



HAL
open science

Asymptotic analysis of different covariance matrices estimation for minimum variance portfolio

Linda Chamakh, Emmanuel Gobet, Jean-Philippe Lemor

► **To cite this version:**

Linda Chamakh, Emmanuel Gobet, Jean-Philippe Lemor. Asymptotic analysis of different covariance matrices estimation for minimum variance portfolio. 2021. hal-03207061

HAL Id: hal-03207061

<https://hal.science/hal-03207061v1>

Preprint submitted on 23 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic analysis of different covariance matrices estimation for minimum variance portfolio*

Linda Chamakh[†]

Emmanuel Gobet[‡]

Jean-Philippe Lemor[§]

April 23, 2021

Abstract

In dynamic minimum variance portfolio, we study the impact of the sequence of covariance matrices taken in inputs, on the realized variance of the portfolio computed along a sample market path. The allocation of the portfolio is adjusted on a regular basis (every H days) using an updated covariance matrix estimator. In a modelling framework where the covariance matrix of the asset returns evolves as an ergodic process, we quantify the probability of observing an underperformance of the optimal dynamic covariance matrix compared to any other choice. The bounds depend on the tails of the returns, on the adjustment period H , and on the total number of rebalancing times N . These results provide asset managers with new insights into the optimality of their choice of covariance matrix estimators, depending on the depth of the backtest NH and the investment period H . Experiments based on GARCH modelling support our theoretical results.

KEYWORDS: covariance matrix estimation, portfolio theory, deviation inequality, ergodic process.

MSC2010: 62M10, 91G10, 37Axx, 62G05

1 Introduction

The mean-variance efficient portfolio theory by Markowitz [Mar59] has had a profound impact on modern finance. The Markowitz portfolio selection requires estimates of the covariance matrix and of the average expected returns. The covariance matrix can be either estimated in a non-parametric way, using sample-based empirical estimation, or in a parametric way, using factor models for example. In both cases, we may require sliding moving averages on the historical data.

Yet, the selected allocations are very sensitive to the values of the covariance matrix and expected returns used for its computation, and small changes in the inputs can lead to large changes in the allocations [Mic89].

*Acknowledgements. This research is supported by the *Chair Stress Test, RISK Management and Financial Steering of the Foundation Ecole Polytechnique* and by the *Association Nationale de la Recherche Technique*. This work is part of the first author's doctoral thesis (funded by BNP Paribas), under the supervision of the other authors.

[†]Global Markets Quantitative Research – BNP Paribas, France. Email: linda.chamakh@bnpparibas.com

[‡]CMAP, CNRS, École Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France. Email: emmanuel.gobet@polytechnique.edu

[§]Global Markets Quantitative Research – BNP Paribas, France. Email: jean-philippe.lemor@bnpparibas.com

On the other hand, it is well known that financial data exhibit heteroscedasticity, that is to say time dependent conditional covariance. This statistical property, referred to as stylized fact in the financial data setting, has been largely documented in the literature ([Con01], [EP07]). This implies that the expected covariance matrix in the near future can be very different from the average of the expected covariance matrix over a long time horizon. In this work, we aim to address the problem of the covariance matrix choice only.

In this work, we are interested in the forecast of the *realized covariance* for a specific time and period of investment. The minimum variance portfolio is the portfolio taking as input a covariance matrix and giving as output the allocation minimizing the associated variance. The key quantity is then the realized covariance over the period of investment. We consider a model for the returns of type

$$\mathbf{r}_{t|\mathcal{F}_{t-1}} \sim \mathcal{N}(0, V_t),$$

denoting t the time of investment and H the period of investment, the realized covariance corresponds to the future matrix $\frac{1}{H} \sum_{k=1}^H \mathbf{r}_{t+k} \mathbf{r}_{t+k}^\top$. Its best estimation at time t is the *conditional realized covariance*:

$$\frac{1}{H} \mathbb{E} \left[\sum_{k=1}^H \mathbf{r}_{t+k} \mathbf{r}_{t+k}^\top | \mathcal{F}_t \right] = \frac{1}{H} \sum_{k=1}^H \mathbb{E} [V_{t+k} | \mathcal{F}_t].$$

For $H = 1$, it coincides with the conditional covariance V_{t+1} ; when the period of investment exceeds the period of observation of the returns ($H > 1$), this quantity can still be estimated at time t . In practice, t and H are measured in days and $H = 21$ would correspond to an investment over a month (in business days) for example.

Usually, the asset manager might also consider a historical based covariance $\frac{1}{T} \sum_{k=1}^T \mathbf{r}_{t-k} \mathbf{r}_{t-k}^\top$ based on the past returns. When the backtest size T goes to infinity, this estimator converges to the *stationary covariance matrix* V_∞ .

The purpose of this work is to study the impact of the choice of the covariance matrix on the performance of the strategy.

In particular, we would like to give optimality guarantees under the form of concentration of measure inequalities of the outperformance of the portfolio based on the conditional realized covariance, versus any other covariance estimate V_{ref} . In the context of minimum variance investment, the natural portfolio metric is the realized variance, also called out-of-sample variance of the portfolio. Our result takes the form of a high probability event over the sums of the realized variance (RV) for N rebalancing dates of the portfolios:

$$\mathbb{P} \left(\sum_{n=1}^N \text{RV} \left(\frac{1}{H} \mathbb{E} \left[\sum_{k=1}^H \mathbf{r}_{t_n+k} \mathbf{r}_{t_n+k}^\top | \mathcal{F}_{t_n} \right] \right) \leq \sum_{n=1}^N \text{RV} (V_{\text{ref}}) \right) \approx 1 - c \left(\frac{H}{N} \right)^q$$

where $c > 0$ is a constant independent from H and N and q is a positive constant depending on the integrability of the process.

It means that the realized variance of the portfolio based on the conditional realized covariance will on the long term be lower than with other covariance with a high probability. The probability is even higher as the number N of times the portfolio is rebalanced grows, and decreases with the period of investment H , with a convergence rate increasing with the integrability of the process. The fact that the probability decreases with H is due to the fact that as H goes to infinity, the conditional realized covariance tends to the stationary covariance, which becomes optimal. But as

we may illustrate later, the half-life time associated to financial returns is often larger than the usual order of magnitude for H used by practitioners. Hence, considering non infinite H as we do is meaningful.

1.1 Literature Background

Our article falls within the line of the sensitivity analysis for Markowitz allocation articles.

In Guigues [Gui11] and Kan et al [KZ07], the authors give bounds on the portfolio risk as function of the variation of the input. In their approach, the metric is the mean-variance utility: $U(\mathbf{w}, \boldsymbol{\mu}, \Sigma) = t\mathbf{w}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{w}'\Sigma\mathbf{w}$. Using perturbation analysis results from Bonnans and Shapiro [BS00], Guigues shows that for two pairs of inputs $(\boldsymbol{\mu}_1, \Sigma_1)$ and $(\boldsymbol{\mu}_2, \Sigma_2)$, the optimal utilities difference can be bounded by the norm of the inputs differences:

$$|U(\mathbf{w}_2^*, \boldsymbol{\mu}_2, \Sigma_2) - U(\mathbf{w}_1^*, \boldsymbol{\mu}_1, \Sigma_1)| \leq \frac{1}{2}|\Sigma_2 - \Sigma_1|_\infty + k|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1|_\infty$$

where \mathbf{w}^* is the portfolio which maximizes the mean-variance utility in \mathbf{w} . In Kan and Zhou's article, the authors assume a multivariate Gaussian model for the returns and provide bounds on the expected difference between the utility of the mean variance portfolio with sample based estimates for $\boldsymbol{\mu}$ and Σ versus the optimal utility with the true parameters $\boldsymbol{\mu}$ and Σ (also called population parameters). The bound is linear in the number of assets and proportional to the inverse of the time of estimation.

Their bound holds on the expected value of the risk whereas in our approach, it is on the empirical risk, which is closer to practitioners needs.

A first attempt could be to use concentration of measures results for matrices (Tropp [Tro12], El Karoui [EK18]), and combine them with the aforementioned sensitivity results. But usually, this approach relies on independence properties between the matrices, which we don't have since V_i is a stochastic process, and on high order integrability (like sub-Gaussian tails) which we don't have since usually, asset returns have heavy tails.

The optimal horizon-adapted covariance matrix is similar to the one proposed in De Nard et al [DNLW18]. In this article, the authors compare empirically the performance of the minimum variance portfolio with different dynamic and static covariance estimation methods, including the GARCH-Dynamic Conditional Correlation (GARCH-DCC) models for returns or for residuals of a static factor models. Their experiments show that these two models outperform 10 other parametric estimation models in term of realized out-of-sample variance, which is in line with the theoretical results we have obtained in this work.

1.2 Contribution and outline of this work

In this work:

- We give guarantees of the optimality of the conditional realized covariance in the minimum variance portfolio setting in the form of a large deviation inequality, with a convergence rate polynomially decreasing in the number of rebalancing times of the portfolio;

- We give a scheme for estimating this covariance matrix in the specific GARCH-Constant Conditional Correlation (GARCH-CCC) model;
- We display results of numerical experiments to illustrate our statements.

Section 2 provides a presentation of the problem, introduces the notations and states the main result of this work. In Section 3, the proof of the main result is given. In Section 4, we introduce the GARCH-CCC model and verify that it satisfies the conditions of Section 2. Section 5 presents the result of the numerical experiments.

2 Formulation of the problem

2.1 Problem Setup

The model

Consider a pool of d assets, with daily centered returns processes $r_{1,t}, \dots, r_{d,t}$. Denote by \mathbf{r}_t the vector of the returns processes from day $t-1$ and t . Let $\{\boldsymbol{\eta}_t\}_{t \geq 0}$ i.i.d., $\boldsymbol{\eta}_t \sim \mathcal{N}(0, I_d)$ the *innovations* process and $\mathcal{F}_t = \sigma\{\boldsymbol{\eta}_s, 0 \leq s \leq t\}$ the associated filtration. We assume that \mathbf{r}_t is given by:

$$\mathbf{r}_t = V_t^{1/2} \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}(0, I_d), \quad t \geq 1, \quad (2.1)$$

where $V_t \in \mathbb{R}^{d \times d}$ is a \mathcal{F}_{t-1} measurable, positive definite matrix. It means that $\mathbf{r}_{t|\mathcal{F}_{t-1}} \sim \mathcal{N}(0, V_t)$. The initial condition V_1 is deterministic.

We assume that $\{V_t\}_{t > 1}$ is square integrable, and we denote $p_{\max} \geq 2$ s.t. $\mathbb{E}[|V_t|^{p_{\max}}]$ is finite for $t > 1$.

Covariance notation

A given portfolio with the allocation vector \mathbf{w} at time T_{reb} and a holding period of H has the realized variance:

$$\mathbf{w}^\top \left(\sum_{k=1}^H \mathbf{r}_{T_{reb}+k} \mathbf{r}_{T_{reb}+k}^\top \right) \mathbf{w} = \mathbf{w}^\top RC_{H, T_{reb}} \mathbf{w} \quad (2.2)$$

with $RC_{H, T_{reb}}$ denoting the *realized covariance* over the period of investment. We seek for the portfolio which minimizes this realized variance.

At the time of the investment, the best estimation of $RC_{H, T_{reb}}$ is:

$$cRC_{H, T_{reb}} := \mathbb{E}[RC_{H, T_{reb}} | \mathcal{F}_{T_{reb}}] = \sum_{k=1}^H \mathbb{E}[V_{T_{reb}+k} | \mathcal{F}_{T_{reb}}]$$

which we call the *conditional realized covariance*. Given a covariance matrix C which we assume to be definite positive, we will consider the following risk optimization under constraints

$$\text{mv}(C) := \arg \min_{\mathbf{w} \in \mathcal{W}} \mathbf{w}^\top C \mathbf{w},$$



Figure 1: Illustration of quantities time-dependence in our setting: since r_t and V_{t+1} are \mathcal{F}_t measurable, they are the known quantities at each time t . The red dots dates correspond to the rebalancing times.

where \mathcal{W} is the set of constraints containing at least the budget constraint and a maximal allocation constraint: $\{\mathbf{w} = \{w_i\}_{i \in \{1, \dots, d\}} \in \mathbb{R}^d : \sum_{i=1}^d w_i = 1, \text{ and } |w_i| \leq c_w, i \in \{1, \dots, d\}\} \subset \mathcal{W}$, $c_w > 0$. We assume that \mathcal{W} is closed and convex, which ensures the existence and uniqueness of $\text{mv}(C)$. We aim at showing that $\text{mv}(cRC_{H, T_{reb}})$ is a good $\mathcal{F}_{T_{reb}}$ -measurable allocation for minimizing the realized variance (2.2).

We will consider multiple rebalancing times of the portfolio: $T_{reb} = t_1, \dots, t_N$, $t_{n+1} - t_n = H$, $t_0 = 0$.

Let us introduce the following processes:

$$\begin{aligned}
 R_{N,H} &:= \sum_{n=1}^N \text{mv}(cRC_{H,t_n})^\top RC_{H,t_n} \text{mv}(cRC_{H,t_n}), \\
 cRV_{N,H} &:= \sum_{n=1}^N \text{mv}(cRC_{H,t_n})^\top cRC_{H,t_n} \text{mv}(cRC_{H,t_n}), \\
 R_{N,H}^{\text{ref}} &:= \sum_{n=1}^N \text{mv}(V_{\text{ref}})^\top RC_{H,t_n} \text{mv}(V_{\text{ref}}), \\
 cRV_{N,H}^{\text{ref}} &:= \sum_{n=1}^N \text{mv}(V_{\text{ref}})^\top cRC_{H,t_n} \text{mv}(V_{\text{ref}}),
 \end{aligned} \tag{2.3}$$

with V_{ref} is a deterministic, positive definite covariance matrix which we take as the benchmark covariance the asset manager considers for his optimization.

- $R_{N,H}$ (resp. $R_{N,H}^{\text{ref}}$) is the realized variance of the cRC_{H,t_n} -based portfolios (resp. the V_{ref} -based portfolio);
- $cRV_{N,H}$ (resp. $cRV_{N,H}^{\text{ref}}$) denotes the sum of conditional realized variances of the cRC_{H,t_n} -based portfolios (resp. the V_{ref} -based based portfolio).

For $p_{\max} \geq 2$, our main result Theorem 2.2 takes the form:

$$\mathbb{P}\left(R_{N,H} < R_{N,H}^{\text{ref}}\right) \geq 1 - C \left(\frac{H}{N}\right)^{\frac{p_{\max}}{2}}.$$

This inequality is of the form of the probability of the realized variance being lower (e.g. better) for the conditional realized covariance-based portfolio than for the reference covariance-based portfolio.

The probability bound is polynomially decreasing in N , the number of rebalancing times of the portfolio and polynomially increasing in H , the investment horizon, with an exponent $\frac{p_{\max}}{2}$ equal

to a quarter of the maximum finite moment of the returns. This relatively slow convergence (polynomial rather than exponential) stems from the low integrability on the process \mathbf{r}_t .

The main message of this result is that, on the long-run (N going to infinity), the conditional realized covariance-based portfolio is outperforming with high probability. For very large H though, this bound might become loose, since the infinite-horizon estimate of $\frac{RC_{H,T_{reb}}}{H}$ coincides with the stationary covariance V_∞ .

2.2 Model assumptions

We recall that we are under the model (2.1) for the returns. We are now going to specify the assumptions on the conditional covariance matrix V_t .

We denote $\mathcal{S} \subset \mathcal{S}_+^d$ the state space on which V_t takes its values, where \mathcal{S}_+^d is the set of symmetric positive definite matrices.

H_{stat}. $\{V_t\}_{t \in \mathbb{N}^*}$ possesses a strictly stationary and ergodic distribution μ with at least L_2 moment and $V_\infty := \int_{\mathcal{S}} v \mu(dv)$.

H_S. $\{V_t\}_{t \in \mathbb{N}}$ is a time homogeneous, aperiodic Lebesgue-irreducible Markov chain¹ on the state-space \mathcal{S} .

H_L. There exist some constant $\delta \in (0, 1)$, $b \in \mathbb{R}$ and a measurable function $\mathcal{L} : \mathcal{S} \rightarrow [1, +\infty)$ s.t. $\lim_{|x| \rightarrow \infty} \mathcal{L}(x) = +\infty$, and an accessible small set $\mathcal{C} \subset \mathcal{B}(\mathcal{S})$, such that for all $V_1 \in \mathcal{S}$,

$$\mathbb{E} [\mathcal{L}(V_2) | V_1] \leq \delta \mathcal{L}(V_1) + b \mathbf{1}_{\mathcal{C}}(V_1).$$

H_{p_{max}}. Under **H_L**, the growth of \mathcal{L} at infinity is polynomial of order p_{\max} : $\exists c_{\mathcal{L}}, C_{\mathcal{L}} > 0$, $c_{\mathcal{L}} |x|^{p_{\max}} \leq \mathcal{L}(x) \leq C_{\mathcal{L}}(1 + |x|^{p_{\max}})$.

H_x. $\{\mathbf{x}_t\}_{t \in \mathbb{N}^*}$ denotes a portfolio allocation process, e.g. a \mathcal{F}_t -measurable process with values in \mathcal{W} , hence bounded: $\mathbb{P}(|x_{i,t}| > c_{\mathbf{w}}) = 0$ for $i \in \{1, \dots, d\}$.

As will be shown later in the proof, these assumptions are compatible with the existence of moment $p_{\max} < \infty$.

2.3 Main results

We will call *performance gap* the quantity ℓ_H defined by the expected value of $\frac{cRV_{N,H}^{\text{ref}} - cRV_{N,H}}{N}$ under the stationary law:

$$\ell_H := \mathbb{E} \left[\frac{cRV_{N,H}^{\text{ref}} - cRV_{N,H}}{N} \middle| V_1 \sim \mu \right].$$

It is a key quantity since it can be interpreted as a performance gap between the benchmark and the estimated conditional realized covariance portfolio.

Let us first state a result on the sign of ℓ_H .

¹We refer the reader to paragraph 3.2.1 for reminders on Markov chains elements of language.

Proposition 2.1 (Non-negativity of the performance gap). *Assume H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$. Then, ℓ_H is finite, deterministic, and non-negative.*

Our main result is the following:

Theorem 2.2. *Assume that $\{V_t, t \in \mathbb{N}^*\}$ satisfies assumptions H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$. Assume that ℓ_H is strictly positive, then for any $N, H \in \mathbb{N}^*$, the processes $R_{N,H}, R_{N,H}^{ref}$ defined in (2.3) satisfy:*

$$\mathbb{P}\left(R_{N,H} > R_{N,H}^{ref}\right) \leq \frac{C}{\ell_H^{p_{\max}}} \frac{H^{\frac{p_{\max}}{2}}}{N^{\frac{p_{\max}}{2}}},$$

where $C > 0$ depends on H , p_{\max} , d and \mathcal{L} .

Comments:

- This inequality is of the form of an upper bound on the probability that the realized variance is higher (e.g. worse) for the conditional covariance-based portfolio than for the reference covariance-based portfolio. It is a probability bound on the *underperformance* of the estimated covariance based-portfolio versus the reference portfolio.
- It is polynomially decreasing in N , the number of rebalancing times of the portfolio, and in ℓ_H , the performance gap, and at least polynomially increasing in H , the investment horizon. As we show in the proof of Theorem 2.2, the constant C is linear in a quantity $C_{FM}^{(H)}$ on which we provide a bound in Proposition 3.6 which has an exponential growth in H and p_{\max} .

Interpretations:

- As mentioned before, ℓ_H can be interpreted as a performance gap between the benchmark and the estimated conditional realized covariance portfolio. The higher the ℓ_H , the more discriminant the impact of using the estimated realized covariance.
- When $CH/(N\ell_H^2)$ is large, the bound is uninformative. If $N \ll CH/\ell_H^2$, there is not enough observations to statistically distinguish which covariance matrix brings the best performance. As we show in Lemma 2.3, the average performance gap ℓ_H/H , when using the estimated realized covariance versus the stationary covariance V_∞ , goes to zero when H goes to infinity. It means that in the $(NH) \ll C(H/\ell_H)^2$ regime, the asset manager should not bother much using a sophisticated estimation for V_t : a good approximation of V_∞ is enough.
- When N tends to infinity, the probability goes to zero: this is a concentration of measure effect, and since the expected realized variance difference ℓ_H is non-negative (see Proposition 2.1,) the probability of having a negative empirical difference goes quickly towards zero. It is coherent with the intuition than with more data, the historical measure of the realized variance difference will be more likely to be of the same sign than its expected value.

When H goes to infinity, the average conditional realized covariance converges towards the mean-value of the process: $\frac{cRC_{H,T_{reb}}}{H} \rightarrow V_\infty$. If $V_{ref} = V_\infty$, we expect that $\lim_{H \rightarrow \infty} \frac{\ell_H}{H} = 0$. This is what is stated in the next Lemma.

Lemma 2.3 (Convergence of $\frac{\ell_H^\infty}{H}$ to zero). *Assume that $\{V_t, t \in \mathbb{N}^*\}$ satisfies assumptions H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$, and let $V_{ref} = V_\infty$. In that case, denote ℓ_H^∞ the performance gap. Then*

$$\lim_{H \rightarrow \infty} \frac{\ell_H^\infty}{H} = 0.$$

3 Proofs and auxiliary results

3.1 Proof of Theorem 2.2

To show this concentration of measures result, we will need the following auxiliary results (whose proofs are postponed to Subsections 3.2, 3.3 and 3.4).

Let us first give a hint of the result by rewriting the realized variance difference $R_{N,H}^{ref} - R_{N,H}$. From (2.3), the realized variance difference breaks down into:

$$R_{N,H}^{ref} - R_{N,H} = \underbrace{R_{N,H}^{ref} - cRV_{N,H}^{ref}}_{D_{N,H}^{ref}} - \underbrace{(R_{N,H} - cRV_{N,H})}_{D_{N,H}} + \underbrace{cRV_{N,H}^{ref} - cRV_{N,H}}_{\mathcal{E}_{N,H}}. \quad (3.1)$$

From the processes definition, it is easy to see that:

- by the definition of $mv(\cdot)$: $(mv(V_{ref})^\top cRC_{H,t_n} mv(V_{ref})) \geq (mv(cRC_{H,t_n})^\top cRC_{H,t_n} mv(cRC_{H,t_n}))$ so $\mathcal{E}_{N,H} = cRV_{N,H}^{ref} - cRV_{N,H} \geq 0$ almost surely,
- by the definition of cRC_{H,t_n} , $\mathbb{E}[RC_{H,t_n} | \mathcal{F}_{t_n}] = cRC_{H,t_n}$, so $D_{N,H}^{ref} = R_{N,H}^{ref} - cRV_{N,H}^{ref}$ and $D_{N,H} = R_{N,H} - cRV_{N,H}$ are \mathcal{F}_{t_n} -centered martingales which concentrate around zero.

The following result gives a concentration of measure result on $\mathcal{E}_{N,H}$:

Proposition 3.1 (Concentration of measure for the ergodic conditional realized variance). *Assume that $\{V_t, t \in \mathbb{N}^*\}$ satisfies assumptions H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$. Then, for any $2 \leq q \leq p_{\max}$, there exists $C_{q,\mathcal{L},H,d} > 0$ such that:*

$$\mathbb{E}[|\mathcal{E}_{N,H} - N\ell_H|^q] \leq C_{q,\mathcal{L},H,d}(NH)^{q/2}.$$

A direct application via the Markov inequality gives, for any $a > 0$:

$$\mathbb{P}\left(\frac{\mathcal{E}_{N,H}}{N} - \ell_H > a\right) \leq C_{q,\mathcal{L},H,d} \left(\frac{H}{Na^2}\right)^{\frac{q}{2}}$$

and

$$\mathbb{P}\left(\frac{\mathcal{E}_{N,H}}{N} - \ell_H < -a\right) \leq C_{q,\mathcal{L},H,d} \left(\frac{H}{Na^2}\right)^{\frac{q}{2}}. \quad (3.2)$$

We can also show that ℓ_H is non-negative and finite, and as a consequence of the previous Proposition, $\frac{\mathcal{E}_{N,H}}{N}$ can be shown to converge towards ℓ_H :

Proposition 3.2 (Convergence of the average conditional realized variance). *Assume H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$. Then, for $\ell_H = \mathbb{E} \left[\frac{cRV_{N,H}^{ref} - cRV_{N,H}}{N} \middle| V_1 \sim \mu \right]$,*

- $\frac{cRV_{N,H}^{ref} - cRV_{N,H}}{N}$ converges in $L_{p_{\max}}$ norm towards ℓ_H :

$$\frac{cRV_{N,H}^{ref} - cRV_{N,H}}{N} \xrightarrow[N \rightarrow \infty]{L_{p_{\max}}} \ell_H.$$

- if $p_{\max} > 2$, $\frac{cRV_{N,H}^{ref} - cRV_{N,H}}{N}$ converges almost surely towards ℓ_H :

$$\frac{cRV_{N,H}^{ref} - cRV_{N,H}}{N} \xrightarrow[N \rightarrow \infty]{a.s.} \ell_H.$$

We now state the concentration of measure results for the martingale processes of type $D_{N,H}$ and $D_{N,H}^{ref}$.

Lemma 3.3. *Assume that $\{V_t, t \in \mathbb{N}^*\}$ satisfies (2.1), H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$. Let $\{\mathbf{x}_{t_n}\}_{n \in \mathbb{N}^*}$, $t_{n+1} - t_n = H$ be a portfolio allocation satisfying H_x and $D_{N,H}(\mathbf{x}) := \sum_{n=1}^N \mathbf{x}_{t_n}^\top (RC_{H,t_n} - cRC_{H,t_n}) \mathbf{x}_{t_n}$. Then, there exists $C_{p_{\max},d,\mathcal{L}} > 0$ such that:*

$$\mathbb{E} [|D_{N,H}(\mathbf{x})|^{p_{\max}}] \leq C_{p_{\max},d,\mathcal{L}} (NH)^{\frac{p_{\max}}{2}}.$$

A direct application of the Markov inequality gives, for any $a > 0$:

$$\mathbb{P} \left(\frac{D_{N,H}(\mathbf{x})}{N} > a \right) \leq C_{p_{\max},d,\mathcal{L}} \left(\frac{H}{Na^2} \right)^{\frac{p_{\max}}{2}}, \quad (3.3)$$

and

$$\mathbb{P} \left(\frac{D_{N,H}(\mathbf{x})}{N} < -a \right) \leq C_{p_{\max},d,\mathcal{L}} \left(\frac{H}{Na^2} \right)^{\frac{p_{\max}}{2}}. \quad (3.4)$$

We can now move to the proof of our main result.

Proof of Theorem 2.2. We aim at showing that $R_{N,H} < R_{N,H}^{ref}$ with high probability.

From equation (3.1), we see that the realized variance difference $R_{N,H}^{ref} - R_{N,H}$ boils down to the sum of the martingales $D_{N,H}^{ref} = R_{N,H}^{ref} - cRV_{N,H}^{ref}$ and $-D_{N,H} = cRV_{N,H} - R_{N,H}$ and and of the ergodic term $\mathcal{E}_{N,H} = cRV_{N,H}^{ref} - cRV_{N,H}$:

$$\begin{aligned} R_{N,H}^{ref} - R_{N,H} &= D_{N,H}^{ref} - D_{N,H} + \mathcal{E}_{N,H} \\ &= N \left(\frac{D_{N,H}^{ref}}{N} - \frac{D_{N,H}}{N} + \left(\frac{\mathcal{E}_{N,H}}{N} - \ell_H \right) + \ell_H \right). \end{aligned}$$

From this equality, we see that if each of the three first terms is higher than $-\frac{\ell_H}{3}$, then the sum plus ℓ_H is non-negative:

$$\left. \begin{aligned} \frac{D_{N,H}^{\text{ref}}}{N} &\geq -\frac{\ell_H}{3} \\ -\frac{D_{N,H}}{N} &\geq -\frac{\ell_H}{3} \\ \frac{\mathcal{E}_{N,H}}{N} - \ell_H &\geq -\frac{\ell_H}{3}, \end{aligned} \right\} \Rightarrow R_{N,H}^{\text{ref}} - R_{N,H} \geq 0.$$

This translates into the following inclusion of events:

$$\left\{ \left\{ \frac{D_{N,H}}{N} \leq \frac{\ell_H}{3} \right\} \cap \left\{ -\frac{D_{N,H}^{\text{ref}}}{N} \leq \frac{\ell_H}{3} \right\} \cap \left\{ -\left(\frac{\mathcal{E}_{N,H}}{N} - \ell_H \right) \leq \frac{\ell_H}{3} \right\} \right\} \subset \left\{ R_{N,H} \leq R_{N,H}^{\text{ref}} \right\}.$$

Taking the complementary, we see that the event $\left\{ R_{N,H} > R_{N,H}^{\text{ref}} \right\}$ is included in an union of low probability events:

$$\left\{ R_{N,H} > R_{N,H}^{\text{ref}} \right\} \subset \left\{ \left\{ \frac{D_{N,H}}{N} > \frac{\ell_H}{3} \right\} \cup \left\{ \frac{D_{N,H}^{\text{ref}}}{N} < -\frac{\ell_H}{3} \right\} \cup \left\{ \left(\frac{\mathcal{E}_{N,H}}{N} - \ell_H \right) < -\frac{\ell_H}{3} \right\} \right\}.$$

Hence we can bound the probability of $\left\{ R_{N,H} > R_{N,H}^{\text{ref}} \right\}$ by the sum of the probability of the three events, on which we know explicit bounds via Lemma 3.3 and Proposition 3.1.

Bound on the long horizon martingale term

By Lemma 3.3, taking $a = \frac{\ell_H}{3}$ (which is positive by assumption) in equation (3.3), with $\mathbf{x}_{t_n} = \text{mv}(cRC_{H,t_n})$, we have:

$$\mathbb{P} \left(\frac{D_{N,H}}{N} > \frac{\ell_H}{3} \right) \leq C_{p_{\max},d,\mathcal{L}} \left(\frac{9H}{N\ell_H^2} \right)^{\frac{p_{\max}}{2}}.$$

From equation (3.4), with $\mathbf{x}_{t_n} = \text{mv}(V_{\text{ref}})$,

$$\mathbb{P} \left(\frac{D_{N,H}^{\text{ref}}}{N} < -\frac{\ell_H}{3} \right) \leq C_{p_{\max},d,\mathcal{L}} \left(\frac{9H}{N\ell_H^2} \right)^{\frac{p_{\max}}{2}}.$$

Bound on the long horizon ergodic term

From Proposition 3.1, replacing a by $\frac{\ell_H}{3}$ in equation (3.2) we have:

$$\mathbb{P} \left(\frac{\mathcal{E}_{N,H}}{N} - \ell_H < -\frac{\ell_H}{3} \right) \leq C_{p_{\max},\mathcal{L},H,d} \left(\frac{9H}{N\ell_H^2} \right)^{\frac{p_{\max}}{2}}.$$

By union bound, we conclude:

$$\mathbb{P} \left(R_{N,H} > R_{N,H}^{\text{ref}} \right) \leq 3^{p_{\max}} (2C_{p_{\max},d,\mathcal{L}} + C_{p_{\max},\mathcal{L},H,d}) \left(\frac{H}{N\ell_H^2} \right)^{\frac{p_{\max}}{2}}.$$

□

3.2 Proof of Proposition 3.1

The non-trivial part of our Proposition 3.1 lies in the fact that we want to highlight both the N and H dependence. From the conditional realized variance definition, there is a nested concentration effect, both in N , the number of rebalancing times, and H , the number of days on which the variance is measured, that we can exploit.

3.2.1 Preparatory results

In this paragraph, we state the concentration of measure result for ergodic Markov process that we will adapt to show our Proposition. We recall also some Markov chain elements of vocabulary.

Concentration of measure for irreducible aperiodic Markov chain

We recall here the concentration of measure result for irreducible aperiodic Markov chain as stated in Fort and Moulines article [FM03, Proposition 2].

Proposition 3.4 ([FM03][Proposition 2]). *Let $\{V_t\}_{t \in \mathbb{N}^*}$ be a ϕ -irreducible aperiodic Markov chain on \mathcal{S} , and let $\mathcal{C} \in \mathcal{B}(\mathcal{S})$ be an accessible petite set. Assume that there exist some constants $\delta \in (0, 1)$, $b < \infty$ and a measurable $\mathcal{L} : \mathcal{S} \rightarrow [1, +\infty)$, bounded on \mathcal{C} , such that*

$$\mathbb{E}[\mathcal{L}(V_2)|V_1] \leq \delta \mathcal{L}(V_1) + b\mathbf{1}_{\mathcal{C}}(V_1), \quad \forall V_1 \in \mathcal{S}. \quad (3.5)$$

Let $q \geq 2$. Choose $M > \sup_{\mathcal{C}} \mathcal{L} \vee b / (1 - \delta^{1/q})^q$. Then the set $\{\mathcal{L} \leq M\}$ is ν_m -small with minorizing constant $\epsilon > 0$ and for any Borel function $g : \mathcal{S} \rightarrow \mathbb{R}^d$, $|g| \leq \mathcal{L}^{\frac{1}{q}}$, it holds that for all $V_1 \in \mathcal{S}$, $H \geq 1$,

$$\mathbb{E} \left[\left| \sum_{k=1}^H (g(V_{k+1}) - \mu(g)) \right|^q \middle| V_1 \right] \leq C_{FM} \mathcal{L}(V_1) H^{q/2},$$

where $C_{FM} = C \left(\frac{m+1}{\epsilon} \right)^{q+1} \frac{M^2}{A^{2q}}$, $A = (1 - \delta)^{1/q} - (b/M)^{1/q}$ and C is a constant which depends only on q , and μ is the invariant measure associated to $\{V_t\}_{t \in \mathbb{N}^*}$.

Remark 3.1. *A few remarks on this Proposition and how we will apply it:*

- For simplicity, we have denoted C_{FM} the constant $C \left(\frac{m+1}{\epsilon} \right)^{q+1} \frac{M^2}{A^{2q}}$. This constant depends on the Lyapunov condition (3.5) (so on \mathcal{L}, δ, b), on the petite set \mathcal{C} , on q and on ν_m and ϵ and on the dynamic of $\{V_t\}_{t \in \mathbb{N}}$. It is hard to quantify it explicitly.
- The proposition is stated for functions g such that $|g| \leq \mathcal{L}^{\frac{1}{q}}$. It can be extended to functions which are bounded in $\mathcal{L}^{1/q}$ norm, e.g. such that: $\left(\sup_{V \in \mathcal{S}} \frac{|g(V)|^q}{\mathcal{L}(V)} \right)^{1/q} < \infty$. Indeed, denoting $\tilde{g} = g / \left(\sup_{V \in \mathcal{S}} \frac{|g(V)|^q}{\mathcal{L}(V)} \right)^{1/q}$, then for any $V' \in \mathcal{S}$,

$$\frac{|\tilde{g}(V')|^q}{\mathcal{L}(V')} = \frac{|g(V')|^q / \mathcal{L}(V')}{\sup_{V \in \mathcal{S}} |g(V)|^q / \mathcal{L}(V)} \leq 1$$

so $|\tilde{g}| \leq \mathcal{L}^{\frac{1}{q}}$. We can apply the proposition on \tilde{g} and express it in terms of g :

$$\mathbb{E} \left[\left| \sum_{k=1}^H (g(V_{k+1}) - \mu(g)) \right|^q \middle| V_1 \right] \leq C_{FM} \left(\sup_{V \in \mathcal{S}} \frac{|g(V)|^q}{\mathcal{L}(V)} \right) \mathcal{L}(V_1) H^{q/2}.$$

Markov chain: elements of vocabulary

We recall the definition of *time of first return*, *irreducibility*, *petite set*, *small set*, *aperiodicity* and *accessibility*, which can be found in the Meyn and Tweedie's Book [MT09, pages 71, 82, 117, 102, 114 and 86]:

Let $\{X_n, n \in \mathbb{N}\}$ be a Markov chain on a state space \mathcal{S} , P its transition probability. Let $A \subset \mathcal{B}(\mathcal{S})$. We denote $\tau_A := \min\{n \geq 1 : X_n \in A\}$ the *first return* time on A and for $x \in \mathcal{S}$, $L(x, A) := \mathbb{P}(\tau_A < \infty | X_0 = x)$ the probability to access A from a specific x . Given ϕ a Borelian measure, we say that X_n is ϕ -irreducible if $\phi(A) > 0 \Rightarrow L(y, A) > 0$ for any $y \in \mathcal{S}$.

A *petite set* is a set $\mathcal{C} \in \mathcal{B}(\mathcal{S})$ such that there is a probability distribution a on \mathbb{N} and a non-trivial measure ν_a such that $\forall x \in \mathcal{C}, \forall A \subset \mathcal{B}(\mathcal{S}), \sum_{n=0}^{\infty} a(n)P^n(x, A) \geq \nu_a(A)$.

A *small set* is a particular case of a petite set in which a only charges a specific $m \in \mathbb{N}$: the definition becomes: $P^m(x, A) \geq \nu_m(A)$. When there exists a small set with $m = 1$ and $\nu_1(\mathcal{C}) > 0$, then the chain is called *strongly aperiodic*.

A set $A \subset \mathcal{B}(\mathcal{S})$ is said *accessible* if it can be accessed from any $x \in \mathcal{S}$: $L(x, A) > 0, \forall x \in \mathcal{S}$.

Concentration of measure for function of lagged Markov chain: motivation

Notice that since $\{V_{t+1}, t \in \mathbb{N}^*\}$ is Markovian, the random variable $cRC_{H,t} = \mathbb{E}\left[\sum_{k=1}^H V_{t+k} | \mathcal{F}_t\right]$ can be seen as a function of the \mathcal{F}_t -measurable V_{t+1} : $cRC_{H,t} = \phi(V_{t+1})$. Hence $\mathcal{E}_{N,H}$ can be identically written:

$$\begin{aligned} \mathcal{E}_{N,H} &= \sum_{n=1}^N \text{mv}(V_{\text{ref}})^\top cRC_{H,t_n} \text{mv}(V_{\text{ref}}) - \text{mv}(cRC_{H,t_n})^\top cRC_{H,t_n} \text{mv}(cRC_{H,t_n}) \\ &= \sum_{n=1}^N [\text{mv}(V_{\text{ref}}) - \text{mv}(cRC_{H,t_n})]^\top cRC_{H,t_n} [\text{mv}(V_{\text{ref}}) + \text{mv}(cRC_{H,t_n})] \\ &= \sum_{n=1}^N \mathbf{x}_{t_n}^\top \phi(V_{t_n+1}) \mathbf{y}_{t_n} \end{aligned} \quad (3.6)$$

where

$$\begin{aligned} \mathbf{x}_{t_n} &= \text{mv}(V_{\text{ref}}) - \text{mv}(cRC_{H,t_n}), \\ \mathbf{y}_{t_n} &= \text{mv}(V_{\text{ref}}) + \text{mv}(cRC_{H,t_n}). \end{aligned} \quad (3.7)$$

In the following result, we show that the concentration of measure result Proposition 3.4 remains valid for function of the lagged chain: $g(V_k) \rightarrow g(V_{kH})$.

Proposition 3.5. [Fort Moulines proposition extension to lagged-Markov chains]

Assume that $\{V_{t+1}, t \in \mathbb{N}^*\}$ satisfies assumptions H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$. Then for any $q \geq 2$, for any Borel function $g : \mathcal{S} \rightarrow \mathbb{R}^{d \times d}$ bounded in $\mathcal{L}^{1/q}$ -norm, $T \in \mathbb{N}^*$, $H \in \mathbb{N}^*$, we have:

$$\mathbb{E} \left[\left| \sum_{t=1}^T (g(V_{tH+1}) - \mu(g)) \right|^q \middle| V_1 \right] \leq C_{FM}^{(H)} \left(\sup_{\mathcal{S}} \frac{|g|^q}{\mathcal{L}} \right) \mathcal{L}(V_1) T^{q/2}, \quad (3.8)$$

where $C_{FM}^{(H)}$ is a constant on which we provide a bound in H and q in Proposition 3.6.

Proof. Let $H \in \mathbb{N}^*$. First, let us show that the lagged Markov chain $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$ satisfies the assumptions of Proposition 3.4.

1. Irreducibility and aperiodicity of the lagged Markov chain

$\{V_{tH+1}\}_{t \in \mathbb{N}^*}$ is a Lebesgue-irreducible aperiodic Markov chain on \mathcal{S} by application of Proposition A.8 ([MT09, Proposition 5.4.5]: extension of irreducibility and aperiodicity of irreducible and aperiod chains) to the chain $\{V_{t+1}\}_{t \in \mathbb{N}^*}$ which is Lebesgue-irreducible and aperiodic by assumption \mathbf{H}_{stat} .

2. Lyapunov condition on the lagged Markov chain

In Meyn and Tweedie's book, Theorem A.6 ([MT09, Theorem 15.3.4]) states that if $\{V_{t+1}\}_{t \in \mathbb{N}^*}$ satisfies a drift condition $\mathbf{H}_{\mathcal{L}}$ with a Lyapunov function \mathcal{L} and a petite set \mathcal{C} , then $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$ also satisfies a drift condition with the same Lyapunov function \mathcal{L} and some set $\mathcal{C}^{(H)}$ which is petite for the H -skeleton.

In our Proposition A.7, we give a quantitative assessment of the Lyapunov condition for $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$ by giving explicitly the constants d_H and b_H s.t.:

$$\mathbb{E}[\mathcal{L}(V_{H+1})|V_1] \leq d_H \mathcal{L}(V_1) + b_H \mathbf{1}_{\mathcal{C}^{(H)}}(V_1).$$

Our computation yields:

$$d_H = \frac{1 + \delta^H}{2}, \quad \mathcal{C}^{(H)} = \mathcal{C} \cup \{x \in \mathcal{S} \mid |x| \leq R^{(H)}\},$$

$$b_H = \sup_{x \in \mathcal{C}^{(H)}} \left(b \frac{1 - \delta^H}{1 - \delta} - \frac{1 - \delta^H}{2} \mathcal{L}(x) \right),$$

and $R^{(H)} > 0$ s.t. for every $x \in \mathcal{S}$ s.t. $|x| > R^{(H)}$, $\frac{1 - \delta^H}{2} \mathcal{L}(x) \geq b \frac{1 - \delta^H}{1 - \delta}$.

3. Smallness of the set $\mathcal{C}^{(H)}$

By irreducibility and aperiodicity, petite sets are also small sets [MT09, Theorem 5.5.7] Since \mathcal{C} is included in the new one $\mathcal{C}^{(H)}$ (whether in Meyn and Tweedie's book [MT09, Lemma 14.2.8] and in our Proposition A.7), it will be enough to show that \mathcal{C} is a small set for $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$. A key assumption to prove this result will be that we can consider a measure associated to \mathcal{C} for $\{V_{t+1}\}_{t \in \mathbb{N}^*}$ satisfying $\nu(\mathcal{C}) > 0$.

Indeed, let us denote ν_m and $\epsilon > 0$ the measure and minorizing constant s.t. $\nu_m(\mathcal{C}) > 0$ (w.l.o.g. we can assume that $\nu_m(\mathcal{C}) = 1$) and $P^m(x, A) \geq \epsilon \nu_m(A)$ for any $x \in \mathcal{C}$ and $A \subset \mathcal{B}(\mathcal{S})$. They exist by application of [MT09, Proposition 5.2.4 - (iii)] (existence of a measure positive on the small set for irreducible chains).

Let us show that \mathcal{C} is a small set for $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$. Indeed, we can show recursively that $P^{mk}(x, A) \geq \epsilon^k \nu_m(A)$ for any $k \in \mathbb{N}^*$. It is satisfied for $k = 1$. Let us assume the property for $k \in \mathbb{N}^*$. Then for $k + 1$:

$$\begin{aligned} P^{(k+1)m}(x, A) &= \int_{\mathcal{S}} P^{mk}(x, dy) P^m(y, A) \geq \epsilon^k \int_{\mathcal{S}} \nu_m(dy) P^m(y, A) \\ &\geq \epsilon^k \int_{\mathcal{C}} \nu_m(dy) \underbrace{P^m(y, A)}_{\substack{\geq \epsilon \nu_m(A) \\ \text{since } y \in \mathcal{C}}} \\ &\geq \epsilon^{k+1} \underbrace{\nu_m(\mathcal{C})}_{=1} \nu_m(A) = \epsilon^{k+1} \nu_m(A). \end{aligned} \tag{3.9}$$

So for $k = H$, \mathcal{C} is a small set for $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$. To show that $\mathcal{C}^{(H)}$ is a small set for $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$, it suffices to take $\nu_m^{(H)} = \nu_m$ on \mathcal{C} and $\nu_m^{(H)} = 0$ on $\mathcal{C}^{(H)} \setminus \mathcal{C}$.

4. Accessibility of the set

Since $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$ is Lebesgue-irreducible and since $\lambda(\mathcal{C}^{(H)}) > 0$ (because we have $\{x \leq R^{(H)}\} \subset \mathcal{C}^{(H)}$ with $R^{(H)} > 0$ so $\lambda(\mathcal{C}^{(H)}) \geq \lambda(\{x \leq R^{(H)}\}) > 0$) then by the irreducibility definition, $L(x, \mathcal{C}^{(H)}) > 0$ for any $x \in \mathcal{S}$ so $\mathcal{C}^{(H)}$ is accessible.

5. Boundedness of the Lyapunov function

\mathcal{L} is bounded on $\mathcal{C}^{(H)}$ because $\mathcal{C}^{(H)} \subsetneq \mathcal{S}$ is bounded and we have assumed \mathcal{L} of polynomial growth.

The assumptions of Proposition 3.4 are verified so by application of the Proposition to the lagged chain $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$ the announced inequality (3.8) is true. □

Parameters dependence in H

In this paragraph, we are going to exhibit the dependence of $C_{FM}^{(H)}$ in H and q .

From Proposition 3.4,

$$C_{FM}^{(H)} = C \left(\frac{m_H + 1}{\epsilon_H} \right)^{q+1} \frac{M_H^2}{A_H^{2q}}$$

where A_H, M_H, m_H and ϵ_H depend on the Lyapunov constants and set d_H, b_H and $\mathcal{C}^{(H)}$.

Alternative choice of d_H, b_H and $\mathcal{C}^{(H)}$ uniform in H

Let us show that we can find d_{\max}, b_{\max} and \mathcal{C}^{\max} independent from H s.t.

$$\mathbb{E} [\mathcal{L}(V_{H+1}) | V_1 = x] \leq d_{\max} \mathcal{L}(x) + b_{\max} \mathbf{1}_{x \in \mathcal{C}^{\max}}.$$

From Proposition A.7's proof, $\mathbb{E} [\mathcal{L}(V_{H+1}) | V_1 = x]$ satisfies the inequality (A.9):

$$\mathbb{E} [\mathcal{L}(V_{H+1}) | V_1 = x] \leq \delta^H \mathcal{L}(x) + b \frac{1 - \delta^{H-1}}{1 - \delta} + \delta^{H-1} b \mathbf{1}_{\mathcal{C}}(x).$$

We can bound this inequality uniformly in H :

$$\begin{aligned} \mathbb{E} [\mathcal{L}(V_{H+1}) | V_1 = x] &\leq \delta \mathcal{L}(x) + \frac{b}{1 - \delta} + b \mathbf{1}_{\mathcal{C}}(x) \\ &= \frac{1 + \delta}{2} \mathcal{L}(x) - \frac{1 - \delta}{2} \mathcal{L}(x) + \frac{b}{1 - \delta} + b \mathbf{1}_{\mathcal{C}}(x). \end{aligned}$$

Taking R^{\max} s.t. $\frac{b}{1 - \delta} + b < \frac{1 - \delta}{2} \mathcal{L}(x)$ for $|x| > R^{\max}$ and $\mathcal{C}^{\max} = \mathcal{C} \cup \{x \in \mathcal{S} \mid |x| \leq R^{\max}\}$,

$$\mathbb{E} [\mathcal{L}(V_{H+1}) | V_1 = x] \leq \underbrace{\frac{1 + \delta}{2}}_{d^{\max}} \mathcal{L}(x) + \underbrace{\left(\frac{b}{1 - \delta} + b - \frac{1 - \delta}{2} \mathcal{L}(x) \right)}_{b^{\max}} \mathbf{1}_{x \in \mathcal{C}^{\max}}.$$

Choice of M_H and A_H uniform in H

By definition,

$$\begin{aligned} M_H &> \sup_{\mathcal{C}^{(H)}} \mathcal{L} \vee \left(b_H / \left(1 - d_H^{\frac{1}{q}} \right)^q \right), \\ A_H &= (1 - d_H)^{\frac{1}{q}} - \left(\frac{b_H}{M_H} \right)^{\frac{1}{q}}. \end{aligned} \tag{3.10}$$

We can replace b_H , d_H and $\mathcal{C}^{(H)}$ by b_{\max} , d_{\max} and \mathcal{C}^{\max} in (3.10) to have M_H and A_H uniform in H .

Behavior of m_H and ϵ_H

By application of Proposition 3.4 to $\{V_{t+1}, t \in \mathbb{N}\}$, $\{\mathcal{L} \leq M\}$ is a ν_m -small set with minorizing constant ϵ . W.l.o.g., we can assume that $M_{\max} \geq M$. Then, as we have shown in the proof of Proposition 3.5 equation (3.9), a ν_m -small set with minorizing constant ϵ for $\{V_{t+1}, t \in \mathbb{N}\}$ is a ν_m -small set with minorizing constant $\epsilon_H := \epsilon^H$ for $\{V_{tH+1}, t \in \mathbb{N}\}$. As we did in the proof of Proposition 3.5, we can take $\nu_{m_H} = \nu_m$ on $\{\mathcal{L} \leq M\}$ and null on $\{\mathcal{L} \leq M^{\max}\} \setminus \{\mathcal{L} \leq M\}$ so that $\{\mathcal{L} \leq M^{\max}\}$ is a ν_{m_H} small set for $\{V_{t+1}, t \in \mathbb{N}\}$ with minorizing constant ϵ^H .

We give our conclusion in the following proposition:

Proposition 3.6. *Under the assumption of Proposition 3.5, the constant $C_{FM}^{(H)}$ can be upper bounded by:*

$$C \left(\frac{m+1}{\epsilon^H} \right)^{q+1} \frac{M_{\max}(q)^2}{A_{\max}(q)^{2q}}$$

where $M_{\max}(q) > \sup_{\mathcal{C}^{(max)}} \mathcal{L} \vee \left(b_{\max} / \left(1 - d_{\max}^{\frac{1}{q}} \right)^q \right)$, and $A_{\max}(q) := (1 - d_{\max})^{\frac{1}{q}} - \left(\frac{b_{\max}}{M_{\max}(q)} \right)^{\frac{1}{q}}$, R^{\max} s.t. $\frac{b}{1-\delta} + b < \frac{1-\delta}{2} \mathcal{L}(x)$, $\mathcal{C}^{\max} = \mathcal{C} \cup \{x \in \mathcal{S} \mid |x| \leq R^{\max}\}$, $d^{\max} = \frac{1+\delta}{2}$, $b^{\max} = \sup_{x \in \mathcal{C}^{\max}} \left(\frac{b}{1-\delta} + b - \frac{1-\delta}{2} \mathcal{L}(x) \right)$ and ϵ , m are given by the application of Proposition 3.4 on $\{V_{t+1}, t \in \mathbb{N}\}$, and C depends on q only.

Remark 3.2. *This bound goes to infinity when H and q go to infinity, because*

- $\epsilon \in (0, 1)$ so $1/\epsilon^H$ goes to infinity when H goes to infinity,
- $\left(1 - d_{\max}^{\frac{1}{q}} \right)$ goes to 0 when q goes to infinity, so $M_{\max}(q)$ goes to infinity when q goes to infinity.

It means that the control becomes loose when H and q are too large. However, this bound is a uniform bound which can be far from the minimal constant one could get with a more refined analysis.

3.2.2 Completion of Proof of Proposition 3.1

We can now prove our Proposition 3.1.

Proof. Let $q \in [2, p_{\max}]$, let $\mathcal{E}_{N,H}$ be defined by (3.6):

$$\mathcal{E}_{N,H} = \sum_{n=1}^N \mathbf{x}_{t_n}^\top \phi(V_{t_{n+1}}) \mathbf{y}_{t_n} = \sum_{n=1}^N g(V_{t_{n+1}}),$$

$$\phi(V_{t_{n+1}}) = \mathbb{E} \left[\sum_{k=1}^H V_{t_n+k} | \mathcal{F}_{t_n} \right], \quad g(V_{t_n}) = \mathbf{x}_{t_n}^\top \phi(V_{t_{n+1}}) \mathbf{y}_{t_n},$$

with \mathbf{x}_{t_n} and \mathbf{y}_{t_n} defined in (3.7) by $\mathbf{x}_{t_n} = \text{mv}(V_{\text{ref}}) - \text{mv}(cRC_{H,t_n})$ and $\mathbf{y}_{t_n} = \text{mv}(V_{\text{ref}}) + \text{mv}(cRC_{H,t_n})$. We want to control $\mathbb{E} [|\mathcal{E}_{N,H} - N\ell_H|^q]$ in N and H , where $\ell_H = \mathbb{E} \left[\frac{\mathcal{E}_{N,H}}{N} \middle| V_1 \sim \mu \right]$.

In order to apply Proposition 3.4 on $\mathcal{E}_{N,H}$, we have to verify that $g(\cdot)$ is bounded in $\mathcal{L}^{1/q}$ norm. From $\mathbf{H}_{p_{\max}}$, a function is bounded in $\mathcal{L}^{1/q}$ norm if it can be bounded by a polynomial function of order $\frac{p_{\max}}{q} \geq 1$ (because \mathcal{L} is larger than one and of polynomial growth of order p_{\max} at infinity). We are going to show that $g(\cdot)$ is sub-linear. Hence $g(\cdot)$ will be bounded in $\mathcal{L}^{1/q}$ norm.

Sublinearity of g

By triangle inequality, using that $|x_{i,t}| \leq 2c_w$ and $|y_{i,t}| \leq 2c_w$ by their definition as sum and difference of portfolios in \mathcal{W} , for $t \in \mathbb{N}^*$,

$$|g(V_{t+1})| \leq (2c_w)^2 d \sum_{k=1}^H \mathbb{E} [|V_{t+k}| | \mathcal{F}_t].$$

To show that $g(\cdot)$ is sub-linear, we will show that each $\mathbb{E} [|V_{t+k}| | V_{t+1}]$ is sub-linear (e.g. linearly bounded with respect to $|V_{t+1}|$). By Markov property, $\mathbb{E} [|V_{t+k}| | \mathcal{F}_t] = \mathbb{E} [|V_{t+k}| | V_{t+1}]$, the sub-linearity of $g(\cdot)$ will ensue.

Let $k \in \mathbb{N}^*$. By the Jensen inequality in (*) and $\mathbf{H}_{p_{\max}}$:

$$\mathbb{E} [|V_{t+k}| | V_{t+1}] \stackrel{(*)}{\leq} \left(\mathbb{E} [|V_{t+k}|^{p_{\max}} | V_{t+1}] \right)^{\frac{1}{p_{\max}}} \stackrel{\mathbf{H}_{p_{\max}}}{\leq} \left(\frac{1}{c_{\mathcal{L}}} \mathbb{E} [\mathcal{L}(V_{t+k}) | V_{t+1}] \right)^{\frac{1}{p_{\max}}}.$$

By the extended Lyapunov condition (A.9) and $\mathbf{H}_{p_{\max}}$:

$$\begin{aligned} \mathbb{E} [\mathcal{L}(V_{t+k}) | V_{t+1}] &\leq \delta^{k-1} \mathcal{L}(V_{t+1}) + b \underbrace{\frac{1 - \delta^{k-2}}{1 - \delta}}_{b^{(k)}} + \delta^{k-2} b \\ &\leq \delta^{k-1} (C_{\mathcal{L}}(1 + |V_{t+1}|^{p_{\max}})) + b^{(k)}. \end{aligned}$$

Combining these inequalities and applying the inequality $(|x| + |y|)^{\frac{1}{p_{\max}}} \leq |x|^{\frac{1}{p_{\max}}} + |y|^{\frac{1}{p_{\max}}}$,

$$\begin{aligned} \mathbb{E} [|V_{t+k}| | \mathcal{F}_t] &\leq \left(\frac{1}{c_{\mathcal{L}}} (\delta^{k-1} C_{\mathcal{L}} |V_{t+1}|^{p_{\max}}) + \delta^{k-1} C_{\mathcal{L}} + b^{(k)} \right)^{\frac{1}{p_{\max}}} \\ &\leq \left(\frac{\delta^{k-1} C_{\mathcal{L}}}{c_{\mathcal{L}}} \right)^{\frac{1}{p_{\max}}} |V_{t+1}| + \left(\frac{b^{(k)} + \delta^{k-1} C_{\mathcal{L}}}{c_{\mathcal{L}}} \right)^{\frac{1}{p_{\max}}}. \end{aligned}$$

Hence, $\mathbb{E} [|V_{t+k}| | V_{t+1}]$ is sub-linear, and by linear combination, so is $g(\cdot)$.

Proposition 3.5 application: dependence in N

Since $g(\cdot)$ is bounded in $\mathcal{L}^{1/q}$ norm, we can apply the concentration of measure result for lagged chain Proposition 3.5 on $\mathcal{E}_{N,H}$:

$$\mathbb{E} \left[|\mathcal{E}_{N,H} - N\ell_H|^q \middle| V_1 \right] \leq C_{FM}^{(H)} \left(\sup_{\mathcal{S}} \frac{|g - \mu(g)|^q}{\mathcal{L}} \right) \mathcal{L}(V_1) N^{q/2}. \quad (3.11)$$

Bound on $\sup_{\mathcal{S}} \frac{|g - \mu(g)|^q}{\mathcal{L}}$ in terms of \mathbf{H}

Let $v \in \mathcal{S}$. By $g(\cdot)$ definition:

$$\begin{aligned} g(v) - \mu(g) &= \text{mv}(V_{\text{ref}})^\top \phi(v) \text{mv}(V_{\text{ref}}) - \mathbb{E}_\mu \left[\text{mv}(V_{\text{ref}})^\top \phi(V) \text{mv}(V_{\text{ref}}) \right] \\ &\quad - \left(\text{mv}(\phi(v))^\top \phi(v) \text{mv}(\phi(v)) - \mathbb{E}_\mu \left[\text{mv}(V)^\top \phi(V) \text{mv}(V) \right] \right). \end{aligned}$$

Let us denote a the first term:

$$a = \text{mv}(V_{\text{ref}})^\top (\phi(v) - \mathbb{E}_\mu [\phi(V)]) \text{mv}(V_{\text{ref}}).$$

Using that $|\text{mv}(V_{\text{ref}})_i| \leq c_{\mathbf{w}}$, $i \in \{1, \dots, d\}$,

$$|a|^q \leq (c_{\mathbf{w}}^2 d)^q |\phi(v) - \mathbb{E}_\mu [\phi(V)]|^q. \quad (3.12)$$

We can write the second term as:

$$b = \text{mv}(\phi(v))^\top \phi(v) \text{mv}(\phi(v)) - \mathbb{E}_\mu \left[\text{mv}(\phi(V))^\top \phi(V) \text{mv}(\phi(V)) \right].$$

Notice that, by definition of mv , for any $C, \tilde{C} \in \mathcal{S}_+^d$,

$$|\text{mv}(C)^\top C \text{mv}(C) - \text{mv}(\tilde{C})^\top \tilde{C} \text{mv}(\tilde{C})| \leq (c_{\mathbf{w}}^2 d)^q |C - \tilde{C}|. \quad (3.13)$$

Let $C, \tilde{C} \in \mathcal{S}_+^d$. Let us assume that $\text{mv}(C)^\top C \text{mv}(C) - \text{mv}(\tilde{C})^\top \tilde{C} \text{mv}(\tilde{C}) \geq 0$. Then

$$\begin{aligned} &\text{mv}(C)^\top C \text{mv}(C) - \text{mv}(\tilde{C})^\top \tilde{C} \text{mv}(\tilde{C}) \\ &= \underbrace{\text{mv}(C)^\top C \text{mv}(C) - \text{mv}(\tilde{C})^\top C \text{mv}(\tilde{C})}_{\leq 0} + \underbrace{\text{mv}(\tilde{C})^\top C \text{mv}(\tilde{C}) - \text{mv}(\tilde{C})^\top \tilde{C} \text{mv}(\tilde{C})}_{=\text{mv}(\tilde{C})^\top (C - \tilde{C}) \text{mv}(\tilde{C})} \\ &\leq \text{mv}(\tilde{C})^\top (C - \tilde{C}) \text{mv}(\tilde{C}) \leq (c_{\mathbf{w}}^2 d)^q |C - \tilde{C}|. \end{aligned}$$

Conversely, if $\text{mv}(\tilde{C})^\top \tilde{C} \text{mv}(\tilde{C}) - \text{mv}(C)^\top C \text{mv}(C) \geq 0$, then we can do the same reasoning inverting C and \tilde{C} :

$$\text{mv}(\tilde{C})^\top \tilde{C} \text{mv}(\tilde{C}) - \text{mv}(C)^\top C \text{mv}(C) \leq \text{mv}(C)^\top (\tilde{C} - C) \text{mv}(C) \leq (c_{\mathbf{w}}^2 d)^q |C - \tilde{C}|.$$

Applying this inequality with $C = \phi(v)$ and $\tilde{C} = \phi(V)$, we can bound $|b|^q$:

$$|b|^q \leq (c_{\mathbf{w}}^2 d)^q |\phi(v) - \mathbb{E}_\mu [\phi(V)]|^q.$$

Hence we arrived at the same bound as (3.12).

By double conditioning, and invariance of the stationary law: $V_t|V_1 \sim \mu \stackrel{(d)}{=} \mu$,

$$\mathbb{E}_\mu [\phi(V)] = \mathbb{E}_\mu \left[\mathbb{E} \left[\sum_{k=1}^H V_{t+k} | V_{t+1} = V \right] \right] = \mathbb{E}_\mu \left[\sum_{k=1}^H V_{t+k} \right] = HV_\infty.$$

By the former remark and applying the Jensen inequality,

$$\begin{aligned} |\phi(v) - \mathbb{E}_\mu [\phi(V)]|^q &= \left| \mathbb{E} \left[\sum_{k=1}^H V_{t+k} - HV_\infty | V_{t+1} = v \right] \right|^q \\ &\leq \mathbb{E} \left[\left| \sum_{k=1}^H V_{t+k} - HV_\infty \right|^q \middle| V_{t+1} = v \right]. \end{aligned}$$

We can separate $\mathbb{E} \left[\left| \sum_{k=1}^H V_{t+k} - HV_\infty \right|^q \middle| V_{t+1} = v \right]$ between the V_{t+1} measurable term and a term of the form sum of function -here the identity function- of V_{t+2}, \dots, V_{t+H} conditioned on V_{t+1} , i.e. in the right form to apply Proposition 3.4 :

$$\mathbb{E} \left[\left| \sum_{k=1}^H (V_{t+k} - V_\infty) \right|^q \middle| V_{t+1} = v \right] \leq 2^{q-1} \left(|v - V_\infty|^q + \mathbb{E} \left[\left| \sum_{k=2}^H (V_{t+k} - V_\infty) \right|^q \middle| V_{t+1} = v \right] \right).$$

To summarize, we obtain:

$$|g(v) - \mu(g)|^q \leq 2^{2q-1} (c_w^2 d)^q \left(|v - V_\infty|^q + \mathbb{E} \left[\left| \sum_{k=2}^H (V_{t+k} - V_\infty) \right|^q \middle| V_{t+1} = v \right] \right).$$

Proposition 3.4 application: dependence in H

Since $\mathcal{L}(x) \geq c_{\mathcal{L}}|x|^{p_{\max}}$ for large x , and since $\mathcal{L}(x) > 1$ for any $x \in \mathcal{S}$, $\sup_{x \in \mathcal{S}} \frac{|x|^q}{\mathcal{L}(x)}$ is bounded. We can then apply the concentration of measure result for standard chain Proposition 3.4 on the function $\tilde{g}(V_{t+k}) = V_{t+k}$ and $\mu(\tilde{g}) = \mathbb{E}_\mu [V_{t+k}] = V_\infty$:

$$\mathbb{E} \left[\left| \sum_{k=2}^H (V_{t+k} - V_\infty) \right|^q \middle| V_{t+1} = v \right] \leq C_{FM} \sup_{v' \in \mathcal{S}} \frac{|v'|^q}{\mathcal{L}(v')} \mathcal{L}(v) (H-1)^{q/2}. \quad (3.14)$$

Hence, there exists a constant $c_{q,d}$ depending on q, d and on the Lyapunov condition $H_{\mathcal{L}}$ such that, for any $H \geq 1$:

$$\begin{aligned} \frac{|g(v) - \mu(g)|^q}{\mathcal{L}(v)} &\leq 2^{2q-1} (c_w^2 d)^q \left(C_{FM} \sup_{v' \in \mathcal{S}} \frac{|v'|^q}{\mathcal{L}(v')} (H-1)^{q/2} + \frac{|v - V_\infty|^q}{\mathcal{L}(v)} \right) \\ &\leq c_{q,d} H^{q/2}. \end{aligned}$$

For example, $c_{q,d} = 2^{2q} (c_w^2 d)^q \left(C_{FM} \sup_{v' \in \mathcal{S}} \frac{|v'|^q}{\mathcal{L}(v')} + \sup_{v' \in \mathcal{S}} \frac{|v' - V_\infty|^q}{\mathcal{L}(v')} \right)$.

Combining this upper bound with (3.11), we obtain:

$$\mathbb{E}[|\mathcal{E}_{N,H} - N\ell_H|^q] \leq C_{q,\mathcal{L},H,d}(NH)^{q/2}.$$

□

3.3 Proof of Propositions 2.1 and 3.2

In this subsection, we are interested in the characterization of $\ell_H = \mathbb{E}_\mu \left[\frac{cRV_{N,H}^{\text{ref}} - cRV_{N,H}}{N} \right]$ and of the limit value of $\frac{cRV_{N,H}^{\text{ref}} - cRV_{N,H}}{N}$ when N goes to infinity.

Proof. The proof goes as follows: first we show the finiteness and non-negativity of ℓ_H . The convergence result will follow from the concentration of measure result on $\frac{cRV_{N,H}^{\text{ref}} - cRV_{N,H}}{N} = \frac{\mathcal{E}_{N,H}}{N}$.

- **Finiteness**

ℓ_H is defined as the expectation under the stationary law of a linear combination of quantities of type $\text{mv}(V_{\text{ref}})^\top \mathbb{E}[V_{t_n+k} | \mathcal{F}_{t_n}] \text{mv}(V_{\text{ref}})$ and $\text{mv}(cRC_{H,t_n})^\top \mathbb{E}[V_{t_n+k} | \mathcal{F}_{t_n}] \text{mv}(cRC_{H,t_n})$, for $k \in \{1, \dots, H\}$ and $n \in \{1, \dots, N\}$.

Since the $\text{mv}(\cdot)$ operator is bounded, and since the $\{V_t\}_{t \in \mathbb{N}^*}$ admit L_1 moment by \mathbf{H}_{stat} , ℓ_H is well defined and finite.

- **Non-negativity**

By definition of the $\text{mv}(\cdot)$ mapping,

$$\text{mv}(V_{\text{ref}})^\top cRC_{H,t_n} \text{mv}(V_{\text{ref}}) \geq \text{mv}(cRC_{H,t_n})^\top cRC_{H,t_n} \text{mv}(cRC_{H,t_n}) \quad \text{a.s.}$$

Hence, summing on n from 1 to N , we have that $cRV_{N,H}^{\text{ref}} \geq cRV_{N,H}$ almost surely.

Taking the expectation: $0 \leq \mathbb{E}_\mu \left[\frac{cRV_{N,H}^{\text{ref}} - cRV_{N,H}}{N} \right] := \ell_H$.

- **Convergence of $\frac{cRV_{N,H}^{\text{ref}} - cRV_{N,H}}{N}$**

From the moment inequality in Proposition 3.1,

$$\mathbb{E} \left[\left| \frac{\mathcal{E}_{N,H}}{N} - \ell_H \right|^{p_{\max}} \right] \leq C_{q,\mathcal{L},H,d} \left(\frac{H}{N} \right)^{\frac{p_{\max}}{2}}.$$

Hence, $\mathbb{E} \left[\left| \frac{\mathcal{E}_{N,H}}{N} - \ell_H \right|^{p_{\max}} \right] \xrightarrow{N \rightarrow +\infty} 0$ hence the convergence in $L_{p_{\max}}$ -norm.

From the Markov inequality in Proposition 3.1,

$$\mathbb{P} \left(\left| \frac{\mathcal{E}_{N,H}}{N} - \ell_H \right| > a \right) \leq \frac{2C_q H^{\frac{p_{\max}}{2}}}{N^{\frac{p_{\max}}{2}} a^{p_{\max}}}. \quad (3.15)$$

If $p_{\max} > 2$, we can apply the Borel-Cantelli Lemma to show that $\frac{\mathcal{E}_{N,H}}{N}$ converges almost surely towards ℓ_H . Indeed, since $p_{\max} > 2$, for every $a > 0$, from (3.15), $\sum_{N=1}^{\infty} \mathbb{P} \left(\left| \frac{\mathcal{E}_{N,H}}{N} - \ell_H \right| > a \right) < \infty$ hence $\frac{\mathcal{E}_{N,H}}{N} \xrightarrow[N \rightarrow +\infty]{\text{a.s.}} \ell_H$.

□

3.4 Proof of Lemma 3.3

Proof. Let $D_{N,H}(\mathbf{x}) := \sum_{n=1}^N \mathbf{x}_{t_n}^\top (RC_{H,t_n} - cRC_{H,t_n}) \mathbf{x}_{t_n}$, where \mathbf{x} satisfies \mathbf{H}_x . To alleviate the notations, we just write $D_{N,H}$ instead of $D_{N,H}(\mathbf{x})$. We want to control $\mathbb{E}[|D_{N,H}|^q]$ in N and H . The difficulty stems from the fact that we want to exhibit the dependence both in N and H . We are going to rely on a double arguments of martingale: $D_{N,H}$ can be seen as a nested martingale, in the long time period N , and in the short time period H .

Dependence in N

Denote $X_n := \mathbf{x}_{t_n}^\top (RC_{H,t_n} - cRC_{H,t_n}) \mathbf{x}_{t_n}$. Then X_n is $\mathcal{F}_{t_{n+1}}$ -measurable and $D_{N,H} := \sum_{n=1}^N X_n$. By cRC_{H,t_n} definition, $\mathbb{E}[RC_{H,t_n} | \mathcal{F}_{t_n}] = cRC_{H,t_n}$, hence since the \mathbf{x}_{t_n} are \mathcal{F}_{t_n} -measurable, $\mathbb{E}[X_n | \mathcal{F}_{t_n}] = 0$, so $D_{N,H}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_{t_n}\}_{n=1}^{N+1}$.

Let us apply the Burkholder inequality (A.10) on $D_{N,H}$, martingale with increments X_n : denoting $C_{p_{\max}}^B$ the Burkholder constant (depending on p_{\max} only), we have:

$$\mathbb{E}[|D_{N,H}|^{p_{\max}}] \leq C_{p_{\max}}^B \mathbb{E} \left[\left| \sum_{n=1}^N X_n^2 \right|^{p_{\max}/2} \right].$$

By convexity inequality on $\left| \sum_{n=1}^N X_n^2 \right|^{p_{\max}/2}$,

$$\mathbb{E}[|D_{N,H}|^{p_{\max}}] \leq C_{p_{\max}}^B N^{\frac{p_{\max}}{2}-1} \sum_{n=1}^N \mathbb{E}[|X_n|^{p_{\max}}].$$

Dependence in H

We can decompose $X_n = \sum_{k=1}^H \mathbf{x}_{t_n}^\top (\mathbf{r}_{t_n+k} \mathbf{r}_{t_n+k}^\top - \mathbb{E}[V_{t_n+k} | \mathcal{F}_{t_n}]) \mathbf{x}_{t_n}$ in a martingale part $Y_n = \sum_{k=1}^H Y_k^{t_n}$ and a remaining part $Z_n = \sum_{k=1}^H Z_k^{t_n}$, where:

$$\begin{aligned} Y_k^{t_n} &:= \mathbf{x}_{t_n}^\top (\mathbf{r}_{t_n+k} \mathbf{r}_{t_n+k}^\top - V_{t_n+k}) \mathbf{x}_{t_n}, \\ Z_k^{t_n} &:= \mathbf{x}_{t_n}^\top (V_{t_n+k} - \mathbb{E}[V_{t_n+k} | \mathcal{F}_{t_n}]) \mathbf{x}_{t_n}, \quad 1 \leq k \leq H. \end{aligned}$$

Notice that $Z_1^{t_n} = \mathbf{x}_{t_n}^\top (V_{t_{n+1}} - \underbrace{\mathbb{E}[V_{t_{n+1}} | \mathcal{F}_{t_n}]}_{=V_{t_{n+1}}}) \mathbf{x}_{t_n} = 0$.

By convexity inequality

$$\mathbb{E}[|X_n|^{p_{\max}}] = \mathbb{E}[|Y_n + Z_n|^{p_{\max}}] \leq 2^{p_{\max}-1} (\mathbb{E}[|Y_n|^{p_{\max}}] + \mathbb{E}[|Z_n|^{p_{\max}}]).$$

Bound on $\mathbb{E}[|Y_n|^{p_{\max}}]$: By the the Burkholder inequality (A.10) on Y_n and by convexity:

$$\begin{aligned} \mathbb{E}[|Y_n|^{p_{\max}}] &\leq C_{p_{\max}}^B \mathbb{E} \left[\left| \sum_{k=1}^H (Y_k^{t_n})^2 \right|^{\frac{p_{\max}}{2}} \right] \\ &\leq C_{p_{\max}}^B H^{\frac{p_{\max}}{2}-1} \sum_{k=1}^H \mathbb{E}[|Y_k^{t_n}|^{p_{\max}}]. \end{aligned}$$

Bound on $\mathbb{E}[|Z_n|^{p_{\max}}]$: since the \mathbf{x}_{t_n} components are bounded by $c_{\mathbf{w}}$ from assumption \mathbf{H}_x ,

$$\mathbb{E}[|Z_n|^{p_{\max}}] \leq (c_{\mathbf{w}}^2 d)^{p_{\max}} \mathbb{E} \left[\left| \sum_{k=2}^H V_{t_n+k} - \mathbb{E}[V_{t_n+k} | \mathcal{F}_{t_n}] \right|^{p_{\max}} \right].$$

Making appear $\mathbb{E}_{\mu}[V_t] = \mathbb{E}[V_{t_n+k} | V_{t_n} \sim \mu] = V_{\infty}$, by convexity inequality:

$$\mathbb{E}[|Z_n|^{p_{\max}}] \leq 2^{p_{\max}-1} (c_{\mathbf{w}}^2 d)^{p_{\max}} \left(\mathbb{E} \left[\left| \sum_{k=2}^H (V_{t_n+k} - V_{\infty}) \right|^{p_{\max}} \right] + \mathbb{E} \left[\left| \sum_{k=2}^H (V_{\infty} - \mathbb{E}[V_{t_n+k} | \mathcal{F}_{t_n}]) \right|^{p_{\max}} \right] \right).$$

Notice that the second term is bounded by the first one, since the conditional expectation is non-expansive in $L_{p_{\max}}$ -norm,

$$\mathbb{E} \left[\left| \sum_{k=2}^H (\mathbb{E}[V_{t_n+k} | \mathcal{F}_{t_n}] - V_{\infty}) \right|^{p_{\max}} \right] = \mathbb{E} \left[\left| \sum_{k=2}^H (V_{t_n+k} - V_{\infty}) \right|^{p_{\max}} \right]. \quad (3.16)$$

Hence

$$\mathbb{E}[|Z_n|^{p_{\max}}] \leq 2^{p_{\max}} (c_{\mathbf{w}}^2 d)^{p_{\max}} \mathbb{E} \left[\left| \sum_{k=2}^H (V_{t_n+k} - V_{\infty}) \right|^{p_{\max}} \right].$$

As we have done in (3.14), we can apply Fort-Moulines Proposition extension (Proposition 3.5) to get the dependence in H :

$$\mathbb{E} \left[\left| \sum_{k=2}^H (V_{t_n+k} - V_{\infty}) \right|^q \middle| V_{t_n+1} \right] \leq C_{FM} \sup_{x \in \mathcal{S}} \frac{|x|^q}{\mathcal{L}(x)} \mathcal{L}(V_{t_n+1}) (H-1)^{q/2}.$$

So,

$$\mathbb{E}[|Z_n|^{p_{\max}}] \leq (2c_{\mathbf{w}}^2 d)^{p_{\max}} C_{FM} \sup_{x \in \mathcal{S}} \frac{|x|^q}{\mathcal{L}(x)} \mathbb{E}[\mathcal{L}(V_{t_n+1}) | V_1] H^{q/2}.$$

Combining the obtained bounds on Z_n and Y_n , we obtain:

$$\mathbb{E} [|X_n|^{p_{\max}}] \leq 2^{p_{\max}-1} \left(C_{p_{\max}}^B \underbrace{\frac{\sum_{k=1}^H \mathbb{E} [|Y_k^{t_n}|^{p_{\max}}]}{H}}_{m_{Y,n}} + (2c_{\mathbf{w}}^2 d)^{p_{\max}} C_{FM} \sup_{x \in \mathcal{S}} \frac{|x|^q}{\mathcal{L}(x)} \underbrace{\mathbb{E} [\mathcal{L}(V_{t_n+1})|V_1]}_{m_{\mathcal{L},n}} \right) H^{\frac{p_{\max}}{2}},$$

$$\mathbb{E} [|D_{N,H}|^{p_{\max}}] \leq \underbrace{2^{p_{\max}-1} C_{p_{\max}}^B \left(C_{p_{\max}}^B \frac{\sum_{n=1}^N m_{Y,n}}{N} + (2c_{\mathbf{w}}^2 d)^{p_{\max}} C_{FM} \sup_{x \in \mathcal{S}} \frac{|x|^q}{\mathcal{L}(x)} \frac{\sum_{n=1}^N m_{\mathcal{L},n}}{N} \right)}_{C_{p_{\max},d,\mathcal{L}}} (NH)^{\frac{p_{\max}}{2}}.$$

□

3.5 Auxiliary result

In what follows, we prove the Lemma 2.3, e.g. that when $V_{\text{ref}} = V_{\infty}$, $\lim_{H \rightarrow \infty} \frac{\ell_H^{\infty}}{H} = 0$.

Proof. Let us first show that $\frac{cRC_{H,T_{reb}}}{H}$ converges to V_{∞} . Let $T_{reb} \geq 2$.

As we did in Lemma 3.3 equation (3.16), by the non-expansivity of the conditional expectation in L_p -norm,

$$\begin{aligned} \mathbb{E} [|cRC_{H,T_{reb}} - HV_{\infty}|^{p_{\max}}] &\leq \mathbb{E} \left[\left| \sum_{k=1}^H V_{T_{reb}+k} - HV_{\infty} \right|^{p_{\max}} \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\left| \sum_{k=1}^H V_{T_{reb}+k} - HV_{\infty} \right|^{p_{\max}} \middle| \mathcal{F}_{T_{reb}-1} \right] \right]. \end{aligned}$$

And as done in (3.14),

$$\mathbb{E} \left[\left| \sum_{k=1}^H V_{T_{reb}+k} - HV_{\infty} \right|^{p_{\max}} \middle| V_{T_{reb}} \right] \leq C_{FM} \sup_{x \in \mathcal{S}} \frac{|x|^{p_{\max}}}{\mathcal{L}(x)} \mathcal{L}(V_{T_{reb}}) H^{\frac{p_{\max}}{2}}.$$

Taking the expectation, $\mathbb{E} [\mathcal{L}(V_{T_{reb}})|V_1]$ is finite by the Lyapunov condition drift $\mathbf{H}_{\mathcal{L}}$.

So dividing by $H^{p_{\max}}$:

$$\mathbb{E} \left[\left| \frac{cRC_{H,T_{reb}}}{H} - V_{\infty} \right|^{p_{\max}} \middle| V_{T_{reb}} \right] \leq C_{FM} \sup_{x \in \mathcal{S}} \frac{|x|^{p_{\max}}}{\mathcal{L}(x)} \mathbb{E} [\mathcal{L}(V_{T_{reb}})|V_1] \frac{1}{H^{\frac{p_{\max}}{2}}}. \quad (3.17)$$

When H goes to infinity, the bound in (3.17) goes to zero so $\frac{cRC_{H,T_{reb}}}{H} \xrightarrow[H \rightarrow +\infty]{L_{p_{\max}}} V_{\infty}$. By homogeneity and continuity of the mv operator which is Lipschitz from equation (3.13), and passage to the limit, since $\frac{cRC_{H,T_{reb}}}{H} \xrightarrow[H \rightarrow +\infty]{L_{p_{\max}}} V_{\infty}$ with $p_{\max} \geq 2$, we have convergence in L_1 norm and we can infer:

$$\frac{\ell_H^{\infty}}{H} = \mathbb{E}_{V_1 \sim \mu} \left[\text{mv}(V_{\infty})^{\top} \frac{cRC_{H,T_{reb}}}{H} \text{mv}(V_{\infty}) - \text{mv} \left(\frac{cRC_{H,T_{reb}}}{H} \right)^{\top} \frac{cRC_{H,T_{reb}}}{H} \text{mv} \left(\frac{cRC_{H,T_{reb}}}{H} \right) \right] \xrightarrow[H \rightarrow +\infty]{} 0.$$

□

4 Model specification

We want to specify models satisfying the assumptions (2.1), \mathbf{H}_{stat} , $\mathbf{H}_{\mathcal{S}}$, $\mathbf{H}_{\mathcal{L}}$ and $\mathbf{H}_{\rho_{\max}}$.

4.1 Motivation: GARCH-CCC model

The GARCH-CCC model [Bol90] is one of the simplest extension of GARCH model to the multidimensional case. For an exhaustive review of multidimensional GARCH, see [BLR06]. In this work, we elaborate our results with the simple GARCH-CCC, presented in [Car01].

This model assumes the following structure for the centered returns \mathbf{r}_t :

$$\begin{aligned} \mathbf{r}_t &= D_t \Gamma^{1/2} \boldsymbol{\eta}_t = D_t \tilde{\boldsymbol{\eta}}_t, \\ \tilde{\boldsymbol{\eta}}_t &= \Gamma^{1/2} \boldsymbol{\eta}_t, \\ V_t &= D_t \Gamma D_t, \end{aligned} \tag{4.1}$$

where

- $D_t = \text{Diag}(\sigma_{1,t}, \dots, \sigma_{d,t})$,
- $\{\sigma_{i,t}\}_{1 \leq i \leq d}$ are one-dimensional GARCH volatilities,
- $\Gamma = \{\rho_{ij}\}_{1 \leq i, j \leq d}$ is a positive definite matrix (the Constant Conditional Correlation matrix),
- $\boldsymbol{\eta}_t$ is a d -dimensional vector with independent components, $\mathbb{E}[\eta_{i,t} \eta_{j,t}] = 0 \forall i \neq j$, $\mathbb{E}[\eta_{i,t}^2] = 1$, and $\boldsymbol{\eta}_t$ independent from D_t .

The original GARCH-CCC assumes a simple GARCH(1,1) volatility recursion for the $\sigma_{i,t}$:

$$\sigma_{i,t}^2 = w_i + \alpha_i r_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2, \quad i = 1, \dots, d, \tag{4.2}$$

where $w_i, \alpha_i, \beta_i \in \mathbb{R}$.

The advantages of the "initial" GARCH-CCC model is its parsimony: it requires $\frac{d(d+5)}{2}$ parameters versus $\frac{d(5d+1)}{2}$ for the interdependent GARCH-CCC with $p = q = 1$ (4.4). The disadvantage is the strong assumption of constant conditional correlation.

In what follows, we will assume that $\{\boldsymbol{\eta}_t\}_{t \in \mathbb{N}}$ is a sequence of independent Gaussian vectors.

4.2 Stationarity, ergodicity and application of results

In this subsection, we show that the GARCH-CCC satisfies the general model equation (2.1) and the model assumptions \mathbf{H}_{stat} , $\mathbf{H}_{\mathcal{S}}$, $\mathbf{H}_{\mathcal{L}}$ and $\mathbf{H}_{\rho_{\max}}$, under the following assumptions and definition on the parameters:

$\mathbf{H}_{\text{param}}$. *The GARCH-CCC parameters $\{w_i, \alpha_i, \beta_i\}_{i=1}^d$ and Γ are deterministic and satisfy:*

- i. $\Gamma \in \mathcal{S}_+^d$ is a Correlation matrix.
- ii. $w_i > 0$, $\alpha_i > 0$ and $\beta_i > 0$ for all $i \in \{1, \dots, d\}$.

iii. $3\alpha_i^2 + \beta_i^2 + 2\alpha_i\beta_i < 1$ for all $i \in \{1, \dots, d\}$,

$q_{\max} \in \mathbb{R}_+$ is defined as:

$$q_{\max} := \min_{1 \leq i \leq d} \arg \max_{p \in \mathbb{R}_+} \left\{ \mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^p] < 1 \right\}, \quad \tilde{\eta} \sim \mathcal{N}(0, \Gamma). \quad (4.3)$$

Remark 4.1. A few remarks on H_{param} :

- The condition H_{param} (iii) is equivalent to the condition $\mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^2] < 1$ (using that $\mathbb{E} [\tilde{\eta}_i^4] = 3$ and $\mathbb{E} [\tilde{\eta}_i^2] = 1$). Hence, it implies that $q_{\max} > 2$. This condition ensures in particular $\alpha_i + \beta_i < 1$, hence the existence of a stationary and ergodic solution to the model, as shown in the Appendix Subsection A.2.
- Continuity and evolution in p of $\mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^p]$:
 - By L_p norms growth in p , if $q > 0$ is s.t. $|\alpha_i \tilde{\eta}_i^2 + \beta_i|_q^q = \mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^q] < 1$, then for $0 \leq p \leq q$,

$$|\alpha_i \tilde{\eta}_i^2 + \beta_i|_p \leq |\alpha_i \tilde{\eta}_i^2 + \beta_i|_q < 1.$$
 - The function $g : p \mapsto \mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^p]$ is continuous (for example, by dominated convergence, for any sequence p_n converging to p , $g(p_n)$ converges to $g(p)$).
 - The function $g(\cdot)$ goes to infinity when p goes to infinity (because $\alpha_i > 0$), hence we are sure to have p s.t. $g(p) \geq 1$.
- Since $g(\cdot)$ is continuous, the quantity $\max_{p \in \mathbb{R}_+} \left\{ \mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^p] : \mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^p] < 1 \right\}$ is equal to 1. To prove the Lyapunov drift condition, we need a p s.t. $\mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^p] < 1$. We denote p_{\max} such a p . From H_{param} (iii) and the first remark, we can take $p_{\max} \geq 2$.

Theorem 4.1. Under H_{param} , the GARCH-CCC variances process $\{\sigma_t^2\}_{t \in \mathbb{N}}$ satisfies the assumption of Theorem 2.2. H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$.

In what follows, we are going to prove the properties H_{stat} , H_S , $H_{\mathcal{L}}$ and $H_{p_{\max}}$ on the GARCH-variances process σ_t^2 , on the stable state-space $\mathcal{S} = \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i}, +\infty \right)$.

Proof. Let $\{\sigma_t^2\}_{t \in \mathbb{N}}$ be a GARCH-CCC variances process.

H_{stat} : Stationarity and ergodicity

From Theorem A.3 explicit condition (A.2), a sufficient condition for stationarity and ergodicity for σ_t^2 is implied by our assumption H_{param} (iii).

H_S on σ_t^2 : By definition, $\{\sigma_t^2\}_{t \in \mathbb{N}^*}$ is Markovian and time homogeneous. We show in our Corollary A.5 that $\{\sigma_t^2\}_{t \in \mathbb{N}^*}$ is Lebesgue-irreducible on $\mathcal{S} = \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i}, +\infty \right)$. The aperiodicity will follow from the existence of a small set of positive measure shown below.

$H_{\mathcal{L}}$ and $H_{p_{\max}}$ on σ_t^2 : For $\mathbf{x} \in \mathcal{S}$, define

$$\mathcal{L}(\mathbf{x}) = 1 + \sum_{i=1}^d |x_i - w_i|^{p_{\max}},$$

and $\rho = \max_{1 \leq i \leq d} \mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^{p_{\max}}]$ and $\delta = \frac{1+\rho}{2}$.

Then, if $\boldsymbol{\sigma}_{t-1}^2 = \mathbf{x}$, $\mathbb{E}_{\mathbf{x}} \mathcal{L}(\boldsymbol{\sigma}_t^2) = 1 + \mathbb{E} \sum_{i=1}^d |(\alpha_i \tilde{\eta}_i^2 + \beta_i) x_i|^{p_{\max}} \leq 1 + \rho (\sum_{i=1}^d |x_i|^{p_{\max}})$.

Since $\lim_{|\mathbf{x}| \rightarrow \infty} \frac{\mathbb{E}_{\mathbf{x}} \mathcal{L}(\boldsymbol{\sigma}_t^2)}{\mathcal{L}(\mathbf{x})} \leq \rho$, the drift condition $\mathbb{E}_{\mathbf{x}} \mathcal{L}(\boldsymbol{\sigma}_t^2) \leq \delta \mathcal{L}(\mathbf{x})$ is verified for \mathbf{x} large enough. Let $R > \max_{1 \leq i \leq d} \frac{w_i}{1-\beta_i}$ s.t. it is verified for $|\mathbf{x}| > R$, $\mathcal{C} := \{\mathbf{x} \in \mathcal{S} \mid |\mathbf{x}| \leq R\}$ and $b := \sup_{\mathbf{x} \in \mathcal{C}} |\mathbb{E}_{\mathbf{x}} \mathcal{L}(\boldsymbol{\sigma}_t^2) - \delta \mathcal{L}(\mathbf{x})|$, then

$$\mathbb{E}_{\mathbf{x}} \mathcal{L}(\boldsymbol{\sigma}_t^2) \leq \delta \mathcal{L}(\mathbf{x}) + b \mathbf{1}_{\mathbf{x} \in \mathcal{C}}, \quad \mathbf{x} \in \mathcal{S}.$$

By definition, $\boldsymbol{\sigma}_t^2 | \boldsymbol{\sigma}_{t-1}^2 = \mathbf{x}$ follows the same law as $\{w_i + (\alpha_i \tilde{\eta}_i^2 + \beta_i) x_i\}$, $\tilde{\boldsymbol{\eta}} \sim N(0, \Gamma)$. It admits a density $g_{\mathbf{x}}$ derived explicitly in Subsection A.4 equation (A.4).

Properties of $g_{\mathbf{x}}$

- $\mathbf{v} \mapsto g_{\mathbf{x}}(\mathbf{v})$ is continuous in \mathbf{v} for $v_i > w_i + \beta_i x_i$,
- For $(\mathbf{x}, \mathbf{v}) \in \underbrace{\left(B(0, R) \cup \prod_{i=1}^d \left[\frac{w_i}{1-\beta_i}, +\infty \right) \right)}_{\mathcal{C}} \times \left(\prod_{i=1}^d \left[\frac{w_i + \beta_i R}{2} + \frac{R}{2}, +\infty \right) \right) =: D$, we have

$$\begin{aligned} \frac{v_i - w_i - \beta_i x_i}{\alpha_i x_i} &\geq \frac{\frac{w_i + \beta_i R}{2} + \frac{R}{2} - w_i - \beta_i R}{\alpha_i R} \\ &= \frac{1}{2\alpha_i} \underbrace{\left(1 - \beta_i - \frac{w_i}{R} \right)}_{>0 \text{ by } R \text{ choice}} > 0. \end{aligned}$$

This function restricted to D is hence continuous and positive.

In particular, $\inf_{\mathbf{x} \in \mathcal{C}} g_{\mathbf{x}}(\mathbf{v}) > 0$ for each $\mathbf{v} \in \prod_{i=1}^d \left[\frac{w_i + \beta_i R}{2} + \frac{R}{2}, +\infty \right)$.

Let

$$\nu(A) := \int_A \inf_{\mathbf{x} \in \mathcal{C}} g_{\mathbf{x}}(\mathbf{v}) \mathbf{1}_{\left\{ v_i > \frac{w_i + \beta_i R}{2} + \frac{R}{2} \right\}_{i=1}^d} d\mathbf{v}, \quad A \in \mathcal{B}(\mathcal{S}).$$

For any $\mathbf{x} \in \mathcal{C}$, $P(\mathbf{x}, A) = \int_A g_{\mathbf{x}}(\mathbf{v}) d\mathbf{v} \geq \nu(A)$ with ν non-null measure with density, hence \mathcal{C} is a petite set.

Since $\frac{w_i + \beta_i R}{2} + \frac{R}{2} < R$ (because $R > \frac{w_i}{1-\beta_i}$), $\prod_{i=1}^d \left[\frac{w_i + \beta_i R}{2} + \frac{R}{2}, R \right]$ is non empty and has a positive Lebesgue measure and is included in \mathcal{C} . Hence:

$$\nu(\mathcal{C}) \geq \int_{\left\{ \frac{w_i + \beta_i R}{2} + \frac{R}{2} < v_i < R \right\}_{i=1}^d} \inf_{\mathbf{x} \in \mathcal{C}} g_{\mathbf{x}}(\mathbf{v}) d\mathbf{v} > 0$$

so \mathcal{C} is accessible and as a consequence of [MT09, Aperiodicity definition - page 114], the chain is aperiodic. Hence the Lyapunov condition is satisfied, with \mathcal{L} polynomial in p_{\max} .

H_{stat}: Existence of L_2 -moment From the Lyapunov drift condition, we are ensured to have finite p_{\max} moments. Since we assumed $p_{\max} \geq 2$, it implies the square integrability of the process $\{\boldsymbol{\sigma}_t^2\}_{t \in \mathbb{N}^*}$ for any initial condition and under the stationary law. \square

Bijection between $\{V_t\}_{t \in \mathbb{N}}$ and σ_t^2

Since $\sigma_{i,t}^2 > 0$ on \mathcal{S} , there exists a function Φ_Γ s.t.

$$(V_t)_{i,j} = \begin{cases} \rho_{ij}\sigma_{i,t}\sigma_{j,t} & \text{if } i \neq j, \\ \sigma_{i,t}^2 & \text{else} \end{cases} =: \Phi_\Gamma(\sigma_t^2).$$

- Φ_Γ is bijective,
- Φ_Γ is sub-linear: $|\Phi_\Gamma(\sigma_t^2)| \leq |\Gamma||\sigma_t^2|$.

By Theorem A.1, if σ_t^2 admits an ergodic and stationary solution, so is $\Phi_\Gamma(\sigma_t^2) = V_t$.

Results application

- By definition of the GARCH-CCC model (4.1), $\{V_t = \Phi_\Gamma(\sigma_t^2)\}_{t \in \mathbb{N}^*}$ satisfies (2.1).
- If $\{\sigma_t^2\}_{t \in \mathbb{N}^*}$ satisfies \mathbf{H}_{stat} , $\mathbf{H}_{\mathcal{S}}$, $\mathbf{H}_{\mathcal{L}}$ and $\mathbf{H}_{p_{\max}}$, then the ergodic concentration results (Propositions 3.5 and 3.4) apply to any function $g(\Phi_\Gamma(\sigma_t^2))$, $g \circ \Phi_\Gamma$ bounded in $L^{1/q}$ norm, $q \geq 2$.
Since Φ_Γ is sub-linear, our Proposition 3.1, Lemma 3.3 and main result Theorem 2.2 apply without any adaptation.

Hence, under parameters condition $\mathbf{H}_{\text{param}}$, by Theorem 4.1 and these remarks, the main result Theorem 2.2 is satisfied with p_{\max} defined in equation (4.3).

4.3 Additional models

In this subsection, we list additional GARCH models on which it is possible to extend our study. We first list one-dimensional GARCH models which can be extended to multidimensional model with a constant correlation matrix. Then we refer to a multidimensional model with inter-dependence between the GARCH volatilities.

4.3.1 One-dimensional GARCH models

The Threshold GARCH (T-GARCH) and the asymmetric power-GARCH are affine models, hence we can directly apply our results on them:

- T-GARCH:

$$\sigma_t = w + [\alpha_+(\eta_{t-1})_+ - \alpha_-(\eta_{t-1})_- + \beta] \sigma_{t-1},$$

where $w, \alpha_+, \alpha_-, \beta > 0$.

- Power GARCH:

$$\sigma_t^\delta = w + [\alpha(|\eta_{t-1}| - \zeta \eta_{t-1})^\delta + \beta] \sigma_{t-1}^\delta,$$

where $w, \alpha, \beta, \delta > 0$, $|\zeta| \leq 1$.

Indeed, we can adapt the assumptions verification taking the same Lyapunov function defined on σ_t instead of σ_t^2 for the T-GARCH, and σ_t^δ for the power GARCH.

The following models are of the form $\sigma_t \leq a(\eta_{t-1})\sigma_{t-1}^\gamma$ with $\gamma < 1$ if $\alpha + \beta < 1$. We can adapt our results on these models by taking a linear Lyapunov function.

- Exponential GARCH:

$$\begin{aligned}\log \sigma_t^2 &= w + \alpha g(\eta_{t-1}) + \beta \log \sigma_{t-1}, \\ g(\eta_{t-1}) &= \theta \eta_{t-1} + \zeta (|\eta_{t-1}| - \mathbb{E}[|\eta_{t-1}|]), \\ \sigma_t^2 &= e^w e^{\alpha g(\eta_{t-1})} (\sigma_{t-1}^2)^\beta,\end{aligned}$$

with $\alpha, \beta > 0$, $\beta < 1$ and $-\zeta < \theta < \zeta$ for g to be increasing in $|\eta_{t-1}|$ (and $\theta < 0$ for negative innovation to have more impact than positive ones).

In this model, in the Gaussian innovation case, moments exist at any order [FZ19, p. 79].

- Log-GARCH: the recursion holds on the logarithm of variances:

$$\begin{aligned}\log \sigma_t^2 &= w + \alpha \log r_{t-1} + \beta \log \sigma_{t-1} \\ &= w + \alpha \log \eta_{t-1}^2 + (\alpha + \beta) \log \sigma_{t-1}, \\ \sigma_t^2 &= e^w (\eta_{t-1})^\alpha \sigma_{t-1}^{\alpha+\beta},\end{aligned}$$

where $w, \alpha, \beta \in \mathbb{R}$.

4.3.2 Multidimensional GARCH models

In Francq and Zakoian's book [FZ19, page 280], the authors define the interdependent GARCH volatilities model:

$$\boldsymbol{\sigma}_t^2 = \boldsymbol{w} + A \boldsymbol{r}_{t-1}^2 + B \boldsymbol{\sigma}_{t-1}^2, \quad (4.4)$$

where $A, B \in \mathbb{R}^{d \times d}$, $\boldsymbol{w} \in \mathbb{R}^d$.

It is possible to adapt our results to this setting.

5 Numerical experiments

In this section, we confront our theoretical results to the real probabilities of better performance of the realized covariance based portfolio. To do so, we base ourselves on GARCH fitted parameters on real financial data to have realistic range of parameters (Subsection 5.1). Then we simulate GARCH-CCC based returns and compute the corresponding cRC_{H,t_n} and my matrices and portfolio, as described in Subsection 5.2. Finally, we exhibit the evolution of $\hat{\mathbb{P}}(R_{N,H} < R_{N,H}^\infty)$ for multiple values of H and N .

5.1 Realistic GARCH values, fit procedure

We have fitted GARCH-CCC parameters on real financial time series, using a two steps procedure:

1. fit of the one-dimensional GARCH(1,1) models by maximum likelihood (where the explicit likelihood for one-dimensional GARCH can be found in [EB86] for example) on the recentered returns (we assumed a fixed drift),
2. fit of the constant conditional correlation on the reconstructed residuals $\tilde{\eta}_{i,t}$.

The financial time series consist in the 12 components of BNPP QIS Momentum Strategy, composed of 4 indices consisting in index futures rolling (indices being EuroStoxx50, S&P 500, Nikkei, HSCEI), 3 indices of rolling of futures of bonds (German Bund, US government bond, Japanese government bond) and 5 indices of rolling of commodity (on gold, brent, S&P GSCI excess return, Goldman Sachs US industrial metal ER) fx hedged in euros, from 12/03/1993 until 02/02/2017 (6046 dates).

	Minimum	Maximum	Average	Standard-deviation
σ_∞^2	0.060	5.15	1.77	1.44
$\sigma_{\infty,an}$	4.04	36.02	18.91	9.47
w	0.00032	0.04320	0.0135	0.0137
α	0.0329	0.0955	0.0622	0.0242
β	0.885	0.966	0.931	0.0271
$\alpha + \beta$	0.981	0.998	0.993	0.00182
ρ	-22%	82%	10%	22%
p_{\max}	1.83	5.72	3.34	1.22
$t_{1/2}$	36.2	437.2	135.8	98.52

Table 1: Summary of variance (σ_∞^2), GARCH(1,1) parameters (α, β, w), constant conditional correlation (ρ), highest moment (p_{\max}) and half-life time ($t_{1/2}$, as defined in (5.2)) fitted on the 12 components of BNPP QIS Momentum Strategy from 12/03/1993 until 02/02/2017 (6046 dates). We give the minimum, maximum, average and standard-deviation of the calibrated parameters. The variance and w parameters are expressed in basis points of daily variances, $\sigma_{\infty,an}$ denotes the annualized volatility and is expressed in percents, $t_{1/2}$ is expressed in days. α , β and p_{\max} have no dimension. What we call standard-deviation is the standard-deviation over the estimated quantities.

We give the minimum, maximum, average and standard-deviation of the obtained values in Table 1. In the same Table, we also give the minimum, maximum, average and standard-deviation of additional quantities computed on the parameters:

- the maximum moment $p_{\max,i}$, computed on the parameters with formula (4.3)

$$p_{\max,i} := \arg \max_{p \in \mathbb{R}_+} \{ \mathbb{E} [|\alpha_i \tilde{\eta}_i^2 + \beta_i|^p] < 1 \}, \quad (5.1)$$

- the half-life $t_{1/2}$, defined as the average time for the gap between the variance level to its long-term level to be reduced by one half:

$$t_{1/2,i} := -\frac{\log 2}{\log(\alpha_i + \beta_i)}, \quad (5.2)$$

- $\sigma_{i,\infty}^2$ the empirical variance measured on the data.

We display in Figure 2 the obtained values for α , β and p_{\max} .

Comments and interpretation

- The average α is relatively small (0.062) and the average β is relatively high (0.931), and the sum $\alpha + \beta$ is very close to 1. According to Campbell [CLM97, page 483], the α parameter measures how much a previous shock (high value of an innovation) will propagate to the future.
- The half-life times are large (because the $\alpha + \beta$ are close to 1): this phenomena corresponds to the persistence of shocks over time. The average half-life time is 135 days hence approximately six months. It means that before this horizon, the instantaneous variance can be far from its expectation under the stationary distribution.
- The w_i values are very close to zero since by stationarity condition:

$$\sigma_{i,\infty}^2 = \mathbb{E} [\epsilon_{i,t}^2] = \mathbb{E} [\sigma_{i,t}^2] = \frac{w_i}{1 - \alpha_i - \beta_i},$$

they should be of the same order of magnitude than $\sigma_{i,\infty}^2(1 - \alpha_i - \beta_i)$, and $\alpha_i + \beta_i \approx 1$.

- We notice that low values of α are associated to high values of β (as shown in Figure 2) and conversely, in such a way that the stationary condition $\alpha + \beta < 1$ is always enforced.
- The average p_{\max} is equal to 3.34, and the minimum is slightly smaller than 2. As shown in Figure 2, most p_{\max} values are higher than 2 which means that the time series have finite kurtosis.

The estimated parameters show that our approach of considering portfolio with horizon H not too large (for example: 21 days) and $p_{\max} \geq 2$ but of small order is relevant since the average half-time is very large (more than a month) and the existence of order-2 moments is almost always verified.

5.2 Simulation procedure

In what follows, we describe how we choose the GARCH-CCC parameters and how we estimate the benchmark covariance V_∞ and the realized conditional covariance $cRC_{H,T_{reb}}$.

- **$\{\alpha_i, \beta_i, \Gamma_{i,j}\}_{i,j=1}^d$ choice**

The $\{\alpha_i, \beta_i, \Gamma_{i,j}\}_{i,j=1}^d$ are simulated uniformly in the ranges indicated in Table 1. The (α_i, β_i) are sorted such that the smallest α_i are associated to the largest β_i . (We enforce Γ to be definite positive by ensuring or capping its eigen-values to 10^{-2} and transforming the matrix to retrieve a 1-diagonal.)

- **V_∞ estimation**

We estimate V_∞ as the empirical covariance of the returns over a long range of time, typically three times the largest half-life time. It amounts to consider a backtest of three times the largest half-life time. Given the half-life times observed range, the largest life-time is around 400 business days, hence 1.6 years: 3 times this period corresponds to a backtest size of 4.8 years.

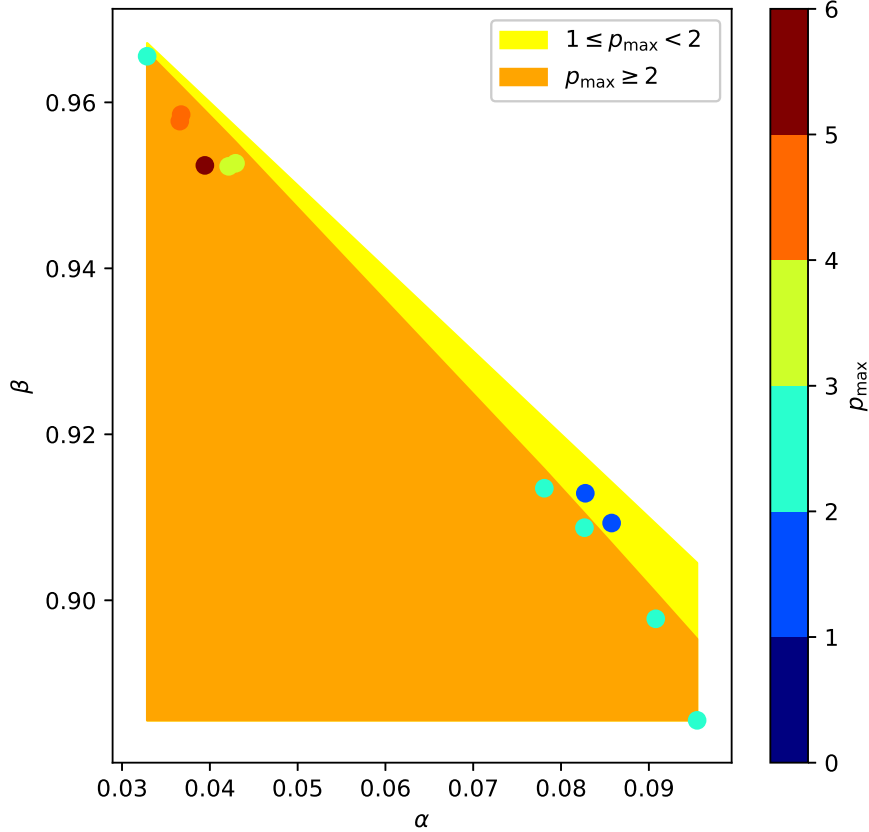


Figure 2: $(\alpha, \beta, p_{\max})$ parameters fitted on 12 financial time series. In orange, we draw the finite variance area (condition $H_{\text{param}}(\text{iii})$), in yellow, we draw the stationary (condition (A.2)) but not finite variance area. A dot corresponds to one (α_i, β_i) fitted parameter. Its color indicates its associated p_{\max} value (as defined in $H_{\text{param}}(p_{\max})$). The color scale to the right corresponds to ranges of observed values for p_{\max} : lower than 2 values are in dark blue, $p_{\max} \in [2, 3]$ is in cyan, higher values are in green, orange and brown. Most of our fitted parameters are colored in cyan, i.e. are associated to $p_{\max} \in [2, 3]$ values.

- **$cRC_{H,T_{reb}}$ estimation**

We evaluate the $cRC_{H,T_{reb}}$ via Monte Carlo:

$$\widehat{cRC}_{H,T_{reb}} = \sum_{k=1}^H \widehat{\mathbb{E}} [V_{T_{reb}+k} | \mathcal{F}_{T_{reb}}] = \frac{1}{N_{mc}} \sum_{n=1}^{N_{mc}} \sum_{k=1}^H V_{T_{reb}+k}^{(n)} | \mathcal{F}_{T_{reb}}$$

with

- $V_{T_{reb}}^{(n)} = V_{T_{reb}}$,
- $V_{T_{reb}+k}^{(n)}$ obtained by application of the recursion formula on $V_{T_{reb}+k}^{(n)}$ and a generated $r_{T_{reb}+k}^{(n)} \sim \mathcal{N}(0, V_{T_{reb}+k-1}^{(n)})$, for $k = 1, \dots, H$.

We took $N_{mc} = 100$.

We consider the minimum variance portfolio without constraints: $\mathcal{W} = \{\mathbf{w} = \{w_i\}_{i=1}^d \in \mathbb{R}^d : \mathbf{1}_d^\top \mathbf{w} = 1\}$, where $\mathbf{1}_d$ is the d -dimensional vector of ones. Then $\text{mv}(\cdot)$ is explicit: for $C \in \mathbb{R}^{d \times d}$ positive definite matrix, $\text{mv}(C) = \frac{C^{-1} \mathbf{1}_d}{\mathbf{1}_d^\top C^{-1} \mathbf{1}_d}$. Since we deal with a non-degenerate GARCH-CCC model, we know that our allocation weights will be bounded during the length of the experiment. This is why we did not enforce a bound on their norm.

We will place in the following settings:

- $d \in \{10, 50\}$,
- $H = 1$ (daily rebalancing), $H = 5$ (weekly), $H = 21$ (monthly),
- $N = 1, \dots, 6 * 21/H$.

For a daily rebalancing period, we will consider up to $N = 126$ rebalancing times of the portfolio, and for monthly time period, up to $N = 6$ rebalancing times.

For one set of d -dimensional GARCH parameters, we reproduce the following experiment multiple times ($N_{MC} = 10^4$ times):

- We let the multidimensional GARCH evolve during 3 times the maximal half-life time associated to the d sets of GARCH parameters. We do so in order to reach plausible GARCH values beyond the burn-in phase of the process. We precise that our results are valid whatever the starting point and that it is not necessary to reach the stationary regime for the results to apply.
- **Computation of the probabilities: trajectory approach (less independent) versus by (H, N) approach (more independent).**

To do our experiments, we first considered a trajectory approach, consisting in the following steps:

We start the allocation procedure by initializing our daily, weekly and monthly portfolios: $\text{mv}(V_1)$, $\text{mv}(\widehat{cRC}_{5,1})$ and $\text{mv}(\widehat{cRC}_{21,1})$. Then for $t = 1, \dots, 126$,

- We update the daily realized variance $R_{t,1}$ and we update the portfolio $\text{mv}(V_t)$, and the realized variance.

- If t modulo 5 is null, we update the weekly realized variance $R_{t,5}$ and portfolio with $\widehat{cRC}_{5,t}$.
- If t modulo 21 is null, we update the monthly realized variance $R_{t,21}$ and portfolio with $\widehat{cRC}_{21,t}$.

Another more independent (but more time consuming) approach consists in generating a new GARCH trajectory for each (H, N) pair considered. This is the approach we have finally taken in the following illustrations (except for the $d = 50$ experiment).

Hence, we can compute iteratively the empirical probabilities:

$$\widehat{\mathbb{P}}(R_{N,H} < R_{N,H}^{\infty}) = \frac{1}{N_{MC}} \sum_{N=1}^{N_{MC}} \mathbf{1}_{R_{N,H} < R_{N,H}^{\infty}}, \quad H \in \{1, 5, 21\}, \quad N = 1, \dots, \frac{126}{H}.$$

5.3 Empirical probability: impact of p_{\max} and d

We display in Figures 3 and 4 the evolution of the empirical $\mathbb{P}(R_{N,H} > R_{N,H}^{\infty})$ with H and N , for $d = 10$ and small versus large p_{\max} (Figure 3), and in the other figures, we compare large dimension ($d = 50$) versus small dimension ($d = 10$) (Figure 4). The time is in business days. We compare daily (red dots), weekly (blue) and monthly (green) rebalancing times of the portfolios. On all the experiments, the probabilities decreased towards 0.

5.3.1 Impact of p_{\max}

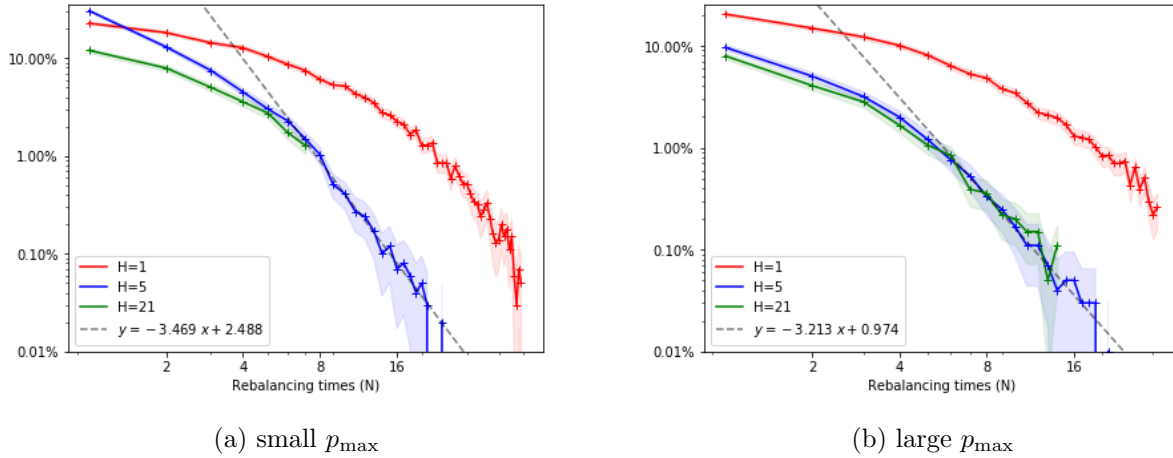


Figure 3: Empirical probabilities $\widehat{\mathbb{P}}(R_{N,H} > R_{N,H}^{\infty})$, for $H = 1$ (red), $H = 5$ (blue) and $H = 21$ (green) as a function of the number of rebalancing times N , in log-log scale, for a set of GARCH parameters such that $d = 10$, $p_{\max} = 2.8$ (left) and $p_{\max} = 7$ (right). In black dots, we plotted the tendency line fitted on $\widehat{\mathbb{P}}(R_{N,H} > R_{N,H}^{\infty})$ for $H = 5$ and $N \geq 5$.

In Figure 3 we compare at fixed dimension $d = 10$ the impact of the integrability. For the figure on the left, it requires more than 7 rebalancing of the monthly or weekly portfolio to reach a 1%

probability level that the GARCH-covariance will outperform the benchmark, whereas in the large p_{\max} case, the 1% level is reached at the 5th rebalancing of the portfolio. Practically speaking, it means that for a period of 5 weeks (weakly portfolio) or 5 months (monthly portfolio), we have a 99% probability that the GARCH-based portfolio has a lowest realized variance than the benchmark. Since a higher p_{\max} implies lighter tails, the deviation are less important and it is easier to see the performance gain. For $N \geq 4$, the tendency lines have a slope of -3.4 (small p_{\max} case) versus -3.2 (large p_{\max} case).

Link with Theorem 2.2: We observe as expected a decreasing shape of the probability $\mathbb{P}\left(R_{N,H} > R_{N,H}^{\infty}\right)$, but with a concave evolution and not the linear, with a $-\frac{p_{\max}}{2}$ slope, expected shape. Our results are not conclusive for N small (the rate of decline is lower than expected). For N large, we even observe a faster convergence than expected in the small p_{\max} case (slope of -3.4 instead of $-\frac{p_{\max}}{2} = -1.4$). This is an indication that there can be several decay regimes depending on N value. This is hard to give a clear, quantitative explanation of this phenomena. It can be argued that our model consists in a vectorized version of multiple unidimensional GARCH processes, each of them being associated to different tail thickness. Since we deal at the portfolio variance level, it is possible than GARCH with lighter tails are associated to higher components in the minimum variance allocation, which could explain this faster than expected convergence rate. Nonetheless, these results are not in contradiction with our theoretical results. Indeed, our results can be interpreted as a "worst-case" bound on the probability, so it is not surprising to do better. And the probability levels are lower for the large p_{\max} case, which is coherent with the bound behavior.

5.3.2 Impact of the dimension

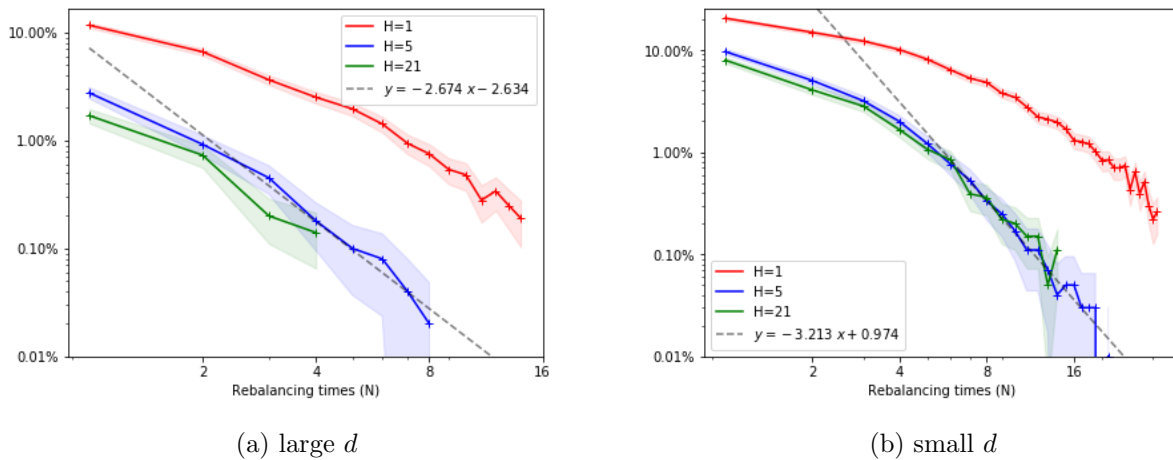


Figure 4: Empirical probabilities $\widehat{\mathbb{P}}(R_{N,H} > R_{N,H}^{\infty})$, for $H = 1$ (red), $H = 5$ (blue) and $H = 21$ (green) as a function of the number of rebalancing times N , in log-log scale, for a set of GARCH parameters such that $d = 50$, $p_{\max} = 5.05$ (left) and $d = 10$, $p_{\max} = 7.03$ (right).

In Figure 4, we display the evolution of the empirical $\mathbb{P}\left(R_{N,H} > R_{N,H}^{\infty}\right)$ for large p_{\max} and very different dimensions: $d = 50$ (Figure 4 (a)) and $d = 10$ (Figure 4 (b)).

Comparing 4 (a) and (b), the dimension does not seem to have a strong impact, or might be

compensated by the stronger impact of the large integrability.

In the large dimension case, the tendency line has a slope of -2.6 which is lower than the theoretical minimum convergence rate of $-\frac{p_{\max}}{2} = -2.025$ (hence the convergence is faster than the theoretical bound).

5.3.3 Impact of the number of rebalancing times

In Figure 5, we display the histogram of the realized variance difference for a monthly portfolio with 10 assets and different N . We see that for $N = 1$, the realized variance difference is almost centered around 0 but for $N \geq 4$, all the distributions are significantly centered on positive values. Indeed, in our setting (initial condition following approximately the stationary law), the average realized variance difference is equal to the renormalized performance gap $\frac{\ell_H^\infty}{H}$ which is positive and deterministic as shown in Proposition 2.1.

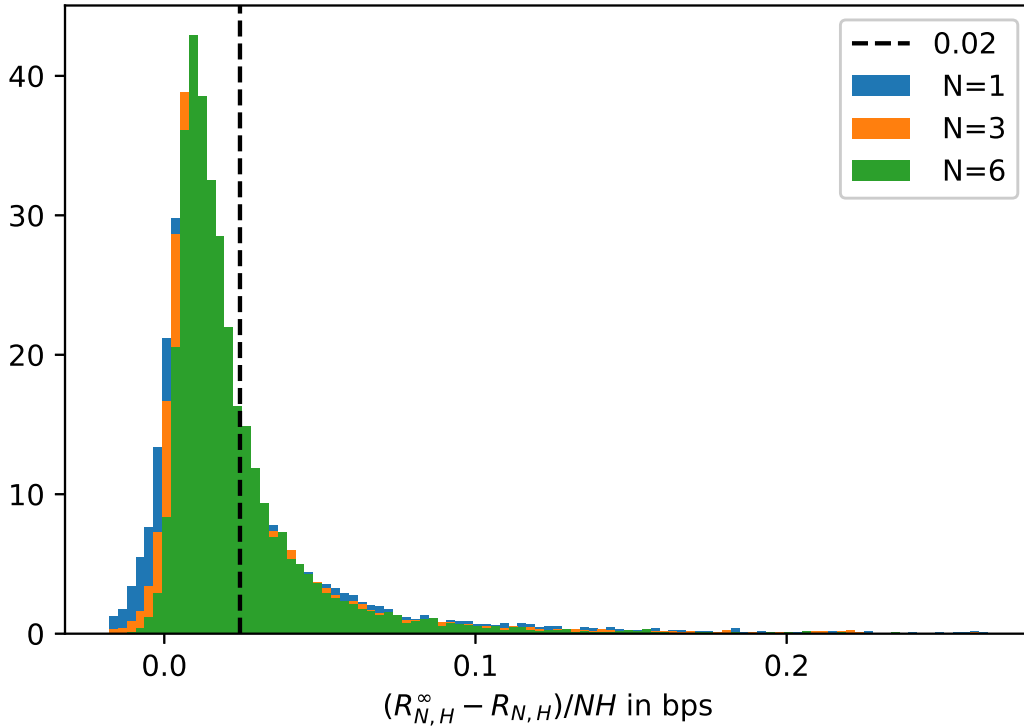


Figure 5: Histograms of $\left\{ \frac{R_{N,H}^\infty - R_{N,H}}{NH} \right\}_{N \in \{1,6,9,12\}}$ with $H = 21$ for a set of GARCH parameters such that $d = 10$, $p_{\max} = 2.8$. Here the variance differences are expressed as the difference in the realized variance over N rebalancing times of the portfolio, in basis points of daily variance. $N = 6$ corresponds to a variance difference over a six months, between the benchmark portfolio and a monthly re-updated portfolio. In black dots, we plot the average of $\frac{R_{N,H}^\infty - R_{N,H}}{NH}$ for $N = 6$ and $H = 21$.

The x scale is a daily variance measure in basis points. We see that in average, the monthly portfolio

implies a drop of daily variance of 0.02 bps. This variance level has to be put in perspective: we consider a range of daily variance of [0.06, 5.2] bps in our experiments.

In terms of associated annualized volatility, we can upper bound the volatility difference by the square-root of the variance difference, which translates to a drop of $\sqrt{0.02 * 252/10^4} = 2.2\%$.

5.3.4 Impact of the investment period

As mentioned in the comments and interpretations paragraph following our main result, if we are in the $(NH) \ll C(H/\ell_H)^2$ regime, the asset manager should not bother much using a sophisticated estimation for V_t : a good approximation of V_∞ is enough. In this Subsubsection, we aim at illustrating this comment. In particular, we aim at answering the question: *starting from which investment period H do we have $R_{N,H} > R_{N,H}^\infty$ in half the cases?* For this H threshold, we are indifferent between the stationary and the conditional realized covariance matrices.

We are interested in the evolution of $\mathbb{P}\left(R_{N,H} < R_{N,H}^\infty\right)$ for increasing H values. For sake of simplicity, we consider only one rebalancing of the portfolio ($N = 1$). For larger N , the probability levels will be even lower according to the previous experiments and theoretical results, so $N = 1$ can be seen as a worst case-scenario.

In Figure 6, we display the evolution of $\mathbb{P}\left(R_{1,H} < R_{1,H}^\infty\right)$ with H in months and in log-log scale, for a universe of 10 assets associated to an integrability of order $p_{\max} = 2.8$ and a half-life of 39 days hence approximately 2 months. As previously observed in Figure 3 (same set of parameters), for $H = 21$ days (so one month) and $N = 1$, the probability is 90%. Then the probability is decreasing with the number of months and it is approximately equal to $\frac{1}{2}$ when H is equal to one-year and a half (18 months). The investment period must therefore be as long as 9 times the half-life of the process for the stationary covariance to be as efficient as the GARCH-based covariance.

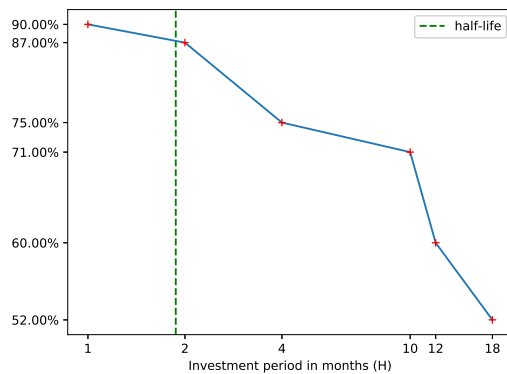


Figure 6: Empirical probabilities $\widehat{\mathbb{P}}(R_{1,H} < R_{1,H}^\infty)$ as a function of the investment period H for a set of GARCH parameters such that $d = 10$, $p_{\max} = 2.8$, $t_{1/2} = 39$ days.

For investment period lower than a year, the GARCH covariance is hence relevant and efficient.

5.4 Tail exponent Hill estimation

From (3.3), there is a $C > 0$ independent from N s.t.

$$\bar{F}_{N,H}(y) = \mathbb{P} \left(\frac{D_{N,H}}{\sqrt{NH}} > y \right) \leq \frac{C}{y^{p_{\max}}}.$$

In what follows, we are going to compare our estimation of p_{\max} via numerical method on equation (5.1) and the one obtained via estimation of the tail exponent of $\frac{D_{N,H}}{\sqrt{NH}}$. To do so, we use the Hill estimator.

Hill estimator

The Hill estimator estimates the distribution tail exponent, e.g. the exponent γ such that its complementary cumulative distribution function \bar{F} verifies:

$$\bar{F}(x) = \frac{\ell(x)}{x^{1/\gamma}}$$

where $\lim_{x \rightarrow \infty} \frac{\ell(tx)}{\ell(x)} = 1$ for any $t > 0$.

For X_1, \dots, X_n random variables of cumulative distribution function F , denoting $X_{1,n} < X_{2,n} < \dots < X_{n,n}$ the ordered variables, the Hill estimator of order $k \in \mathbb{N}$ is defined by:

$$\hat{\gamma}(k) = \frac{1}{k} \sum_{i=0}^{k-1} (\log X_{n-i,n} - \log X_{n-k,n}).$$

It converges in probability towards γ if k increases with n with $k/n \xrightarrow[n \rightarrow \infty]{} 0$ (cf [DHF06, Theorem 3.2.4]) and converges asymptotically under a second order condition, as detailed in Appendix A.5.

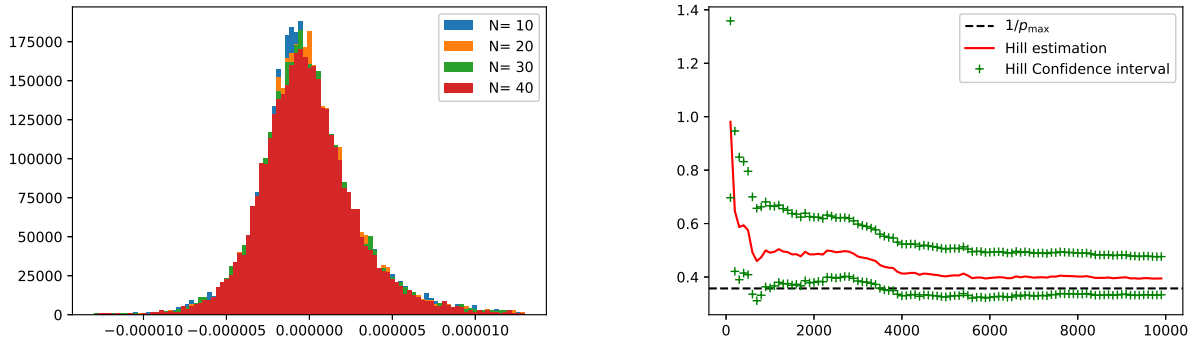
Reasonable choice of k

Assuming a Fréchet domain of attraction (this is the natural attraction domain for Pareto-like distributions), we can get an explicit formulation of the asymptotic variance and minimize it in k . It gives the following optimal k for $\gamma \neq 1$:

$$k_{opt}(\gamma, n) = 2 \left(\frac{\gamma}{1-\gamma} n \right)^{\frac{2}{3}}. \tag{5.3}$$

The details on how we get this result is postponed to Appendix A.5. If we expect a certain value of γ , we can plug it in (5.3) to estimate k_{opt} .

Experiment



(a) Histogram of $\frac{D_{N,H}}{\sqrt{N}}$ for $H = 1$ and varying N , $p_{\max} = 2.8$ and $d = 10$ assets. (b) Hill estimation on 10^4 simulations of $\frac{D_{N,H}}{\sqrt{N}}$, for $H = 1$ and $N = 40$, with $p_{\max} = 2.8$ and $d = 10$ assets.

We display the histogram of $\frac{D_{N,1}}{\sqrt{N}}$ for multiple N and the Gaussian density (black dots) in Figure 7a and the Hill estimator $\hat{\gamma}(k_{\text{opt}}(1/p_{\max}, n))$ in Figure 7b. We have computed the Hill estimator on an increasing number of samples, $n \in \{100, 200, \dots, 10\,000\}$ of the renormalized martingale $\frac{D_{N,1}}{\sqrt{N}}$, $N = 40$.

Comments:

- From Figure 7a, we see that \sqrt{N} is a good renormalization for the $D_{N,1}$ martingale. Since we have finite variance, a Central Limit Theorem seems to be verified.
- From Figure 7b, the Hill estimator is overall close to $1/p_{\max}$. We plotted the confidence interval associated to the Central Limit theorem for the Hill estimator (A.5). $1/p_{\max}$ is inside the confidence interval.

6 Conclusion, perspectives

In this work, assuming a time-dependent conditional model on the returns of the form $r_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, V_t)$, we give the optimal covariance for a given period and time of investment for the minimum variance problem. Using a decomposition between a martingale and a positive ergodic term, we can show that our covariance-based portfolio has a lower realized variance than any other benchmark covariance with a high probability, which is increasing with the number of rebalancing of the portfolio.

We give an explicit recursion scheme for the computation of this covariance matrix for the specific GARCH-CCC model. This recursion scheme could be adapted to many other models. We empirically illustrate our result by computing the empirical probability that the realized variance of our optimal covariance portfolio is smaller than the one with the stationary covariance matrix. The experiment results are not in contradiction with our theoretical analysis: we verify the convergence of the GARCH superperformance probability with the number of rebalancing, at a rate at least equal to $\frac{p_{\max}}{2}$ in the large N regime.

The question the practitioner may ask is *(when) is it relevant to use this more sophisticated covariance matrix, rather than a simple empirical covariance?* Under the assumption that the model is

GARCH with known parameters, we have shown that the performance gap goes to zero when the investment period goes to infinity and the threshold of the investment period for both covariances to perform equally well seems to be around 10 times the half-life of the process. But we do not tackle in this work the fact that in practice, the models have to be fitted, leading to estimation error on the parameters, not to mention model error. We can advocate that estimation error can still be tackled by taking into account the Gaussian uncertainty on the parameters, stemming from the central limit behavior of the parameters when the backtest size is large enough.

A Appendix

A.1 Ergodicity and stationarity

Theorem A.1 ([FZ19, Theorem A.1 page 367]). *If $(Z_t)_{t \in \mathbb{Z}}$ is an ergodic strictly stationary sequence and if $(Y_t)_{t \in \mathbb{Z}}$ is defined by*

$$Y_t = f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots),$$

where f is a measurable function from \mathbb{R}^∞ to \mathbb{R} , then $(Y_t)_{t \in \mathbb{Z}}$ is also an ergodic strictly stationary sequence.

Theorem A.2 (The ergodic theorem for stationary sequences [FZ19, Theorem A.2]). *If $(Z_t)_{t \in \mathbb{Z}}$ is strictly stationary and ergodic, if f is measurable and if*

$$\mathbb{E}[|f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots)|] < \infty$$

then

$$\frac{1}{n} \sum_{t=1}^n f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots) \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots)] \quad a.s.$$

A.2 GARCH-CCC ergodicity

Before recalling Francq and Zakoian's result on the GARCH-CCC stationarity, let us introduce the definition of the *top Lyapunov exponent*.

The GARCH-CCC model (4.1) can be put under the following vector form.

Let $\mathbf{z}_t = \begin{pmatrix} \mathbf{r}_t^2 \\ \boldsymbol{\sigma}_t^2 \end{pmatrix} \in \mathbb{R}^{2d}$ and let $E_t = \text{Diag}(\tilde{\boldsymbol{\eta}}_t^2)$. Then \mathbf{z}_t satisfies the recursion:

$$\begin{aligned} \mathbf{z}_t &= \mathbf{b}_t + A_t \mathbf{z}_{t-1}, \\ \text{where } \mathbf{b}_t &= \begin{pmatrix} E_t \mathbf{w} \\ \mathbf{w} \end{pmatrix}, \quad A_t = \begin{pmatrix} E_t A & E_t B \\ A & B \end{pmatrix}, \end{aligned} \tag{A.1}$$

$\mathbf{w} = \{w_i\}_{i=1}^d$, $A = \text{Diag}(\{\alpha_i\}_{i=1}^d)$, $B = \text{Diag}(\{\beta_i\}_{i=1}^d)$.

The top Lyapunov exponent is defined by [FZ19, Theorem 2.3]:

$$\gamma := \lim_{t \rightarrow \infty} \text{a.s.} \frac{1}{t} \log |A_t A_{t-1} \dots A_1|.$$

Theorem A.3 (Strict stationarity of the CCC model [FZ19, Theorem 10.6]). *A necessary and sufficient condition for the existence of a strict stationary and non-anticipative solution process for the model (4.1) is $\gamma < 0$, where γ is the top Lyapunov exponent of the sequence $\{A_t, t \in \mathbb{Z}\}$ defined in (A.1). This stationary and non-anticipative solution, when $\gamma < 0$, is unique and ergodic.*

Explicit γ formulation and condition:

Notice that γ does not depend on the chosen matricial norm, because norms are equivalent on the finite dimension space of matrices considered, and if $\frac{1}{K}|\cdot|_2 \leq |\cdot|_1 \leq K|\cdot|_2$, $K > 0$,

$$-\frac{\log(K)}{t} + \frac{1}{t} \log|A_t A_{t-1} \dots A_1|_2 \leq \frac{1}{t} \log|A_t A_{t-1} \dots A_1|_1 \leq \frac{\log(K)}{t} + \frac{1}{t} \log|A_t A_{t-1} \dots A_1|_2$$

and by passage to the limit, the term $\frac{\log(K)}{t}$ vanishes when t goes to infinity.

In what follows, we are going to specify an explicit sufficient condition s.t. $\gamma < 0$. W.l.o.g., we consider as matricial norm the infinite norm: $|A| = \max_{1 \leq i, j \leq d} |A_{i,j}|$, $A \in \mathbb{R}^{d \times d}$.

We can specify the top Lyapunov condition by noticing that A_t can be written as: $A_t = \begin{pmatrix} E_t \\ I_d \end{pmatrix} (A \ B)$,

and for $1 \leq s \leq t$, $(A \ B) \begin{pmatrix} E_s \\ I_d \end{pmatrix} = \text{Diag} \left(\left\{ \alpha_i \tilde{\eta}_{s,i}^2 + \beta_i \right\}_{i=1}^d \right)$.

It follows, using the sub-additivity of the infinite norm, that:

$$\begin{aligned} \log|A_t A_{t-1} \dots A_1| &= \log \left| \begin{pmatrix} E_t \\ I_d \end{pmatrix} \prod_{s=1}^{t-1} \text{Diag} \left(\left\{ \alpha_i \tilde{\eta}_{s,i}^2 + \beta_i \right\}_{i=1}^d \right) (A \ B) \right| \\ &\leq \log \left| \begin{pmatrix} E_t \\ I_d \end{pmatrix} \right| + \log |(A \ B)| + \sum_{s=1}^{t-1} \log \left| \text{Diag} \left(\left\{ \alpha_i \tilde{\eta}_{s,i}^2 + \beta_i \right\}_{i=1}^d \right) \right|, \end{aligned}$$

and by the strong law of large numbers, $\gamma \leq \mathbb{E} \log \left| \text{Diag} \left(\left\{ \alpha_i \tilde{\eta}_i^2 + \beta_i \right\}_{i=1}^d \right) \right|$, $\tilde{\eta} \sim \mathcal{N}(0, \Gamma)$.

Taking as matricial norm the infinite norm, the condition

$$\mathbb{E} \log \max_{1 \leq i \leq d} (\alpha_i \tilde{\eta}_i^2 + \beta_i) < 0$$

is a sufficient condition for $\gamma < 0$.

It can be verified via numerical evaluation for example.

Easier to verify sufficient conditions

Since for fixed $\tilde{\eta}_i^2$: $\log \max_{1 \leq i \leq d} (\alpha_i \tilde{\eta}_i^2 + \beta_i) \leq \sum_{i=1}^d \log(\alpha_i \tilde{\eta}_i^2 + \beta_i)$, and taking the expectation:

$$\gamma \leq \mathbb{E} \log \max_{1 \leq i \leq d} (\alpha_i \tilde{\eta}_i^2 + \beta_i) \leq \sum_{i=1}^d \mathbb{E} \log(\alpha_i \tilde{\eta}_i^2 + \beta_i) \leq d \max_{1 \leq i \leq d} \mathbb{E} \log(\alpha_i \tilde{\eta}_i^2 + \beta_i),$$

the following condition is a sufficient condition for $\gamma < 0$:

$$\max_{1 \leq i \leq d} \mathbb{E} \log(\alpha_i \tilde{\eta}_i^2 + \beta_i) < 0.$$

The following condition ensures also the ergodicity condition, as a consequence of the Jensen inequality on $\mathbb{E} \log(\alpha_i \tilde{\eta}_i^2 + \beta_i)$:

$$\alpha_i + \beta_i < 1, \quad i \in \{1, \dots, d\}. \tag{A.2}$$

A.3 Irreducibility and auxiliary results on the GARCH-CCC

We will give very explicit arguments which are the main adding of this study, compared to Chapters 3 "Mixing properties of univariate GARCH" and 10 "Multivariate GARCH" of Francq and Zakoian's book [FZ19].

In this subsection, we assume $\mathbf{H}_{\text{param}}$.

The following lemma enables us to prove that $\mathcal{S} = \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i}, +\infty \right)$ is a stable state space, e.g. that from any point $\mathbf{x} \in \mathcal{S}$, we can reach any point as close from $\frac{w_i}{1-\beta_i}$ as wanted in a finite number of steps.

Lemma A.4. *Let $\mathbf{x} \in \mathcal{S} = \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i}, +\infty \right)$. For all $c > 0$, there exists a finite $N_c \in \mathbb{N}$ such that:*

$$\mathbb{P} \left(\sigma_{i,N_c}^2 < \frac{w_i}{1-\beta_i} + c \mid \boldsymbol{\sigma}_0^2 = \mathbf{x} \right) > 0, \quad \forall i = 1, \dots, d.$$

Proof. Applying recursively equation (4.2) on $\sigma_{i,N}^2$, for $\boldsymbol{\sigma}_0^2 = \mathbf{x} \in \mathcal{S}$, $N \in \mathbb{N}^*$, with $\tilde{\eta}_n \sim \mathcal{N}(0, \Gamma)$, $n = 1, \dots, N-1$,

$$\begin{aligned} \sigma_{i,N}^2 &= w_i + (\alpha_i \tilde{\eta}_{i,N-1}^2 + \beta_i) \left(w_i + (\alpha_i \tilde{\eta}_{i,N-2}^2 + \beta_i) (w_i + \dots + (\alpha_i \tilde{\eta}_{i,0}^2 + \beta_i) \sigma_{i,0}^2) \right) \\ &= w_i \left(1 + (\alpha_i \tilde{\eta}_{i,N-1}^2 + \beta_i) + \dots + \prod_{n=0}^{N-2} (\alpha_i \tilde{\eta}_{i,N-1-n}^2 + \beta_i) \right) \\ &\quad + \prod_{n=0}^{N-1} (\alpha_i \tilde{\eta}_{i,N-1-n}^2 + \beta_i) x_i. \end{aligned} \tag{A.3}$$

Case $\forall i \in \{1, \dots, d\}$, $\frac{w_i}{1-\beta_i} + c > w_i + \beta_i x_i$.

By definition, $\sigma_{i,1}^2 = w_i + (\alpha_i \tilde{\eta}_i^2 + \beta_i) x_i$, $\tilde{\eta} \sim \mathcal{N}(0, \Gamma)$: $\sigma_{i,1}^2$ can be seen as a function of $\tilde{\eta}_i^2 \geq 0$. For small values of $\tilde{\eta}_i^2$, with non null probability, $w_i + \beta_i x_i \leq \sigma_{i,1}^2 < \frac{w_i}{1-\beta_i} + c$: $N_c = 1$ works.

Case $\exists i \in \{1, \dots, d\}$, $\frac{w_i}{1-\beta_i} + c < w_i + \beta_i x_i$

Let $\epsilon > 0$ s.t. $\alpha_i \epsilon^2 + \beta_i < 1$. Let us consider the event $\{|\tilde{\eta}_{i,n}| < \epsilon, n = 0, \dots, N-1, i = 1, \dots, d\}$. From (A.3),

$$\begin{aligned} \sigma_{i,N}^2 &\leq w_i \sum_{n=0}^{N-1} (\alpha_i \epsilon^2 + \beta_i)^n + (\alpha_i \epsilon^2 + \beta_i)^N x_i \\ &= w_i \frac{1 - (\alpha_i \epsilon^2 + \beta_i)^N}{1 - (\alpha_i \epsilon^2 + \beta_i)} + (\alpha_i \epsilon^2 + \beta_i)^N x_i. \end{aligned}$$

This quantity goes to $\frac{w_i}{1-(\alpha_i \epsilon^2 + \beta_i)}$ when N goes to infinity.

Given $c > 0$, let us choose ϵ such that $\frac{w_i}{1-(\alpha_i \epsilon^2 + \beta_i)} < \frac{w_i}{1-\beta_i} + \frac{c}{2}$.

We can take N_c finite such that $(\alpha_i \epsilon^2 + \beta_i)^{N_c} x_i < \frac{c}{2}$.

Then $w_i \frac{1-(\alpha_i \epsilon^2 + \beta_i)^{N_c}}{1-(\alpha_i \epsilon^2 + \beta_i)} + (\alpha_i \epsilon^2 + \beta_i)^{N_c} x_i < \frac{w_i}{1-\beta_i} + c$ and then

$$\mathbb{P}_{\mathbf{x}} \left(\sigma_{i,N_c}^2 \in \left(\frac{w_i}{1-\beta_i}, \frac{w_i}{1-\beta_i} + c \right) \right) \geq \mathbb{P} (|\tilde{\eta}_{i,n}| < \epsilon, n = 1, \dots, N_c, i = 1, \dots, d) > 0.$$

□

Corollary A.5 (Irreducibility). *The Markov process $\{\sigma_t^2\}_{t \in \mathbb{Z}}$ taking values in \mathcal{S} is Lebesgue-irreducible [MT09, Proposition 4.2.1 (ii)], e.g. for any $A \subset \mathcal{B}(\mathcal{S})$ with strictly positive Lebesgue measure, for any $\mathbf{x} \in \mathcal{S}$*

$$\mathbb{P}_{\mathbf{x}}[\arg \min\{n \geq 1 \mid \sigma_n^2 \in A\} < \infty] > 0.$$

Proof. Let $A \subset \mathcal{B}(\mathcal{S})$, $\lambda(A) > 0$ (where $\lambda(\cdot)$ denotes the Lebesgue measure) and $\mathbf{x} \in \mathcal{S}$.

We define $\tau_A := \arg \min\{n \geq 1 \mid \sigma_n^2 \in A\}$.

Let $c > 0$ such that $\lambda\left(A \cap \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i} + c, +\infty\right)\right) > 0$. It exists since

- $c \mapsto \lambda\left(A \cap \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i} + c, +\infty\right)\right)$ is increasing when c decreases,
- $\lim_{c \rightarrow 0} \lambda\left(A \cap \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i} + c, +\infty\right)\right) = \lambda(A) > 0$.

Let ϵ and N_c defined in lemma A.4 s.t. $\mathbb{P}_{\mathbf{x}}\left(\sigma_{i,N_c}^2 \in \left(\frac{w_i}{1-\beta_i}, \frac{w_i}{1-\beta_i} + c\right)\right) > 0$. Then we can show that $\mathbb{P}_{\mathbf{x}}(\tau_A = N_c + 1) > 0$.

Indeed,

$$\begin{aligned} \mathbb{P}_{\mathbf{x}}(\sigma_{N_c+1}^2 \in A) &\geq \mathbb{P}_{\mathbf{x}}\left(\sigma_{N_c+1}^2 \in A \cap \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i} + c, +\infty\right), \sigma_{N_c}^2 \in \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i}, \frac{w_i}{1-\beta_i} + c\right)\right) \\ &= \mathbb{E}\left[\underbrace{\mathbb{P}\left(\sigma_{N_c+1}^2 \in A \cap \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i} + c, +\infty\right) \mid \sigma_{N_c}^2\right)}_{p(\sigma_{N_c}^2)} \mathbf{1}_{\sigma_{N_c}^2 \in \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i}, \frac{w_i}{1-\beta_i} + c\right)}\right] \end{aligned}$$

where $p(\sigma_{N_c}^2) > 0$ for any $\sigma_{N_c}^2$, and $\left\{\sigma_{N_c}^2 \in \prod_{i=1}^d \left(\frac{w_i}{1-\beta_i}, \frac{w_i}{1-\beta_i} + c\right)\right\}$ is of non null probability.

Hence $\mathbb{P}_{\mathbf{x}}(\tau_A = N_c + 1) > 0$ and $\{\sigma_t^2\}_{t \in \mathbb{Z}}$ is Lebesgue-irreducible. □

A.4 GARCH-CCC density

Let ψ be a test function. Let us denote g_{Γ} the density of $\tilde{\boldsymbol{\eta}} \sim \mathcal{N}(0, \Gamma)$. We want to compute the density of the vector $\tilde{\boldsymbol{\eta}}^2 = \{\tilde{\eta}_i^2\}_{i=1}^d$:

$$\begin{aligned} \mathbb{E}\psi(\tilde{\boldsymbol{\eta}}^2) &= \int_{\mathbb{R}^d} \psi\left(\{x_i^2\}_{i=1}^d\right) g_{\Gamma}\left(\{x_i\}_{i=1}^d\right) dx_1 \dots dx_d \\ &= \sum_{\substack{s_i = \pm \\ 1 \leq i \leq d}} \int_{\mathbb{R}_{s_1} \dots \mathbb{R}_{s_d}} \psi\left(\{x_i^2\}_{i=1}^d\right) g_{\Gamma}\left(\{x_i\}_{i=1}^d\right) dx_1 \dots dx_d. \end{aligned}$$

We denote $s_i = \pm$ sign symbol denoting the part of \mathbb{R} interval on which x_i is integrated. We can make a change of variable: denoting

$$z_i = \begin{cases} x_i & \text{if } s_i = +, \\ -x_i & \text{if } s_i = -, \end{cases}$$

and $\phi_{\mathbf{s}}$ s.t. $\mathbf{x} = \phi_{\mathbf{s}}(\mathbf{z})$, with $\mathbf{s} = (s_1, \dots, s_d)$ then

$$\mathbb{E}\psi(\tilde{\boldsymbol{\eta}}^2) = \sum_{\substack{s_i = \pm \\ 1 \leq i \leq d}} \int_{(\mathbb{R}_+)^d} \psi\left(\{z_i^2\}_{i=1}^d\right) g_{\Gamma}(\phi_{\mathbf{s}}(\mathbf{z})) dz_1 \dots dz_d.$$

Taking $u_i = z_i^2$, by change of variable, we retrieve $\tilde{\boldsymbol{\eta}}^2$ density:

$$\mathbb{E}\psi(\tilde{\boldsymbol{\eta}}^2) = \sum_{\substack{s_i = \pm \\ 1 \leq i \leq d}} \int_{(\mathbb{R}_+)^d} \psi\left(\{u_i\}_{i=1}^d\right) g_{\Gamma}\left(\phi_{\mathbf{s}}\left(\{\sqrt{u_i}\}_{i=1}^d\right)\right) \left(\prod_{i=1}^d \frac{1}{2\sqrt{u_i}}\right) du_1 \dots du_d.$$

Denoting p the density of $\tilde{\boldsymbol{\eta}}^2$,

$$p(\mathbf{u}) = \sum_{\substack{s_i = \pm \\ 1 \leq i \leq d}} g_{\Gamma}\left(\phi_{\mathbf{s}}\left(\{\sqrt{u_i}\}_{i=1}^d\right)\right) \left(\prod_{i=1}^d \frac{1}{2\sqrt{u_i}} \mathbf{1}_{u_i > 0}\right).$$

Properties: p is continuous and positive on $(\mathbb{R}_+^*)^d$.

Let us derive the density of $\boldsymbol{\sigma}_t^2 | \boldsymbol{\sigma}_{t-1}^2 = \mathbf{x}$. By definition of $\boldsymbol{\sigma}_t^2$, it is equal to the density of $\{w_i + (\alpha_i \tilde{\eta}_i^2 + \beta_i)x_i\}_{i=1}^d$, for $\tilde{\boldsymbol{\eta}} \sim \mathcal{N}(0, \Gamma)$.

Denoting ϕ a test function, by the change of variable $v_i = w_i + (\alpha_i u_i + \beta_i)x_i$,

$$\begin{aligned} \mathbb{E}[\phi(\boldsymbol{\sigma}_t^2) | \boldsymbol{\sigma}_{t-1}^2 = \mathbf{x}] &= \int_{\mathbb{R}^d} \phi(\{w_i + (\alpha_i u_i + \beta_i)x_i\}_{i=1}^d) p(\mathbf{u}) du_1 \dots du_d \\ &= \int_{\mathbb{R}^d} \phi(\mathbf{v}) \underbrace{p\left(\left\{\frac{v_i - w_i - \beta_i x_i}{\alpha_i x_i}\right\}\right) \prod_{i=1}^d \frac{1}{\alpha_i x_i} \mathbf{1}_{(w_i + \beta_i x_i, +\infty)}(v_i)}_{g_{\mathbf{x}}(\mathbf{v})} dv_1 \dots dv_d. \end{aligned}$$

Hence

$$g_{\mathbf{x}}(\mathbf{v}) := p\left(\left\{\frac{v_i - w_i - \beta_i x_i}{\alpha_i x_i}\right\}\right) \prod_{i=1}^d \frac{1}{\alpha_i x_i} \mathbf{1}_{(w_i + \beta_i x_i, +\infty)}(v_i). \quad (\text{A.4})$$

A.5 Hill estimator

Second order condition

Denoting $U(y) = \bar{F}^{-1}(1/y)$, we say that U satisfies a second order condition if there exists $\gamma > 0$, $\rho \leq 0$ and A of constant sign, $\lim_{t \rightarrow \infty} A(t) = 0$ s.t.

$$\frac{1}{A(t)} \left[\frac{U(tx)}{U(t)} - x^\gamma \right] \xrightarrow{t \rightarrow \infty} x^\gamma \frac{x^\rho - 1}{\rho} \quad \forall x > 0.$$

Under second order condition, if additionally $\sqrt{k}A(n/k) \xrightarrow{n \rightarrow \infty} \lambda < \infty$, $\lim_{n \rightarrow \infty} \frac{k}{n} = \frac{k(n)}{n} = 0$ and $\lim_{n \rightarrow \infty} k = +\infty$, then the Hill estimator satisfies a Central Limit Theorem [DHF06, Theorem 3.2.5] :

$$\sqrt{k}(\hat{\gamma}_n - \gamma) \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{N}(\lambda/(1 - \rho), \gamma^2). \quad (\text{A.5})$$

Choice of k to minimize the asymptotic MSE

From [DHF06, equation 3.2.13 p.77], the asymptotic mean-squared error is given by

$$\text{asMSE} = \frac{\gamma^2}{k} + \frac{A^2(n/k)}{(1 - \rho)^2}. \quad (\text{A.6})$$

Fraga Alves et al [FAGdHN07] propose an explicit function A and constant ρ in the case of Fréchet attraction domain in their Example 4.3, for $\gamma \neq 1$:

$$A(t) = -\frac{1 - \gamma}{2t} \quad \text{and} \quad \rho = -1 \quad (\text{A.7})$$

Minimizing the asymptotic MSE for these functions, we get:

$$k_{opt} = 2 \left(\frac{\gamma}{1 - \gamma} n \right)^{\frac{2}{3}}. \quad (\text{A.8})$$

In particular, for $k = k_{opt}$, $\sqrt{k}A(n/k) = -\sqrt{2\gamma} < \infty$, so the Central Limit Theorem (A.5) holds.

Proof. Plugging (A.7) in the asymptotic MSE equation (A.6), we get:

$$\text{asMSE} = \frac{\gamma^2}{k} + \frac{1}{4} \left(\frac{1 - \gamma}{2n/k} \right)^2 = \frac{\gamma^2}{k} + \left(\frac{1 - \gamma}{4n} \right)^2 k^2.$$

Its derivative in k is:

$$\frac{\partial \text{asMSE}}{\partial k} = -\frac{\gamma^2}{k^2} + 2 \left(\frac{1 - \gamma}{4n} \right)^2 k = -\frac{\gamma^2}{k^2} + \frac{1}{2^3} \left(\frac{1 - \gamma}{n} \right)^2 k$$

which is positive for $k \geq k_{opt}$ and negative $k \leq k_{opt}$, where k_{opt} zero of $\frac{\partial \text{asMSE}}{\partial k}$ is given by (A.8). □

A.6 Additional properties

Extension of the Lyapunov condition to the lagged chain

This paragraph is dedicated to the extension of the Lyapunov condition $\mathbf{H}_{\mathcal{L}}$ to lagged Markov Chains, e.g. the fact that if we have a Lyapunov condition $\mathbf{H}_{\mathcal{L}}$ for $\{V_t\}_{t \geq 1}$ then we have $\mathbf{H}_{\mathcal{L}}$ for lagged chain $\{V_{tH}\}_{t \geq 1}$, $H \in \mathbb{N}^*$.

In Meyn and Tweedie's book, the statement that the extension is possible can be found in [MT09, Theorem 15.3.4] page 383.

First, let us notice that the notion of H -skeleton and H -lagged Markov chain coincide. Skeletons are defined page 62 in [MT09] as the chain $V^{(H)}$ with transition law:

$$\mathbb{P}(V_{tH+1} \in A | V_1 = x) = P^{tH}(x, A), \quad A \subset \mathcal{B}(\mathcal{S}).$$

Hence it is the same Markov chain than $\{V_{tH+1}\}_{t \in \mathbb{N}^*}$.

Let us recall Meyn and Tweedie's result:

Theorem A.6 (Extension of $H_{\mathcal{L}}$ to lagged processes [MT09, Theorem 15.3.4]). *Suppose that V is a Φ -irreducible and aperiodic. If V satisfied $H_{\mathcal{L}}$ with a petite set \mathcal{C} then for any H -skeleton, the function \mathcal{L} also satisfies $H_{\mathcal{L}}$ for some set \mathcal{C}' which is petite for the H -skeleton.*

The Meyn and Tweedie's result is a qualitative result. In what follows, we give a quantitative result on the extension of the Lyapunov condition to lagged Markov chains, in which the Lyapunov drift condition constants for the lagged process are explicit. In this way, we can quantify explicitly the dependence of the constant in the lag period H .

Proposition A.7 (Lyapunov condition for lagged Markov chain). *Assume that the Lyapunov drift criteria $H_{\mathcal{L}}$ is verified for $\{V_t, t \in \mathbb{N}^*\}$. Then, for any $H \in \mathbb{N}^*$, the H -lagged chain $\{V_{nH+1}, n \in \mathbb{N}^*\}$ also satisfies a drift condition:*

$$\mathbb{E}[\mathcal{L}(V_{H+1}) | V_1 = x] \leq d_H \mathcal{L}(x) + b_H \mathbf{1}_{x \in \mathcal{C}^{(H)}}.$$

where

$$d_H = \frac{1 + \delta^H}{2}, \quad \mathcal{C}^{(H)} = \mathcal{C} \cup \{x \in \mathcal{S} \mid |x| \leq R^{(H)}\},$$

$$b_H = \sup_{x \in \mathcal{C}^{(H)}} \left(b \frac{1 - \delta^H}{1 - \delta} - \frac{1 - \delta^H}{2} \mathcal{L}(x) \right),$$

and $R^{(H)} > 0$ s.t. for every $x \in \mathcal{S}$ s.t. $|x| > R^{(H)}$, $\frac{1 - \delta^H}{2} \mathcal{L}(x) \geq b \frac{1 - \delta^{H-1}}{1 - \delta}$.

Proof. Let $H \in \mathbb{N}^*$ be fixed. Let P denotes the transition kernel associated to $\{V_t\}$: using standard Markov chain notations, $\mathbb{E}[g(V_2) | V_1 = x] = Pg(x)$, $\mathbb{E}[g(V_{H+1}) | V_1 = x] = P^H g(x)$. We want to show that there is a $d_H \in (0, 1)$, a constant b_H and a set $\mathcal{C}^{(H)}$ s.t. for all $x \in \mathcal{S}$:

$$P^H \mathcal{L}(x) \leq d_H \mathcal{L}(x) + b_H \mathbf{1}_{\mathcal{C}^{(H)}}(x).$$

Let's apply recursively the initial drift criteria equation $H_{\mathcal{L}}$: $P \mathcal{L}(x) \leq \delta \mathcal{L}(x) + b \mathbf{1}_{\mathcal{C}}(x)$. We get, for $x \in \mathcal{S}$:

$$\begin{aligned} P^2 \mathcal{L}(x) &\leq P(\delta \mathcal{L}(x) + b \mathbf{1}_{\mathcal{C}}(x)) = \delta P \mathcal{L}(x) + bP(x, \mathcal{C}) \\ &\leq \delta^2 \mathcal{L}(x) + \delta b \mathbf{1}_{\mathcal{C}}(x) + bP(x, \mathcal{C}) \\ &\leq \delta^2 \mathcal{L}(x) + b + \delta b \mathbf{1}_{\mathcal{C}}(x), \\ P^3 \mathcal{L}(x) &\leq \delta^3 \mathcal{L}(x) + b + \delta b + \delta^2 b \mathbf{1}_{\mathcal{C}}(x), \\ &\vdots \\ P^H \mathcal{L}(x) &\leq \delta^H \mathcal{L}(x) + b \sum_{k=0}^{H-2} \delta^k + \delta^{H-1} b \mathbf{1}_{\mathcal{C}}(x) \\ &= \delta^H \mathcal{L}(x) + b \frac{1 - \delta^{H-1}}{1 - \delta} + \delta^{H-1} b \mathbf{1}_{\mathcal{C}}(x). \end{aligned} \tag{A.9}$$

Let $d_H = \frac{1+\delta^H}{2} \in (0, 1)$ since $\delta \in (0, 1)$. Let $R^{(H)} > 0$ s.t. for all $x \in \mathcal{S}$, $(|x| > R^{(H)}) \Rightarrow \left(b \frac{1-\delta^{H-1}}{1-\delta} \leq \frac{1-\delta^H}{2} \mathcal{L}(x)\right)$ (it exists since $\lim_{|x| \rightarrow \infty} \mathcal{L}(x) = +\infty$). Let $\mathcal{C}^{(H)} = \mathcal{C} \cup \{x \in \mathcal{S} \mid |x| \leq R^{(H)}\}$.

Then, outside $\mathcal{C}^{(H)}$, using the trick $\delta^H = \frac{1+\delta^H}{2} - \frac{1-\delta^H}{2}$, since $|x| > R^{(H)}$,

$$\begin{aligned} \delta^H \mathcal{L}(x) + b \frac{1-\delta^{H-1}}{1-\delta} &= \frac{1+\delta^H}{2} \mathcal{L}(x) + b \underbrace{\frac{1-\delta^{H-1}}{1-\delta} - \frac{1-\delta^H}{2}}_{\leq 0} \mathcal{L}(x) \\ &\leq \frac{1+\delta^H}{2} \mathcal{L}(x) := d_H \mathcal{L}(x). \end{aligned}$$

For $x \in \mathcal{C}^{(H)}$,

$$\begin{aligned} P^H \mathcal{L}(x) &\leq \delta^H \mathcal{L}(x) + b \frac{1-\delta^{H-1}}{1-\delta} + \delta^{H-1} b \\ &= d_H \mathcal{L}(x) + (\delta^H - d_H) \mathcal{L}(x) + b \frac{1-\delta^{H-1}}{1-\delta} + \delta^{H-1} b. \end{aligned}$$

Making the simplification: $b \frac{1-\delta^{H-1}}{1-\delta} + \delta^{H-1} b = b \frac{1-\delta^H}{1-\delta}$ and $\delta^H - d_H = \delta^H - \frac{1+\delta^H}{2} = -\frac{1-\delta^H}{2}$, denoting

$$b_H = \sup_{x \in \mathcal{C}^{(H)}} \left(b \frac{1-\delta^H}{1-\delta} - \frac{1-\delta^H}{2} \mathcal{L}(x) \right),$$

then

$$P^H \mathcal{L}(x) \leq d_H \mathcal{L}(x) + \left(b \frac{1-\delta^H}{1-\delta} - \frac{1-\delta^H}{2} \mathcal{L}(x) \right) \leq d_H \mathcal{L}(x) + b_H.$$

Thus the wanted inequality $P^H \mathcal{L}(x) \leq d_H \mathcal{L}(x) + b_H \mathbf{1}_{\mathcal{C}^{(H)}}(x)$ is satisfied. □

The following proposition can be found in [MT09, Proposition 5.4.5 - (iii)] page 114.

Proposition A.8 (Extension of \mathbf{H}_S to lagged processes). *Suppose that V is a Φ -irreducible Markov chain. If V is aperiodic then every skeleton is Φ -irreducible and aperiodic.*

Theorem A.9 (Burkholder's inequality [HH80, Theorem 2.10]). *Let $\{S_i, \mathcal{F}_i, 1 \leq i \leq n\}$ be a real-valued martingale, and $1 < p < \infty$. Denoting $X_1 = S_1$ and $X_i = S_i - S_{i-1}, 2 \leq i \leq n$, then there exist constants $c(p)$ and C_p^B depending only on p such that*

$$c(p) \mathbb{E} \left[\left| \sum_{i=1}^n X_i^2 \right|^{p/2} \right] \leq \mathbb{E} \left[\left| S_n \right|^p \right] \leq C_p^B \mathbb{E} \left[\left| \sum_{i=1}^n X_i^2 \right|^{p/2} \right] \quad (\text{A.10})$$

where suitable constants are given by $c(p)^{-1} = (18p^{1/2}q)^p$ and $C_p^B = (18pq^{1/2})^p$, where $p^{-1} + q^{-1} = 1$.

References

- [BLR06] L. Bauwens, S. Laurent, and J.V.K. Rombouts. Multivariate GARCH models: a survey. *J. Appl. Econometrics*, 21(1):79–109, 2006.
- [Bol90] T. Bollerslev. Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model. *The Review of Economics and Statistics*, 72(3):498–505, 1990.
- [BS00] J.F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Series in Operations Research. Springer-Verlag, New York, 2000.
- [Car01] A. Carol. *Market models : a guide to financial data analysis*. John Wiley & Sons, Ltd., Chichester, 2001.
- [CLM97] J.Y. Campbell, A.W. Lo, and A.C. MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, 1997.
- [Con01] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.
- [DHF06] L. De Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Series in Operations Research. Springer-Verlag, New York, 2006.
- [DNLW18] G. De Nard, O. Ledoit, and M. Wolf. Factor models for portfolio selection in large dimensions: the good, the better and the ugly. *Journal of Financial Econometrics*, 2018. nby033.
- [EB86] R.F. Engle and T. Bollerslev. Modelling the persistence of conditional variances. *Econometric Rev.*, 5(1):1–87, 1986. With comments and a reply by the authors.
- [EK18] N. El Karoui. Random matrices and high-dimensional statistics: beyond covariance matrices. In *Proceedings of the International Congress of Mathematicians*, volume 4, pages 2875–2894. World Scientific, 2018.
- [EP07] R.F. Engle and A.J. Patton. What good is a volatility model? In John Knight and Stephen Satchell, editors, *Forecasting Volatility in the Financial Markets (Third Edition)*, Quantitative Finance, pages 47–63. Butterworth-Heinemann, Oxford, third edition edition, 2007.
- [FAGdHN07] M.I. Fraga Alves, M.I. Gomes, L. de Haan, and C. Neves. A note on second order conditions in extreme value theory: linking general and heavy tail conditions. *REVSTAT Statistical Journal*, 5(3):285–304, 2007.
- [FM03] G. Fort and E. Moulines. Convergence of the Monte Carlo expectation maximization for curved exponential families. *Ann. Statist.*, 31(4):1220–1259, 2003.
- [FZ19] C. Francq and J.-M. Zakoïan. *GARCH models*. John Wiley & Sons, Ltd., Chichester, 2nd edition, 2019. Structure, statistical inference and financial applications.
- [Gui11] V. Guigues. Sensitivity analysis and calibration of the covariance matrix for stable portfolio selection. *Computational Optimization and Applications*, 48(3):553–579, 2011.

- [HH80] P. Hall and C.C. Heyde. *Martingale limit theory and its application*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980. Probability and Mathematical Statistics.
- [KZ07] R. Kan and G. Zhou. Optimal portfolio choice with parameter uncertainty. *The Journal of Financial and Quantitative Analysis*, 42(3):621–656, 2007.
- [Mar59] H.M. Markowitz. *Portfolio selection: Efficient diversification of investments*. Cowles Foundation for Research in Economics at Yale University, Monograph 16. John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London, 1959.
- [Mic89] R.O. Michaud. The Markowitz Optimization enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, 45(1):31–42, 1989.
- [MT09] S. Meyn and R.L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. With a prologue by Peter W. Glynn.
- [Tro12] J.A. Tropp. User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.*, 12(4):389–434, 2012.