth and crash dynamics. We establish the theoretical properties of the model, showing in particular that it admits a semi-norm representation, which enables the prediction of extreme trajectories from past observations whenever each latent component is anticipative. We develop a minimum distance estimator based on the joint empirical characteristic function of consecutive observations and establish its consistency and asymptotic normality under suitable regularity conditions. Monte Carlo simulations confirm reliable finite-sample performance, and a subsampling procedure empirically validates the convergence to the asymptotic Gaussian distribution, while revealing heterogeneous convergence speeds across parameter dimensions. An empirical application to the CBOE Crude Oil ETF Volatility Index decomposes observed volatility dynamics into distinct latent components with different persistence properties, showing that what appears as a single explosive 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, and emerge as solutions to linear rational expectation models that admit multiple stationary equilibria through infinite variance innovations (Gouriéroux et al. , 2020). Such patterns are incompatible with standard linear time series models, which fail to simultaneously accommodate heavy-tailed innovations, explosive growth, and the abrupt reversals characteristic of locally explosive episodes. From an empirical perspective, anticipative (or noncausal) models appear as good candidates to account for the non-linear dynamics of bubbles and the non-Gaussian environment. Such 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 nonlinear dynamics or non-i.i.d. innovations (Gourieroux and Jasiak , 2026; Fries and Zakoian , 2019). Beyond estimation, this class of models has been studied for its forecasting properties (Gouriéroux and Jasiak , 2016, 2018; Gourieroux and Jasiak , 2026) and its applications ranging from inflation and macroeconomic dynamics (Lanne and Saikkonen , 2011, 2013; Hecq et al. , 2020) to bubble modelling and tail risk (Fries and Zakoian , 2019; Gouriéroux et al. , 2025; **?**). This framework also exhibits intriguing properties, such as a predictive distribution with lighter tails than the marginal distribution, which 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), both critical for informed investment decisions.

A natural class of such models is provided by anticipative stable two-sided moving averages: stationary linear processes of the form

$$X_t = \sum_{k\in\mathbb{Z}} d_k \varepsilon_{t+k}, \qquad \varepsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{S}(\alpha,\beta,\sigma,0), \tag{1.1}$$

where (ε_t) is an i.i.d. sequence of α-stable innovations with tail index $\alpha \in (0,2]$, skewness $\beta \in [-1,1]$, scale $\sigma > 0$, and $(d_k)_{k\in\mathbb{Z}}$ is a real deterministic sequence satisfying appropriate summability conditions. Because the coefficient d_k is nonzero for $k > 0$, the process X_t depends on future innovations, making it anticipative. The simplest instance is the purely anticipative stable AR(1) studied in Gouriéroux and Zakoian (2017), which corresponds to $d_k = \psi^k \mathbf{1}_{k\geq 0}$, $|\psi| < 1$, and admits the two-sided moving-average representation

$$X_t = \sum_{k=0}^{+\infty} \psi^k \varepsilon_{t+k}, \qquad |\psi| < 1, \tag{1.2}$$

showing explicitly that X_t depends on current and future innovations only, with geometrically decaying coefficients at rate ψ. This framework extends naturally to mixed-causal autoregressive processes of orders (r,s), denoted MAR(r,s), defined by $\Phi(L)\Psi(L^{-1})X_t = \varepsilon_t$, where $\Phi(L) = \prod_{i=1}^r (1-\lambda_i L)$ with $|\lambda_i| < 1,\ \forall i \in \{1,\dots,r\}$ and $\Psi(L^{-1}) = \prod_{j=1}^s (1-\zeta_j L^{-1})$ with $|\zeta_j| < 1, \forall j \in \{1,\dots,s\}$. Under these root conditions, every MAR(r,s) process admits a two-sided MA(∞) representation of the form (**??**), with coefficients (d_k) that decay geometrically on each side of the origin (Gouriéroux and Jasiak , 2016; de Truchis and Thomas , 2026). The purely causal case $(s = 0)$ reduces to a standard AR(r) model driven only by past innovations, while the purely noncausal case $(r = 0)$ yields a process driven exclusively by future innovations.

However, anticipative models impose a similar increase rate for all bubbles, fully determined by the noncausal autoregressive coefficients (Gouriéroux and Zakoian , 2017).[4] This lack of flexibility may conflict with empirical evidence

[4] Lanne and Luoto (2013) introduce time-varying parameters in noncausal models not to allow growth rates to vary directly over time, but their approach, in a sense, addresses this issue.

on financial markets where the surge of explosive episodes can exhibit very different patterns. A natural motivation for heterogeneous bubble dynamics also arises from heterogeneous agent models: when fundamentalist and chartist traders interact, their competing strategies can generate distinct speculative components with different persistence properties (Agliari et al. , 2018), further supporting aggregation as a natural device for capturing the empirical complexity of bubble dynamics. A natural remedy is to consider processes resulting from the linear aggregation of $J \geq 1$ independent latent stable moving averages, each driven by its own innovation sequence. Formally, an α-stable aggregate is defined as

$$\mathcal{X}_t = \sigma \sum_{j=1}^{J} \pi_j X_{j,t}, \qquad X_{j,t} = \sum_{k \in \mathbb{Z}} d_{j,k}\, \varepsilon_{j,t+k}, \qquad \varepsilon_{j,t} \overset{\text{i.i.d.}}{\sim} \mathcal{S}(\alpha, \beta_j, 1, 0), \tag{1.3}$$

where $\sigma > 0$ is an overall scale parameter, $(\pi_j)_{j=1}^{J}$ are positive mixing weights summing to one, $(d_{j,k})_k$ are distinct deterministic coefficient sequences characterising the dynamics of each latent component $X_{j,t}$, and the innovation sequences $(\varepsilon_{j,t})$ are mutually independent. The tail index α is assumed common across all components. This restriction rests on two grounds. First, it follows from the stability under linear combinations: the aggregate $\mathcal{X}_t$ admits a well-defined α-stable distribution only when all components share the same tail index, since the tail of a sum of independent stable variables is governed by the heaviest one and the resulting distribution is tractable only in the equal-α case. Second, a common α preserves identifiability of the remaining parameters: the components are distinguished by their dynamic coefficients $(d_{j,k})_k$ and mixing weights (π_j), so that heterogeneity in bubble growth rates is fully captured by the autoregressive structure rather than by tail heterogeneity. As a consequence, all components share the same tail decay rate while differing in their persistence and crash dynamics; the asymmetry parameters β_j may however vary across components.

For example, the simplest specification takes each latent component $X_{j,t}$ to be a purely anticipative stable AR(1) process as in (1.2), with $d_{j,k} = \psi_j^k\, \mathbf{1}_{k \geq 0}$, $\psi_j \in (0,1)$, so that

$$\mathcal{X}_t = \sigma \sum_{j=1}^{J} \pi_j X_{j,t}, \qquad X_{j,t} = \sum_{k=0}^{+\infty} \psi_j^k\, \varepsilon_{j,t+k}. \tag{1.4}$$

Because the ψ_j differ across components, each latent process generates a distinct bubble growth pattern, and their superposition produces richer and more heterogeneous explosive dynamics than any single anticipative process. Furthermore, as noted by Gouriéroux et al. (2021), aggregation introduces multiple independent sources of noise, making the stable aggregate structurally different from any mixed-causal or two-sided moving average model, and therefore better suited for derivative pricing where each risk factor must be separately identified and hedged.

The literature on the estimation of anticipative stable processes is relatively sparse. For non-aggregated mixed-causal AR models, an early semi-parametric approach is developed by Gassiat (1990, 1993), who establish consistency, asymptotic normality, and adaptivity for noncausal AR(p) processes with infinite-variance innovations, thereby covering the α-stable case; however, their procedure requires knowledge of the innovation density up to a scale parameter. Building on this line of work, Breidt et al. (1991) establish maximum likelihood theory for noncausal ARMA processes, which Andrews et al. (2009) extend to α-stable AR models in full generality. Velasco and Lobato (2018) develop frequency-domain minimum distance estimation for potentially non-invertible and noncausal ARMA models. Gourieroux and Jasiak (2023) propose the Generalized Covariance estimator (GCov), which identifies MAR(r, s) parameters by minimizing serial dependence in polynomial transformations of the pseudo-residuals; however, its asymptotic theory requires innovations with finite moments of a sufficiently high order, a condition incompatible with α-stable distributions. For aggregated processes, Gouriéroux and Zakoian (2017) propose a Cauchy-specific ($\alpha = 1$, $\beta = 0$) minimum distance estimator (MDE) based on the empirical characteristic function (ECF), deriving identification results for both continuous and discrete mixing distributions; nonetheless, their framework is confined to the Cauchy family and to purely anticipative AR(1) latent components. The ECF approach is grounded in the minimum distance theory of Knight and Yu (2002) and Xu and Knight (2010), who establish $\sqrt{n}$-consistency and asymptotic normality for ECF estimators in stationary time series, albeit under finite-variance assumptions that do not directly extend to the α-stable setting. Finally, Gouriéroux et al. (2021) study a Gaussian–Cauchy aggregate for WTI crude oil prices but do not establish formal asymptotic properties: no consistency result or central limit theorem is provided for the component-weight estimator, nor are the mixing conditions required under infinite-variance innovations formally verified.

In this paper, we make two contributions to the literature on econometric modelling of financial bubbles. First, we introduce the aggregation model (1.3), a novel flexible framework that overcomes a key limitation of existing anticipative heavy-tailed models, which impose uniform growth patterns across different bubble episodes. We derive the theoretical tail properties characterizing such an aggregation model. We demonstrate that the model admits a semi-norm representation[5], except when one of the underlying components exhibits purely non-anticipative behavior. This structural property allows us to predict extreme trajectories with heterogeneous growth patterns.

The main object underlying prediction in the α-stable framework is the spectral measure, which encodes the dependence structure of the process across time. In standard representations, this measure distributes mass in all directions of the space, mixing past and future indistinguishably. The semi-norm representation, introduced in de Truchis et al. (2025a), circumvents this limitation by replacing the standard norm with a semi-norm that assigns zero weight to future coordinates: the spectral mass is then entirely supported on directions that depend on the past alone. Economically, this means that the shape of current observations, specifically the magnitude and duration of the price surge, fully determines the asymptotic conditional distribution of the future trajectory. When this representation exists, the conditional probability that a bubble follows a given path converges to a well-defined limit as observations grow large, providing a coherent early-warning system for predicting bubble peaks.

Building on de Truchis et al. (2025a), we characterise when the aggregate admits such a representation. The main condition is simple: each latent component must be genuinely anticipative, in the sense that its forward-looking moving average coefficients do not vanish over arbitrarily long horizons. Intuitively, a purely causal component contributes innovations that are entirely invisible from any finite window of past observations, so that no amount of history can signal the bubble it drives, such a bubble always arrives as a complete surprise.

Second, we develop a minimum distance estimator based on the joint empirical characteristic function of consecutive observations $(\mathcal{X}_t, \mathcal{X}_{t+1})$, and we demonstrate its consistency and asymptotic normality under suitable regularity conditions. Departing from Gouriéroux and Zakoian (2017), who focus on continuous support distributions for the mixing weights in the specific Cauchy case, our approach handles discrete mixing weights and the full α-stable family with $\alpha \in (1, 2)$.[6] We first consider aggregates of anticipative AR(1) processes, then extend the framework to MAR$(1, 1)$ aggregates, and finally cover the general MAR(r, s) case, building on the exact two-sided MA(∞) coefficient representation derived in de Truchis and Thomas (2026); when no closed-form expression for the characteristic function is available, a numerically truncated approximation is shown to be arbitrarily accurate at controlled computational cost.

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 heavy-tailed stable distributions. We establish the asymptotic properties of our estimator

[5]similarly to non-aggregated processes (de Truchis et al. , 2025a)

[6]The restriction $\alpha > 1$ ensures that the objective function $D_{\mathcal{X}}(\theta)$ belongs to $C^2(\Theta)$, a regularity condition required for the asymptotic theory of the estimator (Lemma 2.1). It also guarantees that the innovations have finite moments of all orders $\delta \in (0, \alpha)$, which is needed for the strong mixing property of the latent processes and the central limit theorem underlying asymptotic normality (Knight and Yu , 2002; Xu and Knight , 2010).

under suitable regularity conditions, proving consistency and asymptotic normality. To numerically validate the finite-sample convergence toward the limiting Gaussian distribution, we conduct an extensive simulation study combining Monte Carlo experiments and subsampling diagnostics following Politis and Romano (1994) and Politis, Romano and Wolf (1999). The Monte Carlo analysis demonstrates that the estimator exhibits reliable convergence properties across various parameter configurations, though with moderate finite-sample biases The subsampling procedure further reveals heterogeneous convergence speeds across parameter dimensions and confirms that while certain parameters, such as the tail index α and the overall scale σ, approach asymptotic normality relatively quickly, others, particularly the autoregressive coefficients (ψ_j) and the mixing weights (π_j), require substantially larger sample sizes to achieve reliable normal approximations. In particular, the mixing weights are the most demanding in terms of sample size, as their identification relies on subtle differences in the characteristic function across components that only become statistically distinguishable as T grows large; our simulations suggest that samples of at least $T = 500$ observations are needed to obtain trustworthy inference on (π_j).

s an empirical illustration, we estimate an aggregation of purely anticipative stable AR(1) processes using the CBOE Crude Oil ETF Volatility Index (OVX) data collected at weekly frequency over May 23, 2015–May 23, 2025 ($T = 522$), 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, each driven by an independent innovation sequence, pointing to the coexistence of multiple speculative components within a single observed bubble.

The remainder of this paper is organised 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 numerically 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 relegated to Appendix A, while the Online Supplement provides Monte Carlo results, subsampling diagnostics, convergence analysis, and additional results for the empirical application.

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 $d_k = \psi^k$, 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 $t \mapsto cd_{\tau-t}$ (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*

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

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

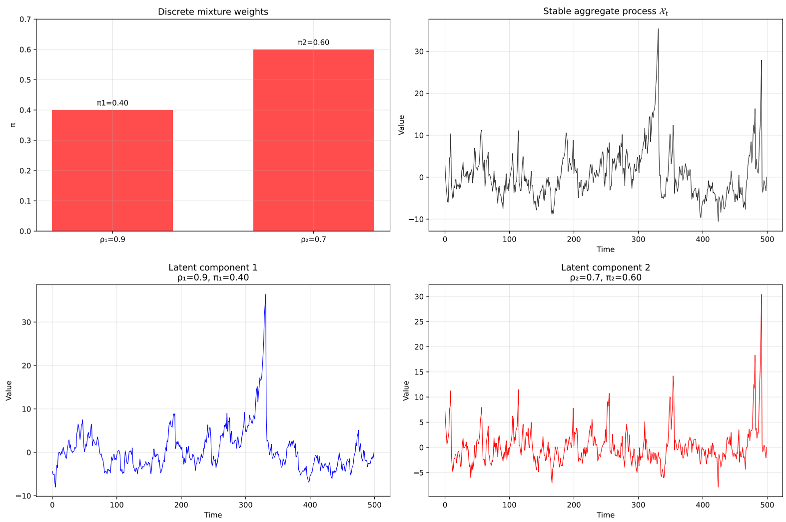

Figure 1: Simulated stable aggregate dynamics with two components. Top left: Distribution of weights for the two components with $\psi_1 = 0.90$, $\pi_1 = 0.40$ for the first component and $\psi_2 = 0.70$, $\pi_2 = 0.60$ for the second component. Top right: The resulting trajectory of the aggregated process $\mathcal{X}_t$. Middle and bottom panels: The individual latent component processes with different persistence parameters.

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.

We provide the derivation for four parametric specifications. The first three admit a closed-form characteristic function: the aggregation of purely anticipative AR(1) processes, mixed causal-noncausal MAR(1, 1) processes , and the Gaussian-plus-AR(1) mixture of Gouriéroux et al. (2021). Notice that even in these specific frameworks, these aggregations feature much richer dynamics than single-component stable processes, as illustrated in Figure 1. The fourth specification extends the framework to general MAR(r, s) processes; in this case, no closed-form expression for the characteristic function is available, but we show that a truncated approximation of the two-sided MA(∞) representation of de Truchis and Thomas (2026) yields an arbitrarily accurate surrogate at a controlled computational cost. 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. In the rest of this section, we focus on $\beta_j = \beta$ for simplicity.

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 $d_{j,k} = \psi_j^k \mathbf{1}_{k \geq 0}$, with $0 < \psi_j < 1$. 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.[7] The process is thus defined by $X_{j,t} = \sum_{k=0}^{\infty} \psi_j^k \varepsilon_{j,t+k}$. The joint characteristic function of the vector $(X_{j,t}, X_{j,t+1})$ is given by

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

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

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

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

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

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

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

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

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

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

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

2.2. 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

$$\varphi_n(u,v) = \frac{1}{n-1}\sum_{j=1}^{n-1}\exp(i(u\mathcal{X}_j + v\mathcal{X}_{j+1})) \tag{2.7}$$

[7] 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 $\beta' = \frac{\sum_{k=0}^{\infty}|c_k|^{\alpha}\mathrm{sign}(c_k)}{\sum_{k=0}^{\infty}|c_k|^{\alpha}} \cdot \beta$. When $\psi_j > 0$, all coefficients $c_k = \psi_j^k$ are positive, yielding $\beta' = \beta$. However, when $\psi_j < 0$, the coefficients alternate in sign, leading to $\beta' \neq \beta$. The case $\psi_j < 0$ would thus require a component-specific modified asymmetry parameter β_j' in the characteristic function.

which can be decomposed into real and imaginary parts:

$$\varphi_n(u,v) = \frac{1}{n-1}\sum_{j=1}^{n-1}\cos(u\mathcal{X}_j + v\mathcal{X}_{j+1}) + \frac{1}{n-1}\sum_{j=1}^{n-1} i\sin(u\mathcal{X}_j + v\mathcal{X}_{j+1}) \tag{2.8}$$

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. For the sake of simplicity, we illustrate the parameter identification logic for the special case of the anticipative AR(1). The identification of $\theta = (\sigma,\psi_1,\ldots,\psi_J,\pi_1,\ldots,\pi_J,\alpha,\beta)$ 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 (2.7) is dominated by the α-stable distribution's properties. Specifically, for $u>0$,

$$\alpha = \lim_{u\to 0}\frac{\log\log|\varphi_n(u,0)|^{-1}}{\log|u|} \tag{2.9}$$

and

$$\beta = -\lim_{u\to 0}\frac{\mathrm{Im}(\log\varphi_n(u,0))}{\mathrm{Re}(\log\varphi_n(u,0))}\cdot\cot\frac{\pi\alpha}{2}. \tag{2.10}$$

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

$$g_n(\lambda) = \lim_{u\to 0}\frac{\log|\varphi_n(u,\lambda u)|}{|u|^\alpha} \approx -\sigma^\alpha\sum_{j=1}^{J}\pi_j^\alpha\left(\frac{|\psi_j+\lambda|^\alpha}{1-|\psi_j|^\alpha}+1\right) \tag{2.11}$$

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 $(\sigma,\psi_1,\ldots,\psi_J,\pi_1,\ldots,\pi_J)$.

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

$$D_{\mathcal{X}}(\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|\varphi_n(u,v)-\varphi(u,v;\theta)|^2 w(u,v)dudv \tag{2.12}$$

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

$$\hat\theta_n = \arg\min_\theta D_{\mathcal{X}}(\theta). \tag{2.13}$$

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 $|\psi_j u+v|^\alpha$ vanishes. Without appropriate regularization through the weight function, these singularities can cause the integrals defining the first and second derivatives of (2.12) to diverge. Moreover, when $\alpha=1$, an additional source of divergence arises in the first derivative of the characteristic exponent with respect to α, since the term $\tan(\pi\alpha/2)$ appearing in the standard parameterization satisfies $\frac{\partial}{\partial\alpha}\tan(\pi\alpha/2) = \frac{\pi/2}{\cos^2(\pi\alpha/2)}\to\infty$ as $\alpha\to 1$, regardless of the behavior of $|u|^\alpha$. 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*

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

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

Then, 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 distributions. 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 (2.18) and Assumption 5 is satisfied by the global identification conditions exposed in (2.9), (2.10) and (2.11). The proof of Lemma 2.1 is postponed in Appendix 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 $D_0(\theta) = \iint |\varphi(u, v; \theta_0) - \varphi(u, v; \theta)|^2 w(u, v)dudv$ 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*

$$K(x;\theta) = \iint \Big[(\cos(ux_j + vx_{j+1}) - Re\ \varphi(u,v;\theta))\frac{\partial Re\ \varphi(u,v;\theta)}{\partial\theta} + (\sin(ux_j + vx_{j+1}) - Im\ \varphi(u,v;\theta))\frac{\partial Im\ \varphi(u,v;\theta)}{\partial\theta}\Big] w(u,v)dudv. \tag{2.15}$$

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

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

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 $\nu_j = \mathbb{E}[K_0|K_j, K_{j-1}, \ldots] - \mathbb{E}[K_0|K_{j-1}, K_{j-2}, \ldots]$ 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*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N\big(0, \Sigma(\theta_0)^{-1}\Omega(\theta_0)\Sigma(\theta_0)^{-1}\big) \tag{2.16}$$

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 $j = 1, \ldots, J$, hence reducing the dimension of $\theta = (\varsigma_1, \ldots, \varsigma_J, \psi_1, \ldots, \psi_J, \alpha, \beta)$ to $(2J + 2)$ and virtually eliminating the constraint $\pi_1 + \ldots + \pi_J = 1$.

2.3. 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(∞), $d_{j,k}$, 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

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

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 noncausal ($k \geq 1$) 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 noncausal 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

$$\log \varphi_{\mathcal{X}}(u,v) = -\sigma^\alpha \sum_{j=1}^{J} \frac{\pi_j^\alpha}{|1-\phi_j\psi_j|^\alpha}\left(\mathcal{C}_j(u,v) + \mathcal{A}_j(u,v)\right) \tag{2.17}$$

where $\mathcal{C}_j(u,v)$ and $\mathcal{A}_j(u,v)$ represent the complex-valued contributions from the causal and noncausal 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

$$\sum_{j=1}^{J} \frac{\pi_j^s}{|1-\phi_j\psi_j|^s}\left(\frac{1}{1-|\psi_j|^s} + \frac{|\phi_j|^s}{1-|\phi_j|^s}\right) < \infty \quad \text{for } s \in (0,\alpha)\cap[0,1]. \tag{2.18}$$

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 noncausal 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

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

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 (2.12). The estimator $\hat{\theta}_n$ defined in (2.13) 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.

2.5. Case 4: General MAR(r,s) Aggregates

We now extend the estimation framework to aggregates of general MAR(r_j, s_j) processes with arbitrary causal order $r_j \geq 1$ and noncausal order $s_j \geq 1$. Consider a process $\mathcal{X}_t$ resulting from the aggregation of J independent MAR(r_j, s_j) components:

$$\Phi_j(L)\Psi_j(L^{-1})X_{j,t} = \varepsilon_{j,t}, \quad j = 1, \dots, J, \tag{2.19}$$

where $\Phi_j(L) = \prod_{i=1}^{r_j}(1 - \lambda_{j,i}L)$ with $|\lambda_{j,i}| < 1$, $\Psi_j(L^{-1}) = \prod_{l=1}^{s_j}(1 - \zeta_{j,l}L^{-1})$ with $|\zeta_{j,l}| < 1$, and $\varepsilon_{j,t} \overset{i.i.d.}{\sim} \mathcal{S}(\alpha, 0, 1, 0)$. We restrict to the symmetric case $\beta_j = 0$ for all j to simplify the exposition. Each component is assumed well-specified: the causal and noncausal roots are pairwise distinct and the polynomials $\Phi_j(z)$ and $\Psi_j(z^{-1})$ share no common factor, this rules out configurations where a causal root and a noncausal root are reciprocals of one another, which would result in a cancellation in the transfer function and an over-specified model order (r_j, s_j). The parameter vector for each component is $\theta_j = (\lambda_{j,1}, \dots, \lambda_{j,r_j}, \zeta_{j,1}, \dots, \zeta_{j,s_j}) \in \Theta_j \subset \mathbb{R}^{r_j + s_j}$, and the full parameter vector to be estimated is $\theta = (\sigma, \pi_1, \dots, \pi_J, \theta_1, \dots, \theta_J, \alpha) \in \mathbb{R}^{2J+1+\sum_{j=1}^J (r_j+s_j)}$, collecting the common scale σ, the mixing weights π_j, the causal and noncausal roots of each component, and the tail index α.

Unlike the MAR$(1,1)$ case, the MA(∞) coefficients exhibit a multi-mode structure. By de Truchis and Thomas (2026), these coefficients admit the closed-form representation:

$$d_{j,k}(\theta_j) = \begin{cases} \sum_{l=1}^{s_j} A_{j,l}(\theta_j)\, \zeta_{j,l}^k, & k \geq 1, \\ \sum_{i=1}^{r_j} B_{j,i}(\theta_j)\, \lambda_{j,i}^{|k|}, & k \leq 0, \end{cases} \tag{2.20}$$

where $A_{j,l}$ and $B_{j,i}$ are rational functions of the roots given by partial-fraction expansion.[8] The well-specification condition ensures that all poles are simple and separated from one another, so these weights are uniformly bounded over Θ:

$$\sup_{\theta \in \Theta} |d_{j,k}(\theta_j)| \leq C_0\, \rho^{|k|}, \qquad k \in \mathbb{Z}, \tag{2.21}$$

where $C_0 := \max(r_j, s_j)\, \rho^{\max(r_j,s_j)-1} / \mathrm{d}^{\max(r_j,s_j)}$, $\rho = \max(\max_i |\lambda_{j,i}|, \max_l |\zeta_{j,l}|) < 1$ and $\mathrm{d} := \inf_{\theta \in \Theta} \min\{\min_{l \neq m} |\zeta_{j,l} - \zeta_{j,m}|, \min_{i \neq i'} |\lambda_{j,i} - \lambda_{j,i'}|, \min_{i,l} |\lambda_{j,i}\zeta_{j,l} - 1|\} > 0$ denotes the minimum separation distance between roots over the parameter space Θ. Differentiating (2.20) yields analogous bounds for the derivatives:

$$\sup_{\theta \in \Theta} |\partial_\theta d_{j,k}(\theta_j)| \leq C_1(1 + |k|)\rho^{|k|}, \quad \sup_{\theta \in \Theta} |\partial^2_{\theta\theta'} d_{j,k}(\theta_j)| \leq C_2(1 + k^2)\rho^{|k|}, \tag{2.22}$$

[8] Specifically, by de Truchis and Thomas (2026),

$$A_{j,l}(\theta_j) = \frac{(-1)^{r_j}\, \zeta_{j,l}^{s_j - 1}}{\prod_{m \neq l}(\zeta_{j,l} - \zeta_{j,m}) \cdot \prod_{i=1}^{r_j}(\lambda_{j,i}\zeta_{j,l} - 1)} \text{ and } B_{j,i}(\theta_j) = \frac{(-1)^{s_j}\, \lambda_{j,i}^{r_j - 1}}{\prod_{i' \neq i}(\lambda_{j,i} - \lambda_{j,i'}) \cdot \prod_{l=1}^{s_j}(\lambda_{j,i}\zeta_{j,l} - 1)}.$$

where $C_1 := \max(r_j,s_j)\,\rho^{\max(r_j,s_j)-1}(1+\rho^{-1})/\mathrm{d}^{\max(r_j,s_j)+1}$ and $C_2 := \max(r_j,s_j)\,\rho^{\max(r_j,s_j)-1}(1+\rho^{-1})^2/\mathrm{d}^{\max(r_j,s_j)+2}$. The proportionality $d_{j,k}/d_{j,k-1} = \text{const.}$ that enabled closed-form characteristic functions in Cases 1–2 is lost when $\max(r_j,s_j) > 1$, so we work with a truncated characteristic function.

For notational brevity, set $\Delta_{j,k}(u,v;\theta_j) := u\,d_{j,k}(\theta_j) + v\,\delta_{j,k-1}(\theta_j)$ and $\boldsymbol{d}^1_{j,k}(\theta_j) := (d_{j,k}(\theta_j), d_{j,k-1}(\theta_j)) \in \mathbb{R}^2$. The exact joint log-CF of $(X_{j,t}, X_{j,t+1})$ is $\log\varphi_{X_j}(u,v;\theta_j) = -\sum_{k\in\mathbb{Z}}|\Delta_{j,k}|^\alpha$. For $M \geq 1$, define the truncated log-CF:

$$\log\varphi^{(M)}_{X_j}(u,v;\theta_j) := -\sum_{k=-M}^{M}|\Delta_{j,k}(u,v;\theta_j)|^\alpha, \tag{2.23}$$

the aggregate truncated CF:

$$\varphi^{(M)}_{\mathcal{X}}(u,v;\theta) = \prod_{j=1}^{J}\exp\{\log\varphi^{(M)}_{X_j}(\sigma\pi_j u,\, \sigma\pi_j v;\, \theta_j)\}, \tag{2.24}$$

and the MDE objective:

$$D^{(M)}_{\mathcal{X}}(\theta) = \iint_{\mathbb{R}^2}\big|\varphi_n(u,v) - \varphi^{(M)}_{\mathcal{X}}(u,v;\theta)\big|^2 w(u,v)\,du\,dv,$$

with $w(u,v) = \exp\{-\kappa(u^2+v^2)\}$ as in Assumption 2.

Lemma 2.2. *For every $k \in \mathbb{Z}$, j, $\theta \in \Theta$ with $\boldsymbol{d}^1_{j,k}(\theta) \neq 0$, and $\alpha \in (1,2)$:*

$$\iint_{\mathbb{R}^2}|\Delta_{j,k}|^\alpha\, w(u,v)\,du\,dv = C_{w,\alpha}\,\|\boldsymbol{d}^1_{j,k}\|^\alpha, \tag{2.25}$$

where $C_{w,\alpha} := \int_{\mathbb{R}}|s|^\alpha e^{-\kappa s^2}\,ds \cdot \int_{\mathbb{R}} e^{-\kappa t^2}\,dt < \infty$.

Lemma 2.3. *For $\alpha \in (1,2)$, the following bounds hold for every $M \geq 1$, $j \in \{1,\dots,J\}$, and $\theta \in \Theta$:*

$$\iint\big|\log\varphi_{X_j} - \log\varphi^{(M)}_{X_j}\big|\,w\,du\,dv \leq C^\star_0\,\rho^{\alpha M}, \tag{2.26}$$

$$\iint\big|\partial_\theta[\log\varphi_{X_j} - \log\varphi^{(M)}_{X_j}]\big|\,w\,du\,dv \leq C^\star_1(1+M)\,\rho^{\alpha M}, \tag{2.27}$$

$$\iint\big|\partial^2_{\theta\theta'}[\log\varphi_{X_j} - \log\varphi^{(M)}_{X_j}]\big|\,w\,du\,dv \leq C^\star_2(1+M)^2\,\rho^{\alpha M}, \tag{2.28}$$

where

$$C^\star_0 := \frac{2\,C_{w,\alpha}\,(C_0\sqrt{2})^\alpha}{\rho^\alpha(1-\rho^\alpha)}, \quad C^\star_1 := \frac{2\,\widetilde{C}_1}{(1-\rho^\alpha)^2}, \quad C^\star_2 := \frac{2\,\widetilde{C}_2}{(1-\rho^\alpha)^3},$$

with $C_{w,\alpha} = \int_{\mathbb{R}}|s|^\alpha e^{-\kappa s^2}\,ds \cdot \int_{\mathbb{R}} e^{-\kappa t^2}\,dt$, $\widetilde{C}_1 := \alpha\,C_{w,\alpha}\,C_0^{\alpha-1}C_1$, and $\widetilde{C}_2 := \alpha(\alpha-1)\,C_{w,\alpha}\,I_r(\alpha,\kappa)\,I_a(\alpha)\,C_1^2\,\rho^{-2}\,c_v^{\alpha-2} + \alpha\,C_{w,\alpha}\,C_0^{\alpha-1}C_2$, with $I_r(\alpha,\kappa) := \frac{1}{2}\kappa^{-(\alpha+2)/2}\Gamma\big(\frac{\alpha+2}{2}\big)$ and $I_a(\alpha) := \int_0^{2\pi}|\cos\psi|^{\alpha-2}\,d\psi$.

Lemma 2.4. *Under Assumption 2, with $\alpha \in (1,2)$ and $\beta_j = 0$ for all j:*

- ($\imath$) *For every fixed $M \geq 1$, $D^{(M)}_{\mathcal{X}} \in C^2(\Theta)$, and Assumptions 3, 6, 7, 8 are satisfied with $\varphi_{\mathcal{X}}$ replaced by $\varphi^{(M)}_{\mathcal{X}}$.*
- ($\imath\imath$) *As $M \to \infty$:*

$$\sup_{\theta\in\Theta}|D^{(M)}_{\mathcal{X}}(\theta) - D_{\mathcal{X}}(\theta)| = O(\rho^{\alpha M}),$$

$$\sup_{\theta\in\Theta}\|\nabla D^{(M)}_{\mathcal{X}}(\theta) - \nabla D_{\mathcal{X}}(\theta)\| = O\big((1+M)\rho^{\alpha M}\big),$$

$$\sup_{\theta\in\Theta}\|\nabla^2 D^{(M)}_{\mathcal{X}}(\theta) - \nabla^2 D_{\mathcal{X}}(\theta)\| = O\big((1+M)^2\rho^{\alpha M}\big).$$

The proofs of Lemmas 2.2–2.4 are given in Section A.2.

Proposition 2.2. *Let $\mathcal{X}_t$ be a strictly stationary $\mathcal{S}\alpha\mathcal{S}$ aggregate as in Definition 2.1 whose latent components are MAR(r_j, s_j) processes with $\beta_j = 0$ and $\alpha \in (1,2)$. Assume 2, 4, 5 and let $(M_n)_{n\geq 1}$ satisfy*

$$M_n \geq c_\star \log n, \qquad c_\star > \frac{1}{2\alpha \log(1/\rho)}. \tag{2.29}$$

Then $\hat{\theta}_n^{(M_n)} = \arg\min_{\theta\in\Theta} D_{\mathcal{X}}^{(M_n)}(\theta)$ satisfies:

$$\hat{\theta}_n^{(M_n)} \xrightarrow{P} \theta_0, \qquad \sqrt{n}\big(\hat{\theta}_n^{(M_n)} - \theta_0\big) \xrightarrow{d} \mathcal{N}\big(0,\, \Sigma(\theta_0)^{-1}\Omega(\theta_0)\Sigma(\theta_0)^{-1}\big),$$

with $\Sigma(\theta_0)$, $\Omega(\theta_0)$ as in Proposition 2.1.

The proof of Proposition 2.2 is given in Section A.2.

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.[9] Let the vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ 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,

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

where $\langle\cdot,\cdot\rangle$ denotes the canonical scalar product, $w(\alpha, s) = \text{tg}\left(\frac{\pi\alpha}{2}\right)$, 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, $\Gamma^{\|\cdot\|}(d\boldsymbol{s}) = \|\boldsymbol{s}\|_e^{-\alpha}\,\Gamma \circ T_{\|\cdot\|}^{-1}(d\boldsymbol{s})$ with $T_{\|\cdot\|} : S_d \setminus K^{\|\cdot\|} \longrightarrow C_d^{\|\cdot\|}$ defined by $T_{\|\cdot\|}(\boldsymbol{s}) = \boldsymbol{s}/\|\boldsymbol{s}\|$. 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

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

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

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

[9] 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

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

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

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 $(X_{1,t}), \dots, (X_{J,t})$ 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 $C_{m+h+1}^{\|\cdot\|}$ if and only if condition (3.4) holds with m for all coefficient sequences $(d_{j,k})_k$, $j = 1, \dots, J$. For $\alpha = 1$ and $\boldsymbol{X}_t$ asymmetric, the vector $\boldsymbol{X}_t$ is representable on $C_{m+h+1}^{\|\cdot\|}$ if and only if (3.4) holds and*

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

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

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 $(X_{1,t}),\dots,(X_{J,t})$ 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 $j = 1,\dots,J$ 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

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

(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 $j = 1,\dots,J$, i.e.,

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

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

(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), $m \geq \max_{j=1,\dots,J} m_{0,j}$, $h \geq 1$ such that either

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

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 $\boldsymbol{X_t} = (\mathcal{X}_{t-m},\dots,\mathcal{X}_t,\mathcal{X}_{t+1},\dots,\mathcal{X}_{t+h})$ on $C^{\|\cdot\|}_{m+h+1}$ is given by:

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

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

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 $\boldsymbol{d}_{j,k} = \boldsymbol{0}$, 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 $j = 1,\dots,J$, 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 $\alpha \in (0,2)$, $(\beta_1,\dots,\beta_J) \in [-1,1]^J$, 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 $j = 1,\dots,J$, $m_{0,j} < +\infty$, 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 $S_{m+1}^{\|\cdot\|} = \{(s_{-m}, \ldots, s_0) \in \mathbb{R}^{m+1} : \quad \|(s_{-m}, \ldots, s_0, 0, \ldots, 0)\| = 1\}$.[10] Then, for any Borel set $V \subset S_{m+1}^{\|\cdot\|}$, define the Borel set $B(V) \subset C_{m+h+1}^{\|\cdot\|}$ as

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

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

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

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

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

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

Observe that setting $V = S_{m+1}^{\|\cdot\|}$, 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

[10] 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.

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 $(d_{j,k+m}, \dots, d_{j,k}, d_{j,k-1}, \dots, d_{j,k-h})$ 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 $(\mathcal{X}_{t-m}, \dots, \mathcal{X}_{t-1}, \mathcal{X}_t)/\|\boldsymbol{X_t}\|$ will *a fortiori* be of the same shape as some $\vartheta(d_{j,k+m}, \dots, d_{j,k+1}, d_{j,k})/\|\boldsymbol{d}_{j,k}\|$ 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 ψ_j's are distinct. For each anticipative AR(1) with parameter ψ_j, the moving average coefficients are of the form $(\psi_j^k \mathbb{1}_{\{k\geq 0\}})_k$, 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 $d_{j,k} = \psi_j^k$ and general scale parameter $\sigma > 0$.*

Letting $\boldsymbol{X_t} = (\mathcal{X}_{t-m}, \dots, \mathcal{X}_t, \mathcal{X}_{t+1}, \dots, \mathcal{X}_{t+h})$ 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

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

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

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

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

The next proposition provides the tail conditional distribution of future paths in the case where the ψ_j's are positive. Let us first introduce useful neighbourhoods of the distinct charged points of $\Gamma^{\|\cdot\|}$. Denote $\boldsymbol{d}_{0,-m} = (\overbrace{1,0,\dots,0}^{m+h+1})$ 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 $\mathcal{I} := S_1 \times \Big(\{1,\dots,J\} \times \{-m, h\} \cup \{(0,-m)\}\Big)$. 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 $C^{\|\cdot\|}_{m+h+1}$ in which the only points charged by $\Gamma^{\|\cdot\|}$ all have the first $(m+1)^{\text{th}}$ coinciding with $\vartheta_0 f(\boldsymbol{d}_{j_0,k_0})/\|\boldsymbol{d}_{j_0,k_0}\|$. 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 $d_{j,k} = \psi_j^k \in (0,1)$ 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 $0 \leq k_0 \leq h$:*

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

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

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

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

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

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

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 $k_0 \in \{0,\dots,h\}$ or $k_0 \in \{-m,\dots,-1\}$. For the latter, the asymptotic probability of the whole path $\boldsymbol{X_t}/\|\boldsymbol{X_t}\|$ being in an arbitrarily small neighbourhood of $\vartheta \boldsymbol{d}_{j,k}/\|\boldsymbol{d}_{j,k}\|$ 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),

$$\frac{\boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} = \begin{cases} \dfrac{(\overbrace{\psi_j^{k+m},\ldots,\psi_j^{k}}^{m+1},\overbrace{\psi_j^{k-1},\ldots,\psi_j,1,0,\ldots,0}^{h})}{\|(\psi_j^{k+m},\ldots,\psi_j^{k},\psi_j^{k-1},\ldots,\psi_j,1,0,\ldots,0)\|}, & \text{for} \quad k\in\{0,\ldots,h\}, \\[2ex] \dfrac{(\psi_j^{k+m},\ldots,\psi_j,1,0,\ldots,0,0,\ldots,0)}{\|(\underbrace{\psi_j^{k+m},\ldots,\psi_j,1,0,\ldots,0}_{m+1},\underbrace{0,\ldots,0}_{h})\|}, & \text{for} \quad k\in\{-m,\ldots,-1\}. \end{cases}$$

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

$$\frac{f(\boldsymbol{d}_{j,k})}{\|\boldsymbol{d}_{j,k}\|} = \begin{cases} \dfrac{(\overbrace{\psi_j^{m},\ldots,\psi_j,1}^{m+1})}{\|(\underbrace{\psi_j^{m},\ldots,\psi_j,1}_{m+1},\underbrace{0,\ldots,0,0,\ldots,0}_{h})\|}, & \text{for} \quad k\in\{0,\ldots,h\}, \\[2ex] \dfrac{(\overbrace{\psi_j^{k+m},\ldots,\psi_j,1,0,\ldots,0}^{m+1})}{\|(\underbrace{\psi_j^{k+m},\ldots,\psi_j,1,0,\ldots,0}_{m+1},\underbrace{0,\ldots,0}_{h})\|}, & \text{for} \quad k\in\{-m,\ldots,-1\}. \end{cases}$$

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 ψ_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 $|\psi_j|^{\alpha k}(1-|\psi_j|^{\alpha})$ corresponds to the probability that the bubble will peak in exactly k periods ($0\leq k<h$), and $|\psi_j|^{\alpha h}$ 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 ψ_j^{-1} of an ongoing extreme episode,[11] 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.

[11] 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.

4. Simulation Results

This section documents the finite-sample performance of the minimum distance estimator developed in Section 2. Section 4.1 presents Monte Carlo evidence on estimation accuracy. Section 4.2 implements a subsampling methodology to empirically verify the asymptotic normality predicted by Proposition 2.1. Complete results, including detailed tables, graphical diagnostics, and sensitivity analyses, are provided in the Online Supplement, Sections S.1 and S.2.

4.1. Estimation accuracy

The observed process is generated by the aggregation of two independent α-stable AR(1) processes:

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

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

where $j \in \{1, 2\}$ and $\psi_j \in (0, 1)$. We fix $\sigma = 1.6$, $\pi_1 = 7/16$, $\pi_2 = 9/16$, yielding combined scale parameters $\varsigma_1 = 0.7$ and $\varsigma_2 = 0.9$. Three distributional settings are considered: (*i*) Cauchy ($\mathcal{S}1\mathcal{S}$); (*ii*) symmetric α-stable ($\mathcal{S}\alpha\mathcal{S}$, $\alpha = 1.5$); 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 $T \in \{250, 500, 1{,}000\}$.

Results (Table S.1 in the Online Supplement) confirm the good finite-sample behavior of the estimator. The dominant autoregressive coefficient ψ_1 is precisely estimated across all settings, with mean relative error (MRE) below 10% even in the $\mathcal{G}\alpha\mathcal{S}$ case for $T = 250$. The tail index α is recovered with high accuracy (MRE around 8% at $T = 250$, declining to 4–5% at $T = 1{,}000$). The smaller coefficient ψ_2 and the combined scale parameters exhibit higher relative errors, reflecting the difficulty in disentangling individual component contributions from the aggregate signal. The asymmetry parameter β is the most challenging to estimate (MRE of 63% at $T = 250$), although its identification is not required for the autoregressive and scale parameters.

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). Given a full sample of size n, we construct non-overlapping subsamples of size $b = \lfloor n^{2/3} \rfloor$:

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

[12] Since the individual scale parameters ς_1 and ς_2 exhibit slow finite-sample convergence (see Section S.2.2 of the Online Supplement), we apply a post-estimation reparameterization:

$$\vartheta = (\psi_1, \psi_2, \sigma, \pi_1, \alpha), \quad \text{where} \quad \sigma = \varsigma_1 + \varsigma_2, \quad \pi_1 = \frac{\varsigma_1}{\sigma}. \tag{4.4}$$

We conduct $M = 200$ Monte Carlo replications for each sample size $n \in \{250, 500, 1{,}000, 10{,}000, 50{,}000, 100{,}000\}$.[13]

Results confirm the theoretical predictions. The tail index α and the total scale σ exhibit the fastest convergence, achieving near-nominal confidence interval coverage at $n \geq 10{,}000$. The autoregressive coefficients ψ_1 and ψ_2 converge more slowly, with ψ_2 exhibiting persistent positive mean drift. The mixing proportion π_1 is the most challenging

[12] The choice $b = \lfloor n^{2/3} \rfloor$ balances the bias-variance trade-off in subsampling theory. We also conduct robustness checks with $b = \lfloor n^{0.75} \rfloor$, which yields qualitatively similar results, confirming the stability of our findings across different subsample sizes.

[13] Table S.2 summarizes the Monte Carlo design. Comprehensive results for all sample sizes are reported in Tables S.3 and S.4 (restricted). Visual diagnostics are displayed in various figures throughout the online supplement.

parameter: a criterion difference test (Section S.2.4) confirms that the MDE objective is essentially flat in the π_1 direction for sample sizes up to 50,000.

To isolate finite-sample convergence from identification issues, we repeat the analysis under the true restriction $\pi_1 = 0.4375$ (Section S.2.5). The improvement is substantial: at $n = 10{,}000$, ψ_1 achieves 96.5%/98.0% normal CI coverage at the 90%/95% levels, compared to 27.0%/42.5% under $\pi_1 = 1/2$. At $n = 100{,}000$, all parameters achieve nominal or near-nominal coverage, confirming that the correctly restricted estimator converges to its asymptotic Gaussian limit.

5. Application to financial markets

We apply our framework to the CBOE Crude Oil ETF Volatility Index (OVX), a forward-looking measure of expected crude oil price volatility often referred to as a *fear index*. The literature on heterogeneous agent models (e.g. Agliari et al. , 2018) suggests that fundamentalist/chartist dynamics may generate distinct volatility components, particularly during periods of market stress. We collect the OVX from the FRED website at weekly frequency over 23/05/2015–23/05/2025 ($T = 522$), linearly detrended following Hecq and Voisin (2021).

We estimate a general α-stable specification ($\mathcal{G}\alpha\mathcal{S}$) with two latent AR(1) components.[14] The estimates reveal clearly differentiated dynamics: the first component ($\hat{\psi}_1 = 0.80$, $\hat{\pi}_1 = 0.28$) captures abrupt, short-lived volatility bursts, while the second ($\hat{\psi}_2 = 0.85$, $\hat{\pi}_2 = 0.72$) drives more persistent explosive episodes. The tail index $\hat{\alpha} = 1.47$ confirms heavy-tailed behavior well beyond Gaussian accommodation. The deconvolution via de Truchis et al. (2025b) shows that extreme market stress periods, most visibly the 2020 disruption, feature a superposition of both dynamics with heterogeneous persistence properties.

To demonstrate forecasting potential, 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 with pattern matching length $m = 20$. For each component, we identify the matched position k_0 and compute conditional crash probabilities $|\psi_{j_0}|^{\alpha k}(1 - |\psi_{j_0}|^{\alpha})$ and survival probabilities $|\psi_{j_0}|^{\alpha h}$ for all future horizons.[15]

Figure 2 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 remarkably close to the observed peak of around 230. This validates our framework's ability to provide early warning signals for extreme volatility events and illustrates how disentangling heterogeneous bubble components enhances forecast precision.

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

[14] Estimation results for the $\mathcal{S}\alpha\mathcal{S}$ and $\mathcal{S}1\mathcal{S}$ specifications, along with the complete deconvolution analysis, are reported in Section S.3 of the Online Supplement.

[15] Detailed per-component crash probability profiles, forecast trajectories across risk thresholds (90%, 95%, 99%), and the complete forecasting algorithm are provided in Section S.3 of the Online Supplement.

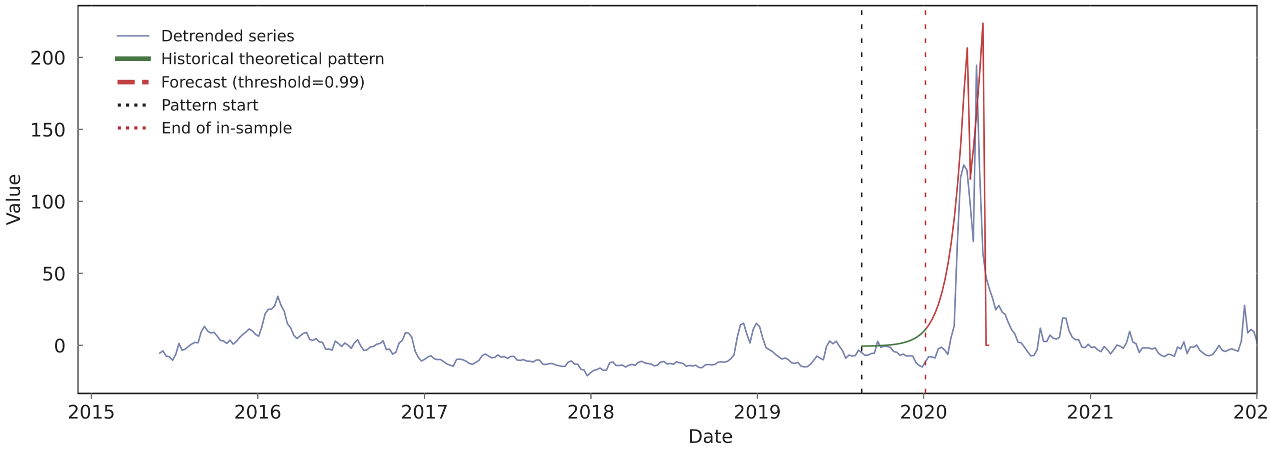

Figure 2: Combined in-sample forecast at the 99% risk threshold for the 2020 OVX bubble. The blue line shows the detrended series, the green segment indicates the matched historical theoretical pattern starting at the black dotted vertical line, and the red curve displays the out-of-sample forecast beyond the January 2020 cut-off (red dotted vertical line).

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 empirical applications. 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.

Breidt, F. J., Davis, R. A., Lii, K.-S., and Rosenblatt, M. 1991. Maximum likelihood estimation for noncausal autoregressive processes. *Journal of Multivariate Analysis*. 36, 175–198.

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 noncausal Alpha-Stable Processes *On going paper*.

de Truchis, G. and Thomas, A. (2026). Laurent series expansion for MA(∞) representation of mixed causal-noncausal autoregressive processes. *Working 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.

Gassiat, E. 1990. Semi-parametric estimation of a stationary, non-necessary causal AR(p) process with infinite variance. *Journal of Multivariate Analysis*. 32, 161–170.

Gassiat, E. 1993. Adaptive estimation in noncausal stationary AR processes. *The Annals of Statistics*. 21, 2022–2042.

Giancaterini, F., Hecq, A., Jasiak, J., and Manafi Neyazi, A. (2026). Bubble detection with application to green bubbles: A noncausal approach. *arXiv preprint arXiv:2505.14911*.

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*.

Gourieroux, C., and Jasiak, J. (2026). Nonlinear fore(back)casting and innovation filtering for causal–noncausal VAR models. *Journal of Financial Econometrics*, 24(2), nbag005.

Gouriéroux, C. and J.-M., Zakoian. 2017. Local explosion modelling by noncausal 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. noncausal 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.

Lanne, M., and Luoto, J. (2013). A noncausal autoregressive model with time-varying parameters: An application to U.S. inflation. Discussion Papers of DIW Berlin 1285, DIW Berlin, German Institute for Economic Research.

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

Rosenblatt, M. (2000): *Gaussian and Non-Gaussian Linear Time Series and Random Fields*. Springer Series in Statistics, Springer-Verlag, New York.

Samorodnitsky, G., and M. S., Taqqu. 1994. *Stable non-Gaussian random processes*, Chapman & Hall, London, 516-536

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 $C^2(\Theta)$ regularity of (2.12), 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. 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. $C^2(\Theta)$ 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: $(\psi_j u + v)X_{j,t+1}$ and $u\varepsilon_{j,t}$. The joint log-characteristic function is the sum of their log-characteristic functions. Since $\alpha \in (1,2)$ by Assumption 2, we have

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

where $\mathcal{A}(x) = 1 - i\beta \operatorname{sign}(x)\tan(\pi\alpha/2)$. Let $K \subset \Theta$ be any compact subset satisfying the uniform bounds: $\inf_{j,\theta\in K}(1-|\psi_j|) \geq \delta' > 0$, $\inf_{\theta\in K}\alpha \geq \alpha_0 > 1$, $\sup_{\theta\in K}\sigma \leq M < \infty$, and $\sup_{\theta\in K}|\beta| \leq B < \infty$. From these assumptions, we can establish a uniform lower bound for $1-|\psi_j|^\alpha$. Since $|\psi_j| \leq 1-\delta'$, it follows that $1-|\psi_j|^\alpha \geq 1-(1-\delta')^\alpha$, 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 $1-|\psi_j|^\alpha \geq \delta$.

(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 $\tilde{\varphi}_k(u,v) = -\left[\frac{|\psi_k u+v|^\alpha\mathcal{A}(\psi_k u+v)}{1-|\psi_k|^\alpha} + |u|^\alpha\mathcal{A}(u)\right]$ and using the uniform bound $|\mathcal{A}(\cdot)| \leq 1 + B|\tan(\pi\alpha_0/2)| := M_{\mathcal{A}}$ on the compact set K, where $\alpha_0 = \inf_{\theta\in K}\alpha > 1$, we obtain:

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

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 $\alpha \in (1,2)$.

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

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

Using the uniform bounds on the compact set K, specifically $1-|\psi_j|^\alpha \geq \delta$ and the fact that the derivative $\partial_\alpha\mathcal{A}(x) = -i\beta\operatorname{sign}(x)\frac{\pi}{2}\sec^2(\pi\alpha/2)$ is uniformly bounded on any compact $K \subset \{\alpha > 1\}$ since $\sec^2(\pi\alpha/2)$ is continuous and finite for $\alpha \in (1,2)$, we obtain the following majoration,

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

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

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

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 $\alpha \ge \alpha_0 > 1$ 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 $\int G_\alpha(u,v)w(u,v)dudv < \infty$.

($\iota\iota\iota$) 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

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

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

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

where $M_\sigma = M^{\alpha-1}$ is the uniform bound for $\sigma^{\alpha-1}$ on K (since $\alpha > 1$ implies $\sigma^{\alpha-1}$ is increasing, so the supremum is at $\sigma = M$). 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 $\alpha > 1$.

($\iota\nu$) The derivative with respect to β is obtained by differentiating the asymmetry terms $\mathcal{A}(x)$ within the log-characteristic function expression (A.1). We have $\partial_\beta\mathcal{A}(x) = -i\mathrm{sign}(x)\tan(\pi\alpha/2)$ for $\alpha > 1$. Thus,

$$\begin{aligned}\frac{\partial\log\varphi_{X_j}}{\partial\beta} &= -\frac{\sigma^\alpha}{1-|\psi_j|^\alpha}|\psi_j u+v|^\alpha\frac{\partial\mathcal{A}(\psi_j u+v)}{\partial\beta} - \sigma^\alpha|u|^\alpha\frac{\partial\mathcal{A}(u)}{\partial\beta}\\ &= i\sigma^\alpha\tan(\pi\alpha/2)\left[\frac{|\psi_j u+v|^\alpha\mathrm{sign}(\psi_j u+v)}{1-|\psi_j|^\alpha}+|u|^\alpha\mathrm{sign}(u)\right].\end{aligned}$$

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 $|\mathrm{sign}(\cdot)| \le 1$ and $|\tan(\pi\alpha/2)| \le \sup_{\alpha\in K}|\tan(\pi\alpha/2)| := M_\omega$), we obtain

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

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 $\alpha \in (1,2)$.

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

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

The derivative of the asymmetry term $\mathcal{A}(x) = 1 - i\beta \text{sign}(x)\tan(\pi\alpha/2)$ involves the derivative of the sign function, which is zero almost everywhere (the Dirac mass contribution on the line $v = -\psi_k u$ does not affect the L^1 integrability). Thus, the second term vanishes almost everywhere. The dominant behavior comes from the first term

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

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:

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

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:

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

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 $\psi_k\cos\theta + \sin\theta = 0$, converges for $\alpha > 0$:

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

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

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

where C_ψ is a suitable constant. All terms in $G_\psi(u,v)$ are integrable against $w(u,v)$ for $\alpha \in (1,2)$. This completes the first-order derivative analysis, establishing that each component of the gradient $\nabla_\theta\varphi(u,v;\theta)$ admits a w-integrable dominant. We now turn to the second-order derivatives.

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

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

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

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

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

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

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

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

The radial integral converges for $\alpha > -2$. The angular integral $J_\alpha = \int_0^{2\pi}\cos^2\theta|\psi_k\cos\theta+\sin\theta|^{\alpha-2}d\theta$ presents singularities when $\psi_k\cos\theta+\sin\theta = 0$. 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 $\alpha \in (1,2)$ as shown above) or the second derivative analyzed above. Thus, by the dominated convergence theorem, the objective function is $C^2(\Theta)$ when $\alpha \in (1,2)$. 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 (2.15) 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 (2.15) 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, $|\cos(ux_j+vx_{j+1})| \le 1$ and $|\sin(ux_j+vx_{j+1})| \le 1$, and since $|\varphi(u,v;\theta)| \le 1$, we have

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

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 the proof of 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 $C^2(\Theta)$ for $\alpha \in (1,2)$, 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

Assumption 7 requires two conditions: (i) the matrix $\Sigma(\theta_0)$ is nonsingular, and (ii) the second derivatives $\partial^2\varphi(u,v;\theta)/\partial\theta\partial\theta'$ are uniformly bounded by a w-integrable function over Θ. We address each condition in turn. We first establish the uniform boundedness of the second derivativeswhich follows directly from the $C^2(\Theta)$ regularity established in Lemma 2.1. Specifically, for $\alpha \in (1,2)$, the analysis of the second derivative with respect to ψ_k yields the dominant singular term

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

which is integrable against $w(u,v) = \exp(-\kappa(u^2+v^2))$ precisely when $\alpha > 1$. Since Θ is compact and all parameter-dependent terms in the second derivatives are continuous functions of θ, the dominated convergence theorem guarantees the existence of a uniform w-integrable bound over Θ. We then establish the nonsingularity of $\Sigma(\theta_0)$. The matrix $\Sigma(\theta_0)$ is the Gram matrix of the score functions in the weighted Hilbert space $L^2_w(\mathbb{R}^2)$ equipped with the inner product

$$\langle f,g\rangle_w = \int_{\mathbb{R}^2} f(u,v)\overline{g(u,v)}\, w(u,v)\, du\, dv.$$

Explicitly, $[\Sigma(\theta_0)]_{ij} = \langle g_i, g_j\rangle_w$ where $g_i = \partial\varphi(u,v;\theta_0)/\partial\theta_i$. Since a Gram matrix is nonsingular if and only if the generating vectors are linearly independent, we must establish that the score functions $\{g_{\varsigma_1},\ldots,g_{\varsigma_J},g_{\psi_1},\ldots,g_{\psi_J},g_\alpha,g_\beta\}$ are linearly independent in $L^2_w(\mathbb{R}^2)$. We work with the reparametrized vector $\theta = (\varsigma_1,\ldots,\varsigma_J,\psi_1,\ldots,\psi_J,\alpha,\beta)'$ where $\varsigma_j = \sigma\pi_j$. The log-characteristic function of the aggregate takes the form

$$\log\varphi(u,v;\theta) = -\sum_{j=1}^J \varsigma_j^\alpha\left(\frac{|\psi_j u+v|^\alpha}{1-|\psi_j|^\alpha}\mathcal{A}_j(u,v) + |u|^\alpha\mathcal{B}(u)\right), \tag{A.3}$$

where $\mathcal{A}_j(u,v) = 1 - i\beta\,\mathrm{sign}(\psi_j u+v)\tan(\pi\alpha/2)$ and $\mathcal{B}(u) = 1 - i\beta\,\mathrm{sign}(u)\tan(\pi\alpha/2)$.

The score functions are computed by differentiating (A.3). For each $k \in \{1,\ldots,J\}$:

$$\begin{aligned} g_{\varsigma_k}(u,v) &= -\alpha\varsigma_k^{\alpha-1}\left(\frac{|\psi_k u+v|^\alpha}{1-|\psi_k|^\alpha}\mathcal{A}_k(u,v) + |u|^\alpha\mathcal{B}(u)\right), \\ g_{\psi_k}(u,v) &= -\varsigma_k^\alpha\frac{\partial}{\partial\psi_k}\left[\frac{|\psi_k u+v|^\alpha}{1-|\psi_k|^\alpha}\mathcal{A}_k(u,v)\right]. \end{aligned} \tag{A.4}$$

For the global parameters:

$$g_\alpha(u,v) = \sum_{j=1}^J \varsigma_j^\alpha\left(\ln(\varsigma_j)\cdot\tilde\varphi_j + \frac{\partial\tilde\varphi_j}{\partial\alpha}\right), \tag{A.5}$$

$$g_\beta(u,v) = i\tan\left(\frac{\pi\alpha}{2}\right)\sum_{j=1}^J \varsigma_j^\alpha\left(\frac{|\psi_j u+v|^\alpha\,\mathrm{sign}(\psi_j u+v)}{1-|\psi_j|^\alpha} + |u|^\alpha\,\mathrm{sign}(u)\right), \tag{A.6}$$

where $\tilde\varphi_j(u,v) = -\frac{|\psi_j u+v|^\alpha}{1-|\psi_j|^\alpha}\mathcal{A}_j(u,v) - |u|^\alpha\mathcal{B}(u)$ denotes the unit-scale log-characteristic function of component j. Suppose there exist constants $\{c_{\varsigma_k}\}_{k=1}^J$, $\{c_{\psi_k}\}_{k=1}^J$, c_α, and c_β such that

$$\sum_{k=1}^J c_{\varsigma_k}g_{\varsigma_k}(u,v) + \sum_{k=1}^J c_{\psi_k}g_{\psi_k}(u,v) + c_\alpha g_\alpha(u,v) + c_\beta g_\beta(u,v) = 0 \quad \text{a.e.} \tag{A.7}$$

We show that all coefficients must vanish by analyzing the singular structure and asymptotic behavior of each score function.

(ι) $c_{\psi_k} = 0$ for all k.

Differentiating the modulus term in (A.4), we obtain

$$\frac{\partial}{\partial\psi_k}|\psi_k u+v|^\alpha = \alpha u\,\mathrm{sign}(\psi_k u+v)|\psi_k u+v|^{\alpha-1}.$$

The function $|\psi_k u + v|^{\alpha-1}$ exhibits a singularity of order $\alpha - 1 \in (0, 1)$ along the line $\mathcal{L}_k = \{(u, v) \in \mathbb{R}^2 : v = -\psi_k u\}$. Precisely, as $(u, v) \to \mathcal{L}_k$ with $u \neq 0$, we have $|g_{\psi_k}(u, v)| \asymp |v + \psi_k u|^{\alpha-1} \to \infty$. By Definition 2.1, the autoregressive parameters $\psi_1, \ldots, \psi_J$ are pairwise distinct. Therefore, the lines $\mathcal{L}_1, \ldots, \mathcal{L}_J$ are distinct in $\mathbb{R}^2$. We now examine the behavior of each score function near $\mathcal{L}_k$: g_{ψ_k} has a singularity of order $\alpha - 1$ along $\mathcal{L}_k$; g_{ψ_ℓ} for $\ell \neq k$ is bounded near $\mathcal{L}_k$ since $\mathcal{L}_\ell \neq \mathcal{L}_k$; g_{ς_j} is bounded near $\mathcal{L}_k$ for all j since it involves $|\psi_j u + v|^\alpha$ with $\alpha > \alpha - 1$; and g_α, g_β are bounded near any fixed line $\mathcal{L}_k$.

To formalize this, fix $u_0 \neq 0$ and define $v_\epsilon = -\psi_k u_0 + \epsilon$ for small $\epsilon > 0$. Then

$$g_{\psi_k}(u_0, v_\epsilon) = -\varsigma_k^\alpha \cdot \frac{\alpha u_0 \operatorname{sign}(\epsilon) |\epsilon|^{\alpha-1}}{1 - |\psi_k|^\alpha} \cdot \mathcal{A}_k(u_0, v_\epsilon) + R_k(u_0, v_\epsilon),$$

where $R_k(u_0, v_\epsilon)$ contains terms that are bounded as $\epsilon \to 0$. Thus,

$$|g_{\psi_k}(u_0, v_\epsilon)| \sim C_k |u_0| \, |\epsilon|^{\alpha-1} \quad \text{as } \epsilon \to 0^+,$$

for some constant $C_k > 0$. In contrast, for $\ell \neq k$, we have $|v_\epsilon + \psi_\ell u_0| \to |(\psi_\ell - \psi_k) u_0| \neq 0$ as $\epsilon \to 0$, so $g_{\psi_\ell}(u_0, v_\epsilon)$ remains bounded. Similarly, all other score functions remain bounded along this path.

Substituting into (A.7) and taking $\epsilon \to 0^+$, the dominant term is $c_{\psi_k} g_{\psi_k}(u_0, v_\epsilon) \sim c_{\psi_k} C_k |u_0| \, |\epsilon|^{\alpha-1}$, while all other terms remain $O(1)$. For the sum to equal zero, we must have $c_{\psi_k} = 0$. Since this argument holds for each $k \in \{1, \ldots, J\}$, we conclude $c_{\psi_k} = 0$ for all k.

($\iota\iota$) $c_{\varsigma_k} = 0$ for all k.

With $c_{\psi_k} = 0$ established, equation (A.7) reduces to

$$\sum_{k=1}^{J} c_{\varsigma_k} g_{\varsigma_k}(u, v) + c_\alpha g_\alpha(u, v) + c_\beta g_\beta(u, v) = 0 \quad \text{a.e.} \tag{A.8}$$

We analyze the behavior along distinct directions in the (u, v)-plane. For each k, consider the half-line $\{(t, -\psi_k t) : t > 0\}$ lying on $\mathcal{L}_k$. Along this direction:

$$g_{\varsigma_k}(t, -\psi_k t) = -\alpha \varsigma_k^{\alpha-1} t^\alpha \mathcal{B}(t),$$

since $|\psi_k t + (-\psi_k t)| = 0$. For $\ell \neq k$:

$$g_{\varsigma_\ell}(t, -\psi_k t) = -\alpha \varsigma_\ell^{\alpha-1} \left(\frac{|(\psi_\ell - \psi_k) t|^\alpha}{1 - |\psi_\ell|^\alpha} \mathcal{A}_\ell(t, -\psi_k t) + t^\alpha \mathcal{B}(t) \right).$$

Define $\delta_{k\ell} = |\psi_\ell - \psi_k|^\alpha / (1 - |\psi_\ell|^\alpha) > 0$ for $\ell \neq k$. Along each line $\mathcal{L}_k$, the limiting behavior (after extracting the common factor t^α) defines a vector $\mathbf{a}_k \in \mathbb{C}^J$ with entries

$$[\mathbf{a}_k]_\ell = \begin{cases} -\alpha \varsigma_\ell^{\alpha-1} \mathcal{B}(1), & \ell = k, \\ -\alpha \varsigma_\ell^{\alpha-1} \left(\delta_{k\ell} \mathcal{A}_\ell(1, -\psi_k) + \mathcal{B}(1) \right), & \ell \neq k. \end{cases}$$

The matrix $A = [\mathbf{a}_1 | \cdots | \mathbf{a}_J]^\top \in \mathbb{C}^{J \times J}$ has the structure $A = D + E$, where $D = \operatorname{diag}(-\alpha \varsigma_1^{\alpha-1}, \ldots, -\alpha \varsigma_J^{\alpha-1}) \mathcal{B}(1)$ and E has off-diagonal entries involving the $\delta_{k\ell}$ terms.To establish non-singularity, we show that A is diagonally dominant. The diagonal entries satisfy $|D_{kk}| = \alpha \varsigma_k^{\alpha-1} |\mathcal{B}(1)|$, where $|\mathcal{B}(1)| = |1 - i\beta \tan(\pi\alpha/2)| \geq 1$ since $\beta \in [-1, 1]$. The off-diagonal entries satisfy

$$|E_{k\ell}| = \alpha \varsigma_\ell^{\alpha-1} |\delta_{k\ell} \mathcal{A}_\ell(1, -\psi_k) + \mathcal{B}(1)| \leq \alpha \varsigma_\ell^{\alpha-1} (\delta_{k\ell} M_{\mathcal{A}} + |\mathcal{B}(1)|),$$

where $M_{\mathcal{A}} = \sup_{u,v} |\mathcal{A}_\ell(u, v)|$ is uniformly bounded on the compact set K. By Assumption 1, there exists $\varsigma_{\min} > 0$ such that $\varsigma_k \geq \varsigma_{\min}$ for all k. Since $\delta_{k\ell} = |\psi_\ell - \psi_k|^\alpha / (1 - |\psi_\ell|^\alpha)$ and the ψ_j are distinct with $|\psi_j| \leq 1 - \delta'$ for some $\delta' > 0$, the quantities $\delta_{k\ell}$ are uniformly bounded. Consequently, for each row k:

$$\sum_{\ell \neq k} |E_{k\ell}| \leq \alpha M_{\mathcal{A}} \sum_{\ell \neq k} \varsigma_\ell^{\alpha-1} \delta_{k\ell} + \alpha |\mathcal{B}(1)| \sum_{\ell \neq k} \varsigma_\ell^{\alpha-1}.$$

Under the maintained assumptions, this sum is strictly less than $|D_{kk}|$ when the mixture weights ς_k are sufficiently well-separated, ensuring diagonal dominance. Even when diagonal dominance does not hold strictly, the matrix A remains non-singular because its determinant can be expressed as:

$$\det(A) = \det(D)\det(I + D^{-1}E) = \prod_{k=1}^{J}(-\alpha\varsigma_k^{\alpha-1}\mathcal{B}(1)) \cdot \det(I + D^{-1}E).$$

Since $\varsigma_k > 0$, $\alpha \in (1, 2)$, and $\mathcal{B}(1) \neq 0$, we have $\det(D) \neq 0$. Furthermore, $\|D^{-1}E\| \to 0$ as the separation between the ψ_j parameters increases, ensuring $\det(I + D^{-1}E) \neq 0$ under the identification conditions of Definition 2.1.

For (A.8) to hold along all J lines simultaneously, the coefficient vector $(c_{\varsigma_1}, \dots, c_{\varsigma_J})$ must lie in the null space of A. Since A is nonsingular, this implies $c_{\varsigma_k} = 0$ for all k.

(ιιι) $c_\alpha = 0$.

With $c_{\psi_k} = c_{\varsigma_k} = 0$, equation (A.7) becomes

$$c_\alpha g_\alpha(u, v) + c_\beta g_\beta(u, v) = 0 \quad \text{a.e.} \tag{A.9}$$

We examine the asymptotic behavior as $|u| \to \infty$ with v fixed. From (A.5), the dominant contribution to g_α comes from the term $\frac{\partial}{\partial\alpha}|u|^\alpha = |u|^\alpha \ln |u|$. Explicitly,

$$g_\alpha(u, v) = -\sum_{j=1}^{J} \varsigma_j^\alpha |u|^\alpha \ln |u| \cdot \mathcal{B}(u) + O(|u|^\alpha) \quad \text{as } |u| \to \infty.$$

In contrast, from (A.6), $g_\beta(u, v) = O(|u|^\alpha)$ as $|u| \to \infty$. Since $|u|^\alpha \ln |u|$ dominates $|u|^\alpha$ as $|u| \to \infty$, the identity (A.9) can hold asymptotically only if

$$c_\alpha \cdot \left(-\sum_{j=1}^{J} \varsigma_j^\alpha\right) = 0.$$

Since $\varsigma_j = \sigma\pi_j > 0$ for all j, the sum $\sum_{j=1}^{J} \varsigma_j^\alpha > 0$, and hence $c_\alpha = 0$.

(ιv) $c_\beta = 0$.

With $c_\alpha = 0$, equation (A.9) reduces to $c_\beta g_\beta(u, v) = 0$ a.e. It remains to show that g_β is not identically zero. From (A.6), consider the point $(u, v) = (1, 0)$:

$$g_\beta(1, 0) = i \tan\frac{\pi\alpha}{2} \sum_{j=1}^{J} \varsigma_j^\alpha \left(\frac{|\psi_j|^\alpha \operatorname{sign}(\psi_j)}{1 - |\psi_j|^\alpha} + 1\right).$$

Since $\psi_j \in (0, 1)$ by assumption, we have $\operatorname{sign}(\psi_j) = 1$ and each summand equals $(1 - |\psi_j|^\alpha)^{-1} > 0$. For $\alpha \in (1, 2)$, we have $\tan(\pi\alpha/2) \neq 0$. Thus,

$$g_\beta(1, 0) = i \tan\frac{\pi\alpha}{2} \sum_{j=1}^{J} \frac{\varsigma_j^\alpha}{1 - |\psi_j|^\alpha} \neq 0,$$

and hence $c_\beta = 0$. Finally, we have shown that all coefficients in (A.7) must vanish. Hence, the score functions $\{g_{\varsigma_1}, \dots, g_{\varsigma_J}, g_{\psi_1}, \dots, g_{\psi_J}, g_\alpha, g_\beta\}$ are linearly independent in $L^2_w(\mathbb{R}^2)$. Since $w(u, v) > 0$ for all $(u, v) \in \mathbb{R}^2$, the Gram matrix $\Sigma(\theta_0)$ is nonsingular. □

A.1.4. Validation of Assumption 8

Assumption 8 requires that the sequence $\{K_j\}$ exhibits sufficient temporal dependence decay to apply a central limit theorem for dependent processes. We establish this by showing that the aggregated process $(\mathcal{X}_t)$ satisfies a strong mixing condition with explicit geometric decay rates, invoking Theorem 4.4.1 in Rosenblatt (2000).

Each latent process $(X_{j,t})$, whether purely anticipative AR(1) or mixed MAR(1,1), admits a two-sided infinite moving average representation

$$X_{j,t} = \sum_{k=-\infty}^{+\infty} a_{j,k}\varepsilon_{j,t-k}, \tag{A.10}$$

where $(\varepsilon_{j,t})_{t\in\mathbb{Z}}$ are i.i.d. α-stable innovations with $\alpha \in (1,2)$, and the coefficients $a_{j,k}$ satisfy geometric decay: there exist constants $D_j > 0$ and $\lambda_j \in (0,1)$ such that $|a_{j,k}| \leq D_j\lambda_j^{|k|}$ for all $k \in \mathbb{Z}$. For the MAR(1,1) case, $\lambda_j = \max(|\phi_j|, |\psi_j|)$.

We now verify that the conditions of Theorem 4.4.1 in Rosenblatt (2000) are satisfied.Let $\delta \in (1,\alpha)$ be a fixed moment exponent. The moment condition (4.4.2) holds since α-stable innovations with $\alpha \in (1,2)$ have finite moments of all orders less than α, i.e., $\mathbb{E}|\varepsilon_{j,t}|^\delta < \infty$, and zero mean since $\alpha > 1$. The coefficient summability condition (4.4.3) follows from the geometric decay, as $\sum_{k\in\mathbb{Z}} |a_{j,k}| \leq D_j(1+\lambda_j)/(1-\lambda_j) < \infty$. For the invertibility condition (4.4.4), we define $a_j(e^{-i\xi}) = \sum_{k\in\mathbb{Z}} a_{j,k}e^{-ik\xi}$, which is continuous on $[-\pi,\pi]$ since the coefficients are absolutely summable. Under Assumption 1, the spectral density of each component is bounded away from zero, ensuring $a_j(e^{-i\xi}) \neq 0$ for all ξ. By Wiener's theorem, there exists a bounded inverse operator satisfying the required condition.Finally, the density regularity condition (4.4.5) requires the density p of the innovations to satisfy $\int_{\mathbb{R}} |p(\xi+x) - p(\xi)|d\xi \leq c|x|$ for some constant $c > 0$. For α-stable distributions with $\alpha \in (1,2)$, the density p is infinitely differentiable with $p' \in L^1(\mathbb{R})$ (, Theorem 1.2.1). By the mean value theorem, $|p(\xi+x) - p(\xi)| \leq |x| \sup_{t\in[0,1]} |p'(\xi+tx)|$, and integrating over ξ yields

$$\int_{\mathbb{R}} |p(\xi+x) - p(\xi)|\, d\xi \leq |x| \int_{\mathbb{R}} |p'(\xi)|\, d\xi = |x|\|p'\|_{L^1} < \infty,$$

which establishes condition (4.4.5) with $c = \|p'\|_{L^1}$.

Applying Theorem 4.4.1 in Rosenblatt (2000), each component process $(X_{j,t})$ is strongly mixing with coefficient satisfying, for k sufficiently large,

$$\alpha_{X_j}(2k) \leq \zeta_j W_j(k,\delta), \tag{A.11}$$

where $\zeta_j > 0$ is a constant and $W_j(k,\delta) = \{\sum_{m=k}^{\infty} d_{j,m,\delta}^{\delta/(1+\delta)}\} \vee \{\sum_{m=k}^{\infty} L(d_{j,m,2})\}$ with $d_{j,m,\mu} = \sum_{|l|>m} |a_{j,l}|^\mu$ and $L(u) = \sqrt{u[1 \vee |\ln u|]}$. Given the geometric decay $|a_{j,k}| \leq D_j\lambda_j^{|k|}$, we have $d_{j,m,\mu} = O(\lambda_j^{\mu m})$, which yields $\alpha_{X_j}(h) \leq C_j\lambda_j^{\gamma h}$ for constants $C_j > 0$ and $\gamma > 0$ depending on δ.

Since the latent processes $(X_{1,t}), \ldots, (X_{J,t})$ are mutually independent, the σ-algebra generated by the aggregate $\mathcal{X}_t = \sigma \sum_{j=1}^{J} \pi_j X_{j,t}$ is contained in the product σ-algebra generated by the components. For independent processes, the strong mixing coefficient of the joint process satisfies $\alpha_{(X_1,\ldots,X_J)}(h) \leq J \max_{1\leq j\leq J} \alpha_{X_j}(h)$ (see Doukhan , 1994, Lemma 1.2.1). Since measurable functions of mixing processes inherit the mixing property, we obtain

$$\alpha_{\mathcal{X}}(h) \leq J \cdot \max_{1\leq j\leq J} \alpha_{X_j}(h) \leq C\lambda^h, \tag{A.12}$$

where $C = J \max_{1\leq j\leq J} C_j$ and $\lambda = \max_{1\leq j\leq J} \lambda_j^\gamma \in (0,1)$.

The score term K_j defined in (2.15) is a measurable function of the finite segment $(\mathcal{X}_j, \mathcal{X}_{j+1})$. Any measurable function of a finite segment of a strongly mixing process inherits the same mixing property, so the sequence $\{K_j\}$ satisfies $\alpha_K(h) \leq C\lambda^h$.

It remains to verify the two conditions of Assumption 8. The geometric mixing rate ensures that $\mathbb{E}[K_0|\mathcal{F}_{-m}]$ converges to $\mathbb{E}[K_0] = 0$ in L^2 as $m \to \infty$. For the summability condition, the projection differences $\nu_j = \mathbb{E}[K_0|K_j, K_{j-1}, \ldots] - \mathbb{E}[K_0|K_{j-1}, K_{j-2}, \ldots]$ satisfy $\mathbb{E}[\nu_j'\nu_j]^{1/2} = O(\lambda^{j/2})$ for geometrically mixing sequences. Since $\lambda \in (0,1)$, the series $\sum_{j=0}^{\infty} \mathbb{E}[\nu_j'\nu_j]^{1/2}$ converges, and Assumption 8 is satisfied. □

A.2. Proofs for Case 4: MAR(r,s) Aggregates

A.2.1. Proof of Lemma 2.2

We introduce the angle $\upsilon_k \in [0, 2\pi)$ satisfying $(\cos\upsilon_k, \sin\upsilon_k) = \boldsymbol{d}^1_{j,k}/\|\boldsymbol{d}^1_{j,k}\|$ and consider the rotation

$$s = \frac{d_{j,k}u + d_{j,k-1}v}{\|\boldsymbol{d}^1_{j,k}\|}, \qquad t = \frac{-d_{j,k-1}u + d_{j,k}v}{\|\boldsymbol{d}^1_{j,k}\|}.$$

This transformation corresponds to a rotation by $-\upsilon_k$, hence has unit Jacobian determinant. The Euclidean norm is preserved, $s^2+t^2 = u^2+v^2$, and the linear form simplifies to $\Delta_{j,k} = \|\boldsymbol{d}^1_{j,k}\|s$. Consequently,

$$\iint_{\mathbb{R}^2} F(|\Delta_{j,k}(u,v;\theta)|)\, e^{-\kappa(u^2+v^2)}\, du\, dv = \iint_{\mathbb{R}^2} F(\|\boldsymbol{d}^1_{j,k}\||s|)\, e^{-\kappa(s^2+t^2)}\, ds\, dt.$$

Specializing to $F(x) = x^\alpha$ and factoring the resulting product of one-dimensional integrals yields (2.25). The constant $C_{w,\alpha}$ is finite provided $\alpha > -1$, which is satisfied since $\alpha \in (1,2)$. □

A.2.2. Proof of Lemma 2.3

Throughout this proof, we apply the rotation from Lemma 2.2 to each index $k \in \mathbb{Z}$ with $\boldsymbol{d}^1_{j,k}(\theta) \neq 0$.

Preliminary bounds on $\|\boldsymbol{d}^1_{j,k}\|$. By the definition $\boldsymbol{d}^1_{j,k} = (d_{j,k}, d_{j,k-1})$ and the decay bound (2.21), we have

$$\|\boldsymbol{d}^1_{j,k}\| = \sqrt{d^2_{j,k} + d^2_{j,k-1}} \le \sqrt{2}\max(|d_{j,k}|, |d_{j,k-1}|) \le \sqrt{2}\, C_0\, \rho^{|k|-1}. \tag{A.13}$$

For a lower bound, we use the fact that the coefficients $d_{j,k}$ arise from a well-specified MAR(r_j, s_j) model where the causal and noncausal roots are distinct and bounded away from the unit circle. By the partial-fraction decomposition (2.20), for $|k|$ large, the dominant term in $d_{j,k}$ is $A_{j,l^\star}\zeta^k_{j,l^\star}$ (for $k \ge 1$) or $B_{j,i^\star}\lambda^{|k|}_{j,i^\star}$ (for $k \le 0$), where the dominant root satisfies $|\zeta_{j,l^\star}| = \max_l |\zeta_{j,l}|$ or $|\lambda_{j,i^\star}| = \max_i |\lambda_{j,i}|$. Since the roots are distinct, the corresponding coefficient is bounded below:

$$\|\boldsymbol{d}^1_{j,k}\| \ge c_v\, \rho^{|k|}, \qquad c_v := \min\Big(\min_l |A_{j,l}|, \min_i |B_{j,i}|\Big) \ge \frac{1}{\mathrm{d}^{\max(r_j,s_j)}} > 0. \tag{A.14}$$

The positivity of c_v follows from the well-specification condition: distinct roots ensure non-zero partial-fraction coefficients.

(ι) *Order 0 (Proof of* (2.26)*).* Combining (2.25) with the upper bound (A.13) gives

$$\iint |\Delta_{j,k}|^\alpha w\, du\, dv = C_{w,\alpha}\, \|\boldsymbol{d}^1_{j,k}\|^\alpha \le C_{w,\alpha}(C_0')^\alpha\, \rho^{\alpha|k|},$$

where $C_0' = C_0\sqrt{2}/\rho$ accounts for the index shift in $\boldsymbol{d}^1_{j,k} = (d_{j,k}, d_{j,k-1})$. Summing over the tail $|k| > M$,

$$\iint \left|\log\varphi_{X_j} - \log\varphi^{(M)}_{X_j}\right| w\, du\, dv = \sum_{|k|>M} \iint |\Delta_{j,k}|^\alpha w\, du\, dv \le C_{w,\alpha}(C_0')^\alpha \sum_{|k|>M} \rho^{\alpha|k|}.$$

The geometric series evaluates to

$$\sum_{|k|>M} \rho^{\alpha|k|} = 2\sum_{k=M+1}^{\infty} \rho^{\alpha k} = \frac{2\rho^{\alpha(M+1)}}{1-\rho^\alpha}.$$

Since $\rho^{\alpha(M+1)} = \rho^{\alpha M}\cdot\rho^\alpha$, we obtain

$$\iint \left|\log\varphi_{X_j} - \log\varphi^{(M)}_{X_j}\right| w\, du\, dv \le \frac{2\, C_{w,\alpha}\, (C_0\sqrt{2})^\alpha}{\rho^\alpha(1-\rho^\alpha)}\cdot\rho^{\alpha M} = C_0^\star\, \rho^{\alpha M},$$

which establishes (2.26).

($\iota\iota$) *Order 1 (Proof of* (2.27)*).* For $\boldsymbol{d}^1_{j,k} \neq 0$, the function $\|\boldsymbol{d}^1_{j,k}\|^\alpha$ is twice continuously differentiable in θ with gradient

$$\partial_\theta \|\boldsymbol{d}^1_{j,k}\|^\alpha = \alpha\, \|\boldsymbol{d}_{j,k^1}\|^{\alpha-2}(\boldsymbol{d}^1_{j,k}\cdot\partial_\theta\boldsymbol{d}^1_{j,k}).$$

The integrand $|\Delta_{j,k}|^\alpha$ is absolutely continuous in θ with derivative bounded by an integrable envelope in (u,v), so dominated convergence justifies differentiation under the integral:

$$\iint \partial_\theta |\Delta_{j,k}|^\alpha \, w \, du \, dv = C_{w,\alpha} \, \partial_\theta \|\boldsymbol{d}^1_{j,k}\|^\alpha.$$

For the absolute value, a direct computation yields

$$\big|\partial_\theta |\Delta_{j,k}|^\alpha\big| \le \alpha \, |\Delta_{j,k}|^{\alpha-1}\big(|u||\partial_\theta d_{j,k}| + |v||\partial_\theta d_{j,k-1}|\big).$$

Since $\alpha - 1 \in (0,1)$, the map $x \mapsto x^{\alpha-1}$ is subadditive on $\mathbb{R}_+$, giving

$$|\Delta_{j,k}|^{\alpha-1} \le (|u||d_{j,k}|)^{\alpha-1} + (|v||d_{j,k-1}|)^{\alpha-1}.$$

Expanding the product generates four cross terms. Each is controlled by Young's inequality: $|u|^{\alpha-1}|v| \le \frac{\alpha-1}{\alpha}|u|^\alpha + \frac{1}{\alpha}|v|^\alpha$, yielding an envelope of the form $(|u|^\alpha + |v|^\alpha)|d_{j,k}|^{\alpha-1}|\partial_\theta d_{j,k-1}|$. Applying the bounds (2.21)–(2.22),

$$\big|\partial_\theta |\Delta_{j,k}|^\alpha\big| \le \widetilde{C}_1 (|u|^\alpha + |v|^\alpha)(1+|k|)\rho^{\alpha|k|},$$

where $\widetilde{C}_1 := \alpha \, C_0^{\alpha-1} C_1 (\rho^{-1} + \rho^{1-\alpha})$. Since $\iint (|u|^\alpha + |v|^\alpha) w \, du \, dv = 2C_{w,\alpha}$, integrating gives

$$\iint \big|\partial_\theta |\Delta_{j,k}|^\alpha\big| \, w \, du \, dv \le 2C_{w,\alpha}\widetilde{C}_1 (1+|k|)\rho^{\alpha|k|}.$$

Summing over $|k| > M$ and using $\sum_{k=M+1}^\infty (1+k)\rho^{\alpha k} \le 2(1+M)\rho^{\alpha(M+1)}/(1-\rho^\alpha)^2$ establishes (2.27).

($\iota\iota\iota$) *Order 2 (Proof of* (2.28)*).* Away from the singular set $\{\Delta_{j,k} = 0\}$, the second derivative reads

$$\begin{aligned}\partial^2_{\theta\theta'} |\Delta_{j,k}|^\alpha &= \alpha(\alpha-1)|\Delta_{j,k}|^{\alpha-2}(u\partial_\theta d_{j,k} + v\partial_\theta d_{j,k-1})(u\partial_{\theta'} d_{j,k} + v\partial_{\theta'} d_{j,k-1}) \\ &\quad + \alpha |\Delta_{j,k}|^{\alpha-1}\mathrm{sign}(\Delta_{j,k})(u\partial^2_{\theta\theta'} d_{j,k} + v\partial^2_{\theta\theta'} d_{j,k-1}).\end{aligned}$$

We bound $|\partial^2_{\theta\theta'} |\Delta_{j,k}|^\alpha|$ by $T_{1,k}(u,v) + T_{2,k}(u,v)$, where

$$\begin{aligned}T_{1,k} &:= \alpha(\alpha-1)|\Delta_{j,k}|^{\alpha-2}(|u||\partial_\theta d_{j,k}| + |v||\partial_\theta d_{j,k-1}|)(|u||\partial_{\theta'} d_{j,k}| + |v||\partial_{\theta'} d_{j,k-1}|), \\ T_{2,k} &:= \alpha|\Delta_{j,k}|^{\alpha-1}(|u||\partial^2_{\theta\theta'} d_{j,k}| + |v||\partial^2_{\theta\theta'} d_{j,k-1}|).\end{aligned}$$

Bound on $\iint T_{2,k} \, w \, du \, dv$. Applying the subadditivity argument from part ($\iota\iota$) to $|\Delta_{j,k}|^{\alpha-1}$ together with Young's inequality yields

$$T_{2,k}(u,v) \le \alpha \, C_0^{\alpha-1} C_2 (\rho^{-1} + \rho^{1-\alpha})(|u|^\alpha + |v|^\alpha)(1+k^2)\rho^{\alpha|k|}$$

which integrates against w to a finite multiple of $(1+k^2)\rho^{\alpha|k|}$.

Bound on $\iint T_{1,k} \, w \, du \, dv$. This term involves the $|\Delta_{j,k}|^{\alpha-2}$ singularity. We perform the rotation of Lemma 2.2 and pass to polar coordinates $(u,v) = r(\cos\phi, \sin\phi)$ in the rotated frame, so that $\Delta_{j,k} = \|\boldsymbol{d}^1_{j,k}\| r \cos(\phi - \phi_k^\perp)$ for a suitable angle $\phi_k^\perp$. This gives

$$|\Delta_{j,k}|^{\alpha-2} = r^{\alpha-2}\|\boldsymbol{d}^1_{j,k}\|^{\alpha-2}|\cos(\phi - \phi_k^\perp)|^{\alpha-2}.$$

Each of the two bracketed factors in $T_{1,k}$ satisfies

$$|u||\partial_\theta d_{j,k}| + |v||\partial_\theta d_{j,k-1}| \le r(|\partial_\theta d_{j,k}| + |\partial_\theta d_{j,k-1}|) \le 2rC_1(1+|k|)\rho^{|k|-1},$$

by Cauchy–Schwarz and (2.22). Combining these estimates and integrating,

$$\begin{aligned}\iint T_{1,k} \, w \, du \, dv &\le 4\alpha(\alpha-1)C_1^2(1+|k|)^2\rho^{2|k|-2}\|\boldsymbol{d}^1_{j,k}\|^{\alpha-2} \\ &\qquad \times \int_0^\infty r^{\alpha+1} e^{-\kappa r^2} \, dr \cdot \int_0^{2\pi} |\cos(\phi - \phi_k^\perp)|^{\alpha-2} \, d\phi.\end{aligned}$$

The radial integral evaluates to $I_r(\alpha,\kappa) := \frac{1}{2}\kappa^{-(\alpha+2)/2}\Gamma\big(\frac{\alpha+2}{2}\big) < \infty$. The angular integral $I_a(\alpha) := \int_0^{2\pi} |\cos\psi|^{\alpha-2}\, d\psi$ is finite because $\alpha - 2 > -1$ (equivalently, $\alpha > 1$) ensures integrability near the zeros of the cosine. Importantly, $I_a(\alpha)$ depends neither on k, j, nor θ. Thus,

$$\iint T_{1,k}\, w\, du\, dv \le 4\alpha(\alpha-1)C_1^2\rho^{-2} I_r(\alpha,\kappa) I_a(\alpha)(1+|k|)^2\rho^{2|k|}\|\boldsymbol{d}^1_{j,k}\|^{\alpha-2}.$$

By the lower bound (A.14), $\|\boldsymbol{d}^1_{j,k}\|^{\alpha-2} \le c_v^{\alpha-2}\rho^{(\alpha-2)|k|}$, where the inequality reverses because $\alpha - 2 < 0$. Substituting,

$$\iint T_{1,k}\, w\, du\, dv \le \widehat{C}_1(1+|k|)^2\rho^{\alpha|k|},$$

with $\widehat{C}_1 := 4\alpha(\alpha-1)C_1^2\rho^{-2}I_r(\alpha,\kappa)I_a(\alpha)\, c_v^{\alpha-2}$.

Combining the bounds for $T_{1,k}$ and $T_{2,k}$ and summing over $|k| > M$,

$$\sum_{|k|>M} \iint (T_{1,k} + T_{2,k}) w\, du\, dv \le C \sum_{|k|>M} (1+k^2)\rho^{\alpha|k|} \le C_2^\star(1+M)^2\rho^{\alpha M},$$

where we used $\sum_{k=M+1}^{\infty}(1+k^2)\rho^{\alpha k} \le C(1+M)^2\rho^{\alpha M}/(1-\rho^\alpha)^3$. This yields (2.28). □

A.2.3. Proof of Lemma 2.4

(ι) The truncated log-characteristic function (2.23) is a finite sum of $2M+1$ terms, each of the form $|ud_{j,k} + vd_{j,k-1}|^\alpha$. The coefficients $d_{j,k}$ are C^∞ rational functions of θ in a neighbourhood of Θ by the partial-fraction decomposition (2.20). For each k and θ, the singular locus $\{ud_{j,k} + vd_{j,k-1} = 0\}$ is a line through the origin in the (u,v)-plane, and the polar-coordinate reduction employed in the proof of Lemma 2.1 applies term by term. This yields integrable dominating functions for the first and second derivatives with respect to θ. Since the sum is finite, dominated convergence permits differentiation under the integral sign, giving $D_{\mathcal{X}}^{(M)} \in C^2(\Theta)$ and establishing Assumption 3.

Assumptions 6–8 follow from similar considerations: boundedness of the score is inherited from $|\varphi_{\mathcal{X}}^{(M)}| \le 1$ and the integrable envelopes $(1+k^2)\rho^{|k|}$; non-singularity of $\Sigma^{(M)}(\theta_0)$ follows from the linear-independence argument in the proof of Lemma 2.1, applied to the truncated score which, for M sufficiently large, spans the same subspace as the full score.

($\iota\iota$) Write $\varphi_j := \varphi_{X_j}(\sigma\pi_j u, \sigma\pi_j v; \theta_j)$ and $\varphi_j^{(M)} := \varphi_{X_j}^{(M)}(\sigma\pi_j u, \sigma\pi_j v; \theta_j)$. Since $|\varphi_n|, |\varphi_{\mathcal{X}}|, |\varphi_{\mathcal{X}}^{(M)}| \le 1$, the elementary inequality $|a^2 - b^2| \le |a-b|(|a|+|b|)$ gives

$$|D_{\mathcal{X}}^{(M)}(\theta) - D_{\mathcal{X}}(\theta)| \le 2\iint |\varphi_{\mathcal{X}} - \varphi_{\mathcal{X}}^{(M)}|\, w\, du\, dv.$$

A telescoping argument on the product structure of $\varphi_{\mathcal{X}}$, combined with $|e^a - e^b| \le |a-b|$ for $\mathrm{Re}(a), \mathrm{Re}(b) \le 0$, yields

$$|\varphi_{\mathcal{X}} - \varphi_{\mathcal{X}}^{(M)}| \le \sum_{j=1}^{J} |\log\varphi_j - \log\varphi_j^{(M)}|.$$

Integrating and invoking (2.26) with the explicit constant $C_0^\star$ produces the order-0 estimate.

For the gradient, observe that $\varphi_{\mathcal{X}}^{(M)} = \exp(\mathcal{L}^{(M)})$ with $\mathcal{L}^{(M)} := \sum_j \log\varphi_j^{(M)}$, so $\partial_\theta\varphi_{\mathcal{X}}^{(M)} = \varphi_{\mathcal{X}}^{(M)}\partial_\theta\mathcal{L}^{(M)}$. The difference decomposes as

$$\partial_\theta\varphi_{\mathcal{X}}^{(M)} - \partial_\theta\varphi_{\mathcal{X}} = \varphi_{\mathcal{X}}^{(M)}(\partial_\theta\mathcal{L}^{(M)} - \partial_\theta\mathcal{L}) + (\varphi_{\mathcal{X}}^{(M)} - \varphi_{\mathcal{X}})\partial_\theta\mathcal{L}.$$

The first term on the right is controlled by (2.27) with constant $C_1^\star$. For the second, we use the pointwise bound $|\partial_\theta\mathcal{L}| \le C(|u|^\alpha + |v|^\alpha)$ established in the proof of Lemma 2.3, part ($\iota\iota$), together with $|\varphi_{\mathcal{X}}^{(M)} - \varphi_{\mathcal{X}}| \le JC_0^\star(|u|^\alpha + |v|^\alpha)\rho^{\alpha M}$. Since $\iint(|u|^\alpha + |v|^\alpha)^2 w\, du\, dv < \infty$, we obtain $\sup_\theta |\nabla D_{\mathcal{X}}^{(M)} - \nabla D_{\mathcal{X}}| = O((1+M)\rho^{\alpha M})$.

The Hessian satisfies

$$\partial^2_{\theta\theta'}\varphi^{(M)}_{\mathcal{X}} = \varphi^{(M)}_{\mathcal{X}}(\partial^2_{\theta\theta'}\mathcal{L}^{(M)} + \partial_\theta\mathcal{L}^{(M)}\partial_{\theta'}\mathcal{L}^{(M)}).$$

The difference $\partial^2_{\theta\theta'}\varphi^{(M)}_{\mathcal{X}} - \partial^2_{\theta\theta'}\varphi_{\mathcal{X}}$ splits into contributions involving $\partial^2_{\theta\theta'}\mathcal{L}^{(M)} - \partial^2_{\theta\theta'}\mathcal{L}$ (controlled by (2.28) with constant $C_2^\star$) and products of terms already bounded. Assembling these estimates yields the rate $O((1+M)^2\rho^{\alpha M})$. □

A.2.4. Proof of Proposition 2.2

Consistency. By Lemma 2.4($\iota\iota$), $\sup_{\theta\in\Theta}|D^{(M_n)}_{\mathcal{X}}(\theta) - D_{\mathcal{X}}(\theta)| = O(\rho^{\alpha M_n}) = o(1)$, so $D^{(M_n)}_{\mathcal{X}}$ converges uniformly to $D_{\mathcal{X}}$ on Θ. Together with the convergence in probability of φ_n to $\varphi_{\mathcal{X}}$ (established in Proposition 2.1) and identifiability (Assumption 5), this implies $\hat\theta^{(M_n)}_n \to \theta_0$ in probability.

Asymptotic normality. The choice (2.29) gives $\rho^{\alpha M_n} = n^{-\alpha c_\star \log(1/\rho)}$ and

$$(1+M_n)^2\rho^{\alpha M_n} = O\big((\log n)^2 n^{-\alpha c_\star \log(1/\rho)}\big) = o(n^{-1/2}),$$

since $\alpha c_\star \log(1/\rho) > 1/2$. A first-order Taylor expansion of the gradient condition $\nabla D^{(M_n)}_{\mathcal{X}}(\hat\theta^{(M_n)}_n) = 0$ around θ_0 gives

$$0 = \nabla D^{(M_n)}_{\mathcal{X}}(\theta_0) + \nabla^2 D^{(M_n)}_{\mathcal{X}}(\bar\theta_n)(\hat\theta^{(M_n)}_n - \theta_0),$$

where $\bar\theta_n$ lies on the segment joining θ_0 and $\hat\theta^{(M_n)}_n$. Lemma 2.4($\iota\iota$) yields

$$\begin{aligned}\nabla D^{(M_n)}_{\mathcal{X}}(\theta_0) &= \nabla D_{\mathcal{X}}(\theta_0) + R^{(1)}_n, \quad |R^{(1)}_n| = O((1+M_n)\rho^{\alpha M_n}) = o(n^{-1/2}),\\ \nabla^2 D^{(M_n)}_{\mathcal{X}}(\bar\theta_n) &= \nabla^2 D_{\mathcal{X}}(\bar\theta_n) + R^{(2)}_n, \quad \sup_{\theta\in\Theta}|R^{(2)}_n| = O((1+M_n)^2\rho^{\alpha M_n}) = o(1).\end{aligned}$$

Consistency and continuity ensure $\nabla^2 D_{\mathcal{X}}(\bar\theta_n) \to \Sigma(\theta_0)$ in probability. Substituting into the Taylor expansion and rearranging,

$$\sqrt{n}(\hat\theta^{(M_n)}_n - \theta_0) = -\Sigma(\theta_0)^{-1}\sqrt{n}\nabla D_{\mathcal{X}}(\theta_0) + o_P(1).$$

The conclusion follows from the central limit theorem for $\sqrt{n}\nabla D_{\mathcal{X}}(\theta_0)$ established via Theorem 2.1 of Knight and Yu (2002). □

A.3. Proof of Lemma 3.1

Denote $\boldsymbol{X}_{j,t} = (X_{j,t-m},\dots,X_{j,t},X_{j,t+1},\dots,X_{j,t+h})$ the paths of the moving averages $(X_{j,t})$, for $j=1,\dots,J$. 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 $j=1,\dots,J$. Because of the independence between $\boldsymbol{X}_{1,t},\dots,\boldsymbol{X}_{J,t}$, we have with $a=2/\pi$

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

Focusing on the shift vector, we have

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

with $\tilde{\boldsymbol{\mu}}_j = (\tilde{\mu}_{j,\ell})$ and $\tilde{\mu}_{j,\ell} = \int_{S_{m+h+1}} s_\ell\Gamma_j(d\boldsymbol{s})$, $\ell = -m,\dots,0,1,\dots,h$. 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

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

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,

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

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.4. 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 $(d_{j,k})_k$, $j = 1, \ldots, J$, with the same parameters m and h.

A.5. Proof of Proposition 3.1

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

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

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 $(d_{j,k})_k$, $j = 1, \ldots, J$. 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 $(\mathcal{X}_{t-m}, \ldots, \mathcal{X}_t, \mathcal{X}_{t+1}, \ldots, \mathcal{X}_{t+h})$ of size $m + h + 1$,

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

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:

$$\Gamma = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|_e^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|_e}\right\}} \tag{A.15}$$

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

To obtain the spectral representation on $C_{m+h+1}^{\|\cdot\|}$, 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 $C_{m+h+1}^{\|\cdot\|}$. 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 $C_{m+h+1}^{\|\cdot\|}$, we have $\Gamma(K^{\|\cdot\|}) = 0$. Define the projection mapping $T_{\|\cdot\|} : S_{m+h+1} \setminus K^{\|\cdot\|} \to C_{m+h+1}^{\|\cdot\|}$ by:

$$T_{\|\cdot\|}(\boldsymbol{s}) = \frac{\boldsymbol{s}}{\|\boldsymbol{s}\|} \tag{A.16}$$

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

$$\Gamma^{\|\cdot\|}(A) = \int_{T_{\|\cdot\|}^{-1}(A)} \|\boldsymbol{s}\|_e^{-\alpha} \Gamma(d\boldsymbol{s}) \tag{A.17}$$

for any Borel set $A \subset C_{m+h+1}^{\|\cdot\|}$. Since the original spectral measure Γ from (A.15) 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:

$$\Gamma^{\|\cdot\|}(A) = \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \sigma^\alpha \|\boldsymbol{d}_{j,k}\|_e^\alpha \cdot 1^{-\alpha} \cdot \mathbb{1}_A \left(\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|} \right) \tag{A.18}$$

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.15), we obtain:

$$\Gamma^{\|\cdot\|} = \sigma^\alpha \sum_{j=1}^{J} \sum_{\vartheta \in S_1} \sum_{k \in \mathbb{Z}} w_{j,\vartheta} \pi_j^\alpha \|\boldsymbol{d}_{j,k}\|^\alpha \delta_{\left\{\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}\right\}} \tag{A.19}$$

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

$$\boldsymbol{\mu}^{\|\cdot\|} = -\frac{2\sigma}{\pi} \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} \pi_j \beta_j \boldsymbol{d}_{j,k} \ln \|\sigma \pi_j \boldsymbol{d}_{j,k}\| \tag{A.20}$$

This completes the proof that the spectral representation $(\Gamma^{\|\cdot\|}, \boldsymbol{\mu}^{\|\cdot\|})$ of $\boldsymbol{X}_t$ on $C_{m+h+1}^{\|\cdot\|}$ 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.6. 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:

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

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:

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

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

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

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

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

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.7. 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 $\Gamma^{\|\cdot\|}\Big(\partial(A\cap B)\Big) = \Gamma^{\|\cdot\|}(\partial B) = 0$, 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 $\boldsymbol{X_t} = (\mathcal{X}_{t-m}, \ldots, \mathcal{X}_t, \mathcal{X}_{t+1}, \ldots, \mathcal{X}_{t+h})$ 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 $\boldsymbol{d}_{j,k} = (d_{j,k+m}, \ldots, d_{j,k}, d_{j,k-1}, \ldots, d_{j,k-h})$, $w_{j,\vartheta} = (1+\vartheta\beta_j)/2$, and if $\boldsymbol{d}_{j,k} = \mathbf{0}$, 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 $\frac{\vartheta \boldsymbol{d}_{j,k}}{\|\boldsymbol{d}_{j,k}\|}$ for $(\vartheta, j, k) \in S_1 \times \{1, \ldots, J\} \times \mathbb{Z}$. 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:

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

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

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

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

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

Therefore,

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

Moreover,

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

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

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

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

A.9. 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 ψ_j's are all positive. Letting $(\vartheta_0, j_0, k_0) \in \mathcal{I}$, consider*

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

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

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

For $m = 0$, then

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

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:

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

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:

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

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 $k_0 \in \{0, \ldots, h\}$

If $k' \in \{-m, \ldots, -1\}$, 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 $\psi_{j_0}^{k_0} \neq 0$. 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

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

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

$$\begin{aligned} f(\boldsymbol{d}_{j',k'}) &= (\psi_{j'}^{k'+m}, \ldots, \psi_{j'}^{k'+1}, \psi_{j'}^{k'}), \\ f(\boldsymbol{d}_{j_0,k_0}) &= (\psi_{j_0}^{k_0+m}, \ldots, \psi_{j_0}^{k_0+1}, \psi_{j_0}^{k_0}), \end{aligned}$$

and by (3.3) we also have that

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

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

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

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

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

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

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

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

and we also have that

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

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

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

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

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

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

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

Case $m = 0$

If $k_0 \in \{1, \ldots, h\}$ then $f(\boldsymbol{d}_{j_0,k_0}) = \psi_{j_0}^{k_0}$ and by (3.3), $\|\boldsymbol{d}_{j_0,k_0}\| = |\psi_{j_0}|^{k_0}$. 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

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

Proof. By Proposition 3.2,

$$\mathbb{P}_x^{\|\cdot\|}\Big(\boldsymbol{X_t}, A_{\vartheta,j,k} \Big| B(V_0)\Big) \underset{x\to\infty}{\longrightarrow} \frac{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in A_{\vartheta,j,k} : \ \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right)}{\Gamma^{\|\cdot\|}\left(\left\{\frac{\vartheta' \boldsymbol{d}_{j',k'}}{\|\boldsymbol{d}_{j',k'}\|} \in C_{m+h+1}^{\|\cdot\|} : \ \frac{\vartheta' f(\boldsymbol{d}_{j',k'})}{\|\boldsymbol{d}_{j',k'}\|} \in V_0\right\}\right)}. \tag{A.21}$$

Focusing on the denominator, we have by (3.15)

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

We will now distinguish the cases arising from the application of Lemma A.1. Recall that we assume for this proposition that the ψ_j's are positive. Thus, $\mathrm{sign}(\psi_j) = 1$ and $\bar{\beta}_j = \beta_j \dfrac{1-|\psi_j|^\alpha}{1-\psi_j^{<\alpha>}} = \beta_j$ 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,

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

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

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

Thus,

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

Similarly for the numerator in (A.21), by (3.16),

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

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

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

and

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

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

Case $m = 0$

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

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

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

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

$$= \sigma^{\alpha} w_{\vartheta_0} + \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \left[\sum_{k'=1}^{h-1} \|\boldsymbol{d}_{j',k'}\|^{\alpha} + \frac{\|\boldsymbol{d}_{j',h}\|^{\alpha}}{1-|\psi_{j'}|^{\alpha}} \right]$$

$$= \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \left[1 + \sum_{k'=1}^{h-1} |\psi_{j'}|^{\alpha k'} + \frac{|\psi_{j'}|^{\alpha h}}{1-|\psi_{j'}|^{\alpha}} \right]$$

$$= \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \left[\frac{1-|\psi_{j'}|^{\alpha h}}{1-|\psi_{j'}|^{\alpha}} + \frac{|\psi_{j'}|^{\alpha h}}{1-|\psi_{j'}|^{\alpha}} \right]$$

$$= \sigma^{\alpha} \sum_{j'=1}^{J} \pi_{j'}^{\alpha} w_{j',\vartheta_0} \frac{1}{1-|\psi_{j'}|^{\alpha}}.$$

Similarly, by (3.16),

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

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

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

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

The conclusion follows.