A High-frequency Approach to Realized Risk Measures

Federico Gatta[1], Fabrizio Lillo[1,2], Piero Mazzarisi[3]

[1]Scuola Normale Superiore, [2]Università di Bologna,
[3]Università di Siena

federico.gatta@sns.it

SCUOLA
NORMALE
SUPERIORE

Summary

Introduction

Problem Setting and General Concept

Let's define:

- $S_j^{(t)}$ be the asset's log price on day t at the j-th minute of the Regular Trading Hours. $t \in 0, \cdots, T$ and $j \in \{0, \cdots, 390\}$.
- $Y^{(t)} = S_{390}^{(t)} - S_0^{(t)}$ be the daily log return;
- $F_{Y^{(t)}}(\cdot)$ the cumulative distribution function of $Y^{(t)}$ conditional on the information set at the beginning of the day.

Problem Setting and General Concept

Let's define:

- $S_j^{(t)}$ be the asset's log price on day t at the j-th minute of the Regular Trading Hours. $t \in 0, \cdots, T$ and $j \in \{0, \cdots, 390\}$.
- $Y^{(t)} = S_{390}^{(t)} - S_0^{(t)}$ be the daily log return;
- $F_{Y^{(t)}}(\cdot)$ the cumulative distribution function of $Y^{(t)}$ conditional on the information set at the beginning of the day.

Let θ be the target probability level. The main measures used to summarize financial risk are:

- **Value at Risk (VaR)** - It is defined as the θ quantile of the conditional distribution:

$$VaR_\theta(Y^{(t)}) = \inf\{F_{Y^{(t)}}^{-1}([\theta, 1])\}$$

  Noticeable levels θ for the VaR are $0.05, 0.025, 0.01$.

- **Expected Shortfall (ES)** - It is defined as the conditional tail expectation, that is:

$$ES_\theta(Y^{(t)}) = \mathbb{E}[Y^{(t)} | Y^{(t)} \leq VaR_\theta(Y^{(t)})]$$

  Noticeable levels θ for the ES are $0.05, 0.025$.

Forecasting Risk Measures- An Overview

So far, the literature has been oriented to **forecast** the risk measures at future timesteps:

- **Autoregressive** - Taylor (2019); Patton, Ziegel, and Chen (2019); Gatta, Lillo, and Mazzarisi (2024).
- **Neural networks** - Barrera et al. (2022).

Now, we change the perspective and try to **filter** the daily risk measures from the data.

Research question

How do we estimate VaR and ES directly from the data at a daily time scale? (Possibly accounting for the **stylized facts** of financial returns, see e.g. Cont (2001), such as fat tails, volatility clustering, etc.)

The literature related to the asset's volatility estimate provides the solution: leverage high-frequency (intra-day) data to estimate the low-frequency (daily) risk measures.

From Intra-day to Daily: Key Idea

Realized Risk Measures (RRMs)

The name *Realized* comes from the fact that we use **high-frequency** data to estimate the daily risk measures.
Big **difference** wrt Realized Variance: RRMs make a **distributional assumption** on the return process as we are interested in the tail behavior.

From Intra-day to Daily: Key Idea

Realized Risk Measures (RRMs)

The name *Realized* comes from the fact that we use **high-frequency** data to estimate the daily risk measures.
Big **difference** wrt Realized Variance: RRMs make a **distributional assumption** on the return process as we are interested in the tail behavior.

General Pipeline

- Focus on a **single day at once** (for simplicity remove the superscript $S_j^{(t)} \rightarrow S_j$).
- Observe the **raw intra-day log-returns**.
- Filter the intra-day data to **remove the microstructural noise**.
- Estimate the **intra-day risk** measures: $\hat{q}_j^{(t)}$ and $\hat{e}_j^{(t)}$.
- **Scale** the measures to daily frequency: $\hat{q}_D = q\_scale(\hat{q}_i)$; $\hat{e}_D = e\_scale(\hat{e}_i)$.

Motivation

Realized Quantiles - Dimitriadis and Halbleib (2022)

Only one paper in this direction: **Dimitriadis and Halbleib (2022)**.
Main assumption: the **subordinated** (i.e., properly sampled) log price process is **self-similar** with **stationary increments** (s.s.s.i.).

Subordinated Process

The idea is to measure the process in a particular time, also known as **intrinsic time**, where the observations should be more regular.

The core is the transformation of the clock time via a non-decreasing function known as **subordinator**:

$$\tau : [0, c] \to [0, 390] \qquad \textit{with} \quad \tau(0) = 0 \quad \tau(c) = 390$$

Values of c: 39, 78, and 130 (in clock time: 10, 5, and 3 minutes sampling).

The subordinated process is thus defined as $\{\tilde{S}_j\}_{j=0}^{c}$ with $\tilde{S}_j = S_{\tau(j)}$.

Dimitriadis and Halbleib (2022) - II

Self-similarity with Stationary increments

The process $\{\tilde{S}_j\}_{j=0}^{c}$ is said to be self-similar iff $\exists H \in (0,1)$ such that:

$$\tilde{S}_{j\Delta} \overset{d}{=} \Delta^H \tilde{S}_j$$

If the increments of a self-similar process are stationary, then

$$\tilde{S}_{j+\Delta} - \tilde{S}_{j-1} \overset{d}{=} \Delta^H \left(\tilde{S}_j - \tilde{S}_{j-1}\right)$$

Let define the intra-day log return as $Y_j = \tilde{S}_j - \tilde{S}_{j-1}$. Furthermore $Y = \sum_{j=1}^{c} Y_j$. Under the above assumptions:

$$\begin{aligned}
&\mathbb{E}\left[\tilde{S}_j - \tilde{S}_{j-1}\right] = 0 && \Longrightarrow \mathbb{E}[Y_j] = 0 \\
&VaR_\theta(\tilde{S}_{j+\Delta} - \tilde{S}_{j-1}) = \Delta^H VaR_\theta(\tilde{S}_j - \tilde{S}_{j-1}) && \Longrightarrow VaR_\theta(Y) = c^H VaR_\theta(Y_j) \\
&ES_\theta(\tilde{S}_{j+\Delta} - \tilde{S}_{j-1}) = \Delta^H ES_\theta(\tilde{S}_j - \tilde{S}_{j-1}) && \Longrightarrow ES_\theta(Y) = c^H ES_\theta(Y_j)
\end{aligned}$$

The empirical estimator is used for $VaR_\theta(Y_j)$ and $ES_\theta(Y_j)$!

Self-similarity Assumption's Breakdown - I

An s.s.s.i. process is **monofractal**.

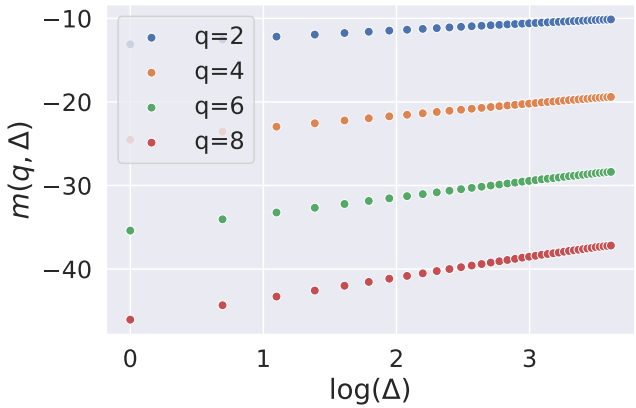
Monofractal process

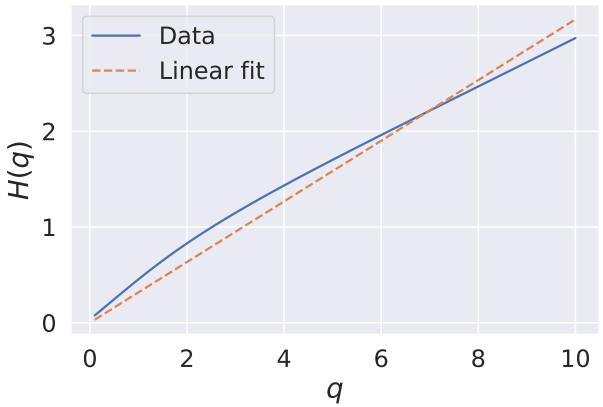
The process $\{\tilde{S}_j\}_{j=0}^{c}$ is said to be multifractal if $\exists H(q)$ such that:

$$m(q,\Delta) := \mathbb{E}\left[|\tilde{S}_{j+\Delta} - \tilde{S}_j|^q\right] = A(q)\Delta^{H(q)} \quad q,\Delta > 0 \tag{1}$$

If $H(q) = H\ q$, then the process is said to be monofractal.

To test this we study $\log(m(q,\Delta)) = \log(A(q)) + H(q)\log(\Delta)$.





Self-similarity Assumption's Breakdown - II

Empirically, the Ljung-Box test detects the presence of a small **autoregressive** component at the **first lag** of the log returns series. A self-similar process is not able to efficiently model it:

Embrechts and Maejima (2000)

Given an s.s.s.i. process with Hurst exponent H: $\cdot H = 0.5$ leads to independent increments $\cdot H > 0.5$ implies long-range dependence $\cdot H < 0.5$ used for noise, unfeasible for daily aggregation.

In no case is it possible to model the empirical correlation efficiently. DH set $H = 0.5$.

Self-similarity Assumption's Breakdown - II

Empirically, the Ljung-Box test detects the presence of a small **autoregressive** component at the **first lag** of the log returns series. A self-similar process is not able to efficiently model it:

Embrechts and Maejima (2000)

Given an s.s.s.i. process with Hurst exponent H: $\cdot H = 0.5$ leads to independent increments $\cdot H > 0.5$ implies long-range dependence $\cdot H < 0.5$ used for noise, unfeasible for daily aggregation.

In no case is it possible to model the empirical correlation efficiently. DH set $H = 0.5$.

An s.s.s.i. process has zero-mean increments. Nonetheless, Y_j mean is $\frac{Y}{c}$ that is not zero $\Longrightarrow$ **inconsistency** in the **scaling** procedure!

Realized Risk Measures

Our Proposal - Realized Risk Measures (RRM)

Issue

Removing the assumption of self-similarity makes harder to scale from high to low frequency! To solve this, we need other assumptions.

Assumptions

- The process for high-frequency log returns Y_j is stationary.
- The conditional probability distribution of $Y_j | Y_{j-1}$ is known.

Our Proposal - Realized Risk Measures (RRM)

Issue

Removing the assumption of self-similarity makes harder to scale from high to low frequency! To solve this, we need other assumptions.

Assumptions

- The process for high-frequency log returns Y_j is stationary.
- The conditional probability distribution of $Y_j | Y_{j-1}$ is known.

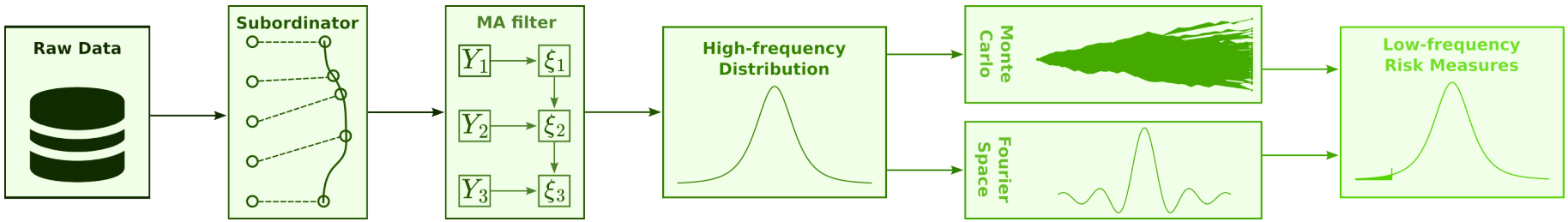
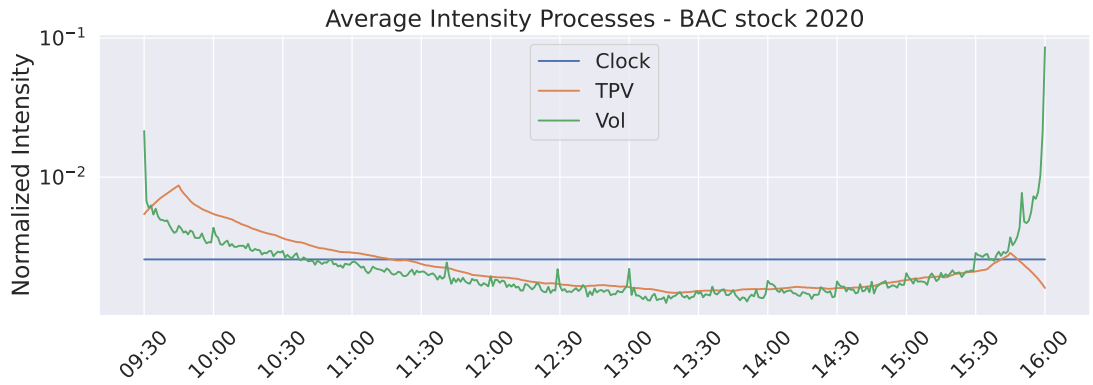


Figure 1: RRM Graphical abstract. First, the **subordinated returns** time series is obtained from raw data. Optionally, the **Moving Average (MA) filter** is used to extract the innovations series. Then, a **fat-tail distribution** is fitted to the high-frequency data. Lastly, the low-frequency distribution and risk measures are obtained via two parallel approaches based on **Monte-Carlo simulation** and **characteristic function**.

Data Preprocessing

The **subordinator** plays a key role in RRM. The idea is to aggregate minute-by-minute data over a coarse-grained grid in such a way that the activity of the market is the same in each interval. The following subordinators are tested:

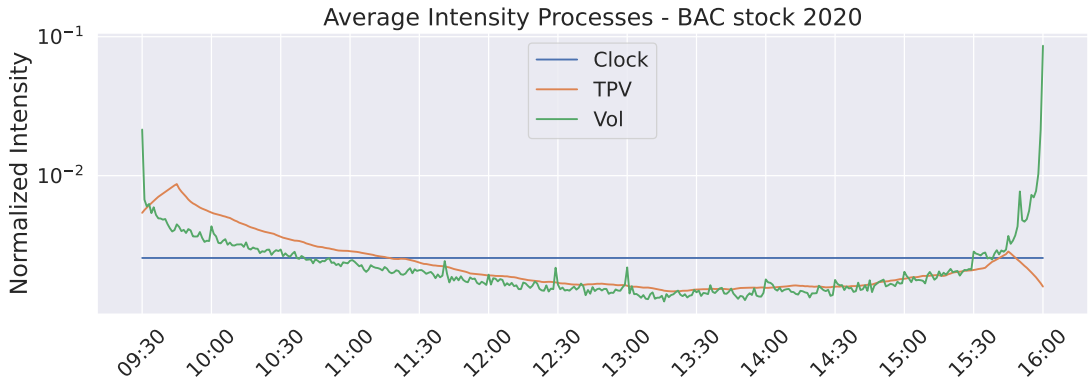
- **Clock** - naive subordinator based on the clock time, that is $\lambda_i = 1 \; \forall i$.
- **TPV** - based on the Tri-Power Variation.
- **Vol** uses the transactions' volume as the intensity process.



Data Preprocessing

The **subordinator** plays a key role in RRM. The idea is to aggregate minute-by-minute data over a coarse-grained grid in such a way that the activity of the market is the same in each interval. The following subordinators are tested:

- **Clock** - naive subordinator based on the clock time, that is $\lambda_i = 1 \; \forall i$.
- **TPV** - based on the Tri-Power Variation.
- **Vol** uses the transactions' volume as the intensity process.



Sometimes it makes sense to model the $\{Y_j\}_{j=1}^{c}$ data as an **MA process** of lag 1.
This choice is driven by: visual inspection of the ACF; and ease of handling.

$$Y_j = \phi\xi_{j-1} + \xi_j \quad \forall j = 1, \cdots, c\,, \quad \textit{with } \xi_j \textit{ iid}$$

The daily return can be written as: $Y = \sum_{j=1}^{c-1}(1+\phi)\xi_j + \phi\xi_0 + \xi_c$.
The innovations ξ_j replace Y_j in the following reasoning.

Our Proposal - Realized Risk Measures (RRM)

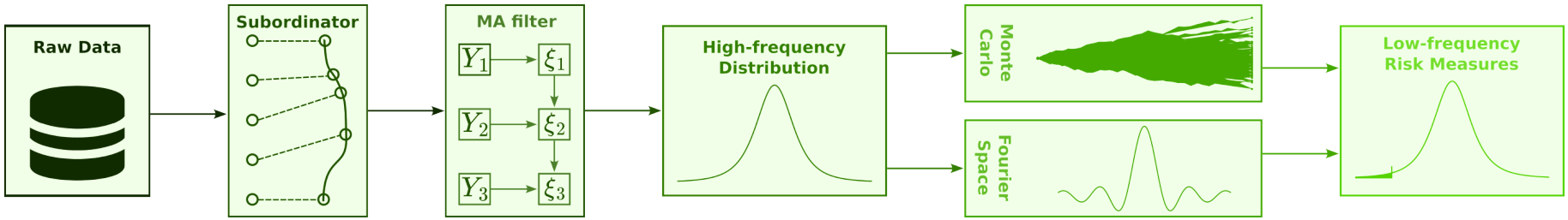


Figure 2: RRM Graphical abstract. First, the **subordinated returns** time series is obtained from raw data. Optionally, the **Moving Average (MA) filter** is used to extract the innovations series. Then, a **fat-tail distribution** is fitted to the high-frequency data. Lastly, the low-frequency distribution and risk measures are obtained via two parallel approaches based on **Monte-Carlo simulation** and **characteristic function**.

High-frequency Distribution (Hyperparameter)

Assumption on either the residuals $\{\xi_j\}$ (when applying the MA filter) or the returns time series $\{Y_j\}$. We use the **Student's t-distribution**: fat-tailed; few parameters; largely used in the literature.

Drift Bias

Fitting the high-frequency drift parameter μ from the high-frequency data would result in $\mu = \frac{Y}{c}$. Thus, we would be calculating the risk of a **low-frequency distribution centred on the observed value**.

Solution

Use an **a priori** assumption on μ to mitigate this risk:

- $\mu = 0$ reasonable for high-frequency returns.
- **Exponential Moving Average** (EMA): $\mu = EMA_{t+1}(\beta) = \frac{2}{\beta+1} \cdot Y^{(t)} + \left(1 - \frac{2}{\beta+1}\right) \cdot EMA_t(\beta)$, where β is the window length.

Our Proposal - Realized Risk Measures (RRM)

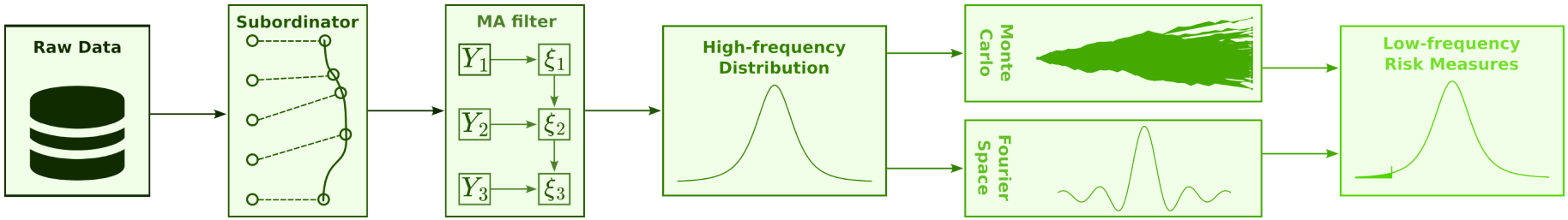


Figure 3: RRM Graphical abstract. First, the **subordinated returns** time series is obtained from raw data. Optionally, the **Moving Average (MA) filter** is used to extract the innovations series. Then, a **fat-tail distribution** is fitted to the high-frequency data. Lastly, the low-frequency distribution and risk measures are obtained via two parallel approaches based on **Monte-Carlo simulation** and **characteristic function**.

From High to Low Frequency: Characteristic Function

Recovering Daily Characteristic

Let φ_Y be the characteristic functions of the daily return. Assume Y_j to be i.i.d. (the MA case is similar) and let φ_{HF} be their characteristic function:

$$\varphi_Y(\omega) = \varphi_{HF}(\omega)^c$$

Gil-Pelaez Formula (1951)

The cumulative distribution function of Y can be recovered from its characteristic function: $F_Y(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty Im\left(\frac{\exp(-i\omega x)\varphi_Y(\omega)}{\omega}\right) d\omega$

From High to Low Frequency: Characteristic Function

Recovering Daily Characteristic

Let φ_Y be the characteristic functions of the daily return. Assume Y_j to be i.i.d. (the MA case is similar) and let φ_{HF} be their characteristic function:

$$\varphi_Y(\omega) = \varphi_{HF}(\omega)^c$$

Gil-Pelaez Formula (1951)

The cumulative distribution function of Y can be recovered from its characteristic function: $F_Y(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty Im\left(\frac{\exp(-i\omega x)\varphi_Y(\omega)}{\omega}\right) d\omega$

Some Remarks

Pro: **Not** prone to **sampling errors** (no randomness involved).

Cons: Rely on the numerical approximations of f_{HF} and the Gil-Pelaez Formula, which introduce **numerical errors** and could be **computationally heavy**.

From High to Low Frequency: Monte Carlo

An alternative way is by Monte-Carlo simulations (antithetic variates method for variance reduction). Main steps:

- Randomly draw a $B \times c$ matrix $\{z_{b,j}\}_{b=1\cdots B}^{j=1,\cdots,c}$. $B \in N$ is the batch size and each entry $z_{b,j}$ is distributed as the high-frequency observations.
- Sum the elements in each row and use the output vector of size B as a proxy for the Y distribution.
- VaR and ES are empirically computed from this.

Some Remarks

Pro: **Reduced numerical errors** and **computational time** as simpler operations are involved.

Cons: Vulnerable to an **approximation error** as the Monte-Carlo simulation replaces the true Y distribution.

From High to Low Frequency: Monte Carlo

An alternative way is by Monte-Carlo simulations (antithetic variates method for variance reduction). Main steps:

- Randomly draw a $B \times c$ matrix $\{z_{b,j}\}_{b=1\cdots B}^{j=1,\cdots,c}$. $B \in N$ is the batch size and each entry $z_{b,j}$ is distributed as the high-frequency observations.
- Sum the elements in each row and use the output vector of size B as a proxy for the Y distribution.
- VaR and ES are empirically computed from this.

Some Remarks

Pro: **Reduced numerical errors** and **computational time** as simpler operations are involved.

Cons: Vulnerable to an **approximation error** as the Monte-Carlo simulation replaces the true Y distribution.

Ensemble (mean)

As both the aggregation procedures have some weaknesses, we average their output. In such a way, we aim to obtain a more robust estimator.

Experimental Stage

The Experimental Framework

The experimental stage is carried out on synthetic and real-world datasets.

Separate experiments with three different numbers of intra-day returns: $c = 39, 78, 130$.

Regarding the target probability level, we consider $\theta = 0.05, 0.025, 0.01$ for VaR estimation and $\theta = 0.05, 0.025$ for ES, as standard in the current literature.

The Experimental Framework

The experimental stage is carried out on synthetic and real-world datasets.

Separate experiments with three different numbers of intra-day returns: $c = 39, 78, 130$.

Regarding the target probability level, we consider $\theta = 0.05, 0.025, 0.01$ for VaR estimation and $\theta = 0.05, 0.025$ for ES, as standard in the current literature.

We compare:

- **DH** - The approach in Dimitriadis and Halbleib (2022).
- **t-iid** - Our approach, using the Student's t-distribution to describe the high-frequency returns, which are assumed to be iid with 0 mean.
- **t-MA** - Our approach, where the Y_j are assumed to follow an MA(1) process whose returns are t-distributed with 0 mean.
- **t-iid (5;21)** and **t-MA 0** The same, but with drift = 0 for a fairer comparison with **DH**.

Synthetic Dataset

The **subordinated returns** are **directly generated** (no need for a subordinator). Generating process: $Y_j = \phi\xi_{j-1} + \sigma\xi_j$. Coefficients: (ϕ, σ) (and ξ_j distribution).

Four experiments: ξ_j normally distributed with $\phi = 0$ and $\phi \neq 0$;
ξ_j t-distributed with $\phi = 0$ and $\phi \neq 0$.
Coefficients fitted from real data. 40 years of simulations (approx. 10k days).

Evaluation: root mean squared error (**rMSE**) between the estimated and true risk measures (in the t-distribution case, the latter are computed via Monte Carlo simulation).

Synthetic Dataset

The **subordinated returns** are **directly generated** (no need for a subordinator). Generating process: $Y_j = \phi\xi_{j-1} + \sigma\xi_j$. Coefficients: (ϕ, σ) (and ξ_j distribution).

Four experiments: ξ_j normally distributed with $\phi = 0$ and $\phi \neq 0$;
ξ_j t-distributed with $\phi = 0$ and $\phi \neq 0$.
Coefficients fitted from real data. 40 years of simulations (approx. 10k days).

Evaluation: root mean squared error (**rMSE**) between the estimated and true risk measures (in the t-distribution case, the latter are computed via Monte Carlo simulation).

Results example: rMSE ($\times 10^3$) between the estimated and true VaR on the Gaussian dataset, iid returns -**DH** is well-specified.

Gaussian iid Returns - Synthetic Dataset (Value-at-Risk)									
Algorithm	$\theta = 0.05$			$\theta = 0.025$			$\theta = 0.01$		
	$c = 39$	$c = 78$	$c = 130$	$c = 39$	$c = 78$	$c = 130$	$c = 39$	$c = 78$	$c = 130$
DH	4.233	3.082	2.378	5.263	3.891	3.023	6.777	5.024	4.045
t-iid	**2.432**	**1.785**	**1.393**	**2.923**	**2.144**	**1.665**	**3.552**	**2.601**	**1.976**
t-MA	3.005	2.391	2.086	3.613	2.839	2.490	4.395	3.374	2.951
t-iid (5)	6.182	6.109	6.203	6.397	6.223	6.270	6.715	6.391	6.355
t-MA (5)	6.431	6.292	6.376	6.747	6.473	6.516	7.213	6.725	6.704
t-iid (21)	3.603	3.329	3.237	3.958	3.517	3.366	4.454	3.765	3.525
t-MA (21)	4.010	3.670	3.573	4.490	3.968	3.818	5.150	4.363	4.134

Synthetic Dataset - Further Insights

Gaussian iid Returns - Synthetic Dataset (Expected Shortfall)						
Algorithm	$\theta = 0.05$			$\theta = 0.025$		
	$c = 39$	$c = 78$	$c = 130$	$c = 39$	$c = 78$	$c = 130$
DH	5.188	3.723	2.882	6.790	4.843	3.826
t-iid	**3.097**	**2.286**	**1.774**	**3.591**	**2.650**	**2.000**
t-MA	3.808	2.927	2.554	4.441	3.328	2.904
t-iid (5)	6.478	6.271	6.298	6.734	6.410	6.364
t-MA (5)	6.857	6.509	6.540	7.245	6.699	6.682
t-iid (21)	4.093	3.580	3.411	4.490	3.770	3.534
t-MA (21)	4.652	4.028	3.854	5.195	4.324	4.094

Synthetic Dataset - Further Insights

Gaussian iid Returns - Synthetic Dataset (Expected Shortfall)						
Algorithm	$\theta = 0.05$			$\theta = 0.025$		
	$c = 39$	$c = 78$	$c = 130$	$c = 39$	$c = 78$	$c = 130$
DH	5.188	3.723	2.882	6.790	4.843	3.826
t-iid	**3.097**	**2.286**	**1.774**	**3.591**	**2.650**	**2.000**
t-MA	3.808	2.927	2.554	4.441	3.328	2.904
t-iid (5)	6.478	6.271	6.298	6.734	6.410	6.364
t-MA (5)	6.857	6.509	6.540	7.245	6.699	6.682
t-iid (21)	4.093	3.580	3.411	4.490	3.770	3.534
t-MA (21)	4.652	4.028	3.854	5.195	4.324	4.094

Performances:

- Improve as c increases (no microstructural noise).
- Worsen as θ becomes smaller (deep tail).

Statistically significant (Diebold-Mariano)!

Synthetic Dataset - Further Insights

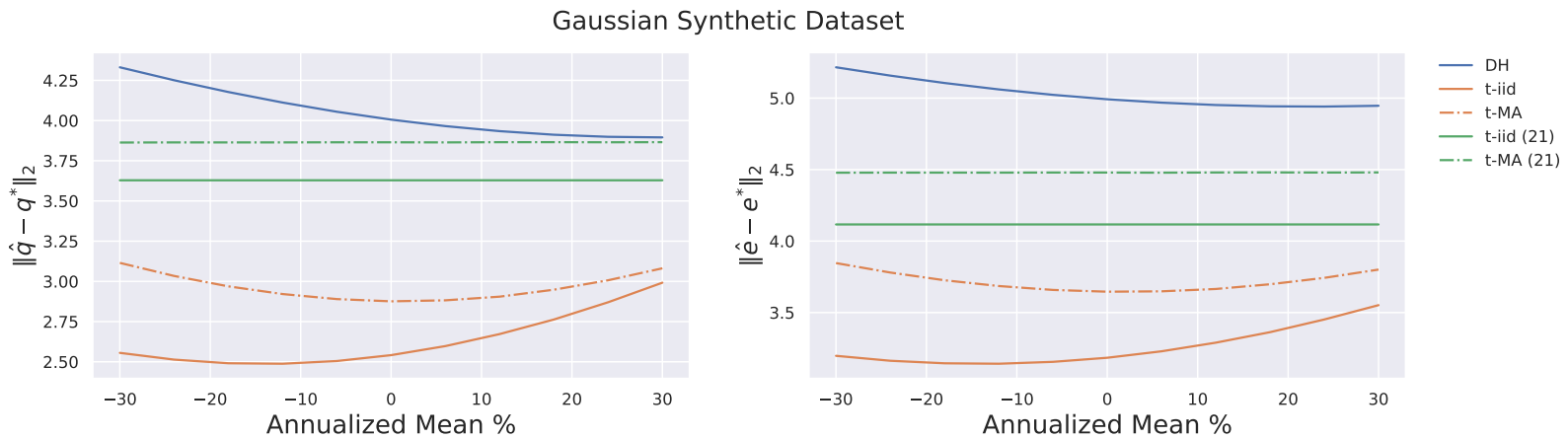
Gaussian iid Returns - Synthetic Dataset (Expected Shortfall)						
Algorithm	$\theta = 0.05$			$\theta = 0.025$		
	$c = 39$	$c = 78$	$c = 130$	$c = 39$	$c = 78$	$c = 130$
DH	5.188	3.723	2.882	6.790	4.843	3.826
t-iid	**3.097**	**2.286**	**1.774**	**3.591**	**2.650**	**2.000**
t-MA	3.808	2.927	2.554	4.441	3.328	2.904
t-iid (5)	6.478	6.271	6.298	6.734	6.410	6.364
t-MA (5)	6.857	6.509	6.540	7.245	6.699	6.682
t-iid (21)	4.093	3.580	3.411	4.490	3.770	3.534
t-MA (21)	4.652	4.028	3.854	5.195	4.324	4.094

Performances:

- Improve as c increases (no microstructural noise).
- Worsen as θ becomes smaller (deep tail).

Statistically significant (Diebold-Mariano)!

So far, the mean has been set to 0. How do the errors scale as a function of the mean? The left plot is for the VaR; the right is for the ES. $\theta = 0.05$ and $c = 39$.



Real-World Dataset

The real-world experiments are carried out on 18 US stocks.
The period spans from 1998 to 2020.

Approach 1: Measuring Statistical Properties (in-sample)

Let $\{(y^{(t)}, \hat{q}^{(t)}, \hat{e}^{(t)})\}_{t=1,\dots,T}$ be the time series of daily observations and estimated risk measures at a given probability level θ. Under the null hypothesis H_0 : *the filtering approach is well-specified*, the followings limits hold as $T \to \infty$.

Hits Frequency - VaR

$$\frac{1}{T}\sum_{t=1}^{T} \mathbb{1}_{\{y^{(t)} \leq \hat{q}^{(t)}\}} \to \theta$$

Acerbi and Szekely (2014) - (VaR, ES)

- $Z_1 := \dfrac{1}{|\{t | y^{(t)} \leq \hat{q}^{(t)}\}|} \displaystyle\sum_{t=1}^{T} \frac{y^{(t)}}{\hat{q}^{(t)}} \mathbb{1}_{\{y^{(t)} \leq \hat{q}^{(t)}\}} \to 1$
- $Z_2 := \dfrac{1}{T} \displaystyle\sum_{t=1}^{T} \frac{y^{(t)}}{\theta \hat{q}^{(t)}} \mathbb{1}_{\{y^{(t)} \leq \hat{q}^{(t)}\}} \to 1$

Real-World Results - Hits Frequency

Algorithm	$\theta = 0.05$			$\theta = 0.025$			$\theta = 0.01$		
	$c = 39$	$c = 78$	$c = 130$	$c = 39$	$c = 78$	$c = 130$	$c = 39$	$c = 78$	$c = 130$
DH (Clock)	0.034	0.033	0.032	0.013	0.011	0.010	0.004	0.002	0.003
DH (TPV)	0.031	0.030	0.028	0.016	0.013	0.010	0.007	0.004	0.003
DH (Vol)	0.030	0.027	0.025	0.013	0.010	0.008	0.005	0.003	0.002
t-iid (Clock)	0.036	0.035	0.032	0.014	0.013	0.013	0.004	0.004	0.004
t-iid (TPV)	0.038	0.035	0.026	0.016	0.014	0.011	0.004	0.004	0.003
t-iid (Vol)	0.034	0.029	0.021	0.013	0.012	0.008	0.004	0.003	0.002
t-MA (Clock)	0.038	0.037	0.036	0.014	0.014	0.015	0.003	0.004	0.004
t-MA (TPV)	0.040	0.038	0.030	0.016	0.016	0.013	0.004	0.004	0.004
t-MA (Vol)	0.035	0.031	0.022	0.013	0.012	0.009	0.003	0.003	0.003
t-iid (5; Clock)	0.056	0.052	**0.048**	**0.025**	0.024	0.022	0.008	0.008	0.008
t-iid (5; TPV)	0.056	**0.051**	0.037	0.026	**0.025**	0.018	**0.008**	0.008	0.006
t-iid (5; Vol)	**0.050**	0.041	0.027	0.023	0.019	0.012	0.007	0.006	0.004
t-MA (5; Clock)	0.059	0.056	0.054	0.026	0.026	**0.025**	0.007	0.008	**0.008**
t-MA (5; TPV)	0.059	0.057	0.042	0.027	0.027	0.020	0.008	**0.009**	0.007
t-MA (5; Vol)	0.051	0.044	0.028	0.023	0.020	0.013	0.006	0.007	0.004
t-iid (21; Clock)	0.042	0.039	0.037	0.017	0.016	0.015	0.005	0.004	0.004
t-iid (21; TPV)	0.043	0.040	0.029	0.018	0.017	0.013	0.005	0.005	0.004
t-iid (21; Vol)	0.038	0.032	0.021	0.015	0.013	0.009	0.004	0.004	0.003
t-MA (21; Clock)	0.044	0.043	0.041	0.017	0.017	0.017	0.004	0.005	0.005
t-MA (21; TPV)	0.046	0.044	0.033	0.019	0.018	0.014	0.005	0.005	0.005
t-MA (21; Vol)	0.039	0.034	0.023	0.016	0.014	0.010	0.004	0.004	0.003

Approach 2: Forecasting the Risk Measures (out-of-sample)

The idea is to use the RRM **out-of-sample**. In this way, we can use the **standard losses without** incurring the **bias**. The price to pay for this is:

Assumptions on Risk Measures Processes

Three separate tests are carried out, one for each of the following models for $\{\hat{q}^{(t)}\}_{t=1,\cdots,T}$ and $\{\hat{e}^{(t)}\}_{t=1,\cdots,T}$ time series:

· **AutoRegressive AR(1)** · **Random Walk** · **EMA**

Loss Functions for VaR and ES

Pinball loss for quantile estimate - Koenker and Bassett Jr (1978)

$$\mathcal{L}_q^\theta(\hat{q}, Y) := (Y - \hat{q})\left(\theta - \mathbb{1}_{\{Y \leq \hat{q}\}}\right)$$

Joint (VaR, ES) loss - Fissler and Ziegel (2016)

$$\mathcal{L}_e^\theta(\hat{q}, \hat{e}, Y) := \frac{\hat{q}}{\hat{e}} - \frac{\hat{q} - Y}{\theta \hat{e}} \mathbb{1}_{\{Y \leq \hat{q}\}} + log(-\hat{e})$$

The results are coherent with the previous findings.

Conclusion

Findings Summary

The take-away message of this talk can be summarized as follows.

- Filtering risk measures provides valuable insights for **evaluating predictive models** and **improving** their **performance** as well.
- We tackle this task by using **high-frequency** (intra-day) data.
- The pre-processing step is carried out by means of **subordinators** and an **MA filter**.
- The **scaling** from high to low frequency is done via an ensemble made up of two approaches. One is based on the **characteristic function** and the other on **Monte Carlo** simulation.
- Assessing the quality of filtered risk measures is trickier than forecasting problems. We exploit three strategies:
  - **Synthetic Datasets**
  - **Empirical Datasets**
    - **Statistical tests** In-sample
    - **Loss functions** (with additional assumptions) Out-of-sample

Next Steps

As for future research, there are two main directions.

- Study the effect of including **jumps** in the **high-frequency** dynamic. Theoretically, one could expect an improvement in the estimated low-frequency measures. However, empirical works in the literature (Clements, Galvão, and Kim (2008); Žikeš and Baruník (2015)) raise some doubts. So, we have not included jumps in the current work, and we want to deepen this analysis in the future.

- Later, we aim to use the **RRM** as an additional **feature** for risk **forecasting**. Indeed, we expect this could enhance state-of-the-art approaches as the RRM contains high-frequency information that traditional forecasters neglect.

- Study the tail-dependencies between assets - common movements direction and networks.

References I

Acerbi, Carlo and Balazs Szekely (2014). "Back-testing expected shortfall". In: *Risk* 27.11, pp. 76–81.

Barrera, D et al. (2022). "Statistical Learning of Value-at-Risk and Expected Shortfall". In: *arXiv preprint arXiv:2209.06476*.

Clements, Michael P, Ana Beatriz Galvão, and Jae H Kim (2008). "Quantile forecasts of daily exchange rate returns from forecasts of realized volatility". In: *Journal of Empirical Finance* 15.4, pp. 729–750.

Cont, Rama (2001). "Empirical properties of asset returns: stylized facts and statistical issues". In: *Quantitative Finance* 1.2, p. 223.

Dimitriadis, Timo and Roxana Halbleib (2022). "Realized quantiles". In: *Journal of Business & Economic Statistics* 40.3, pp. 1346–1361.

Embrechts, Paul and Makoto Maejima (2000). "An introduction to the theory of self-similar stochastic processes". In: *International journal of modern physics B* 14.12n13, pp. 1399–1420.

References II

Fissler, Tobias and Johanna F Ziegel (2016). "Higher order elicitability and Osband's principle". In: *The Annals of Statistics* 44.4, pp. 1680–1707.

Gatta, Federico, Fabrizio Lillo, and Piero Mazzarisi (2024). "CAESar: Conditional Autoregressive Expected Shortfall". In: *arXiv preprint arXiv:2407.06619*.

Koenker, Roger and Gilbert Bassett Jr (1978). "Regression quantiles". In: *Econometrica: journal of the Econometric Society*, pp. 33–50.

Patton, Andrew J, Johanna F Ziegel, and Rui Chen (2019). "Dynamic semiparametric models for expected shortfall (and value-at-risk)". In: *Journal of econometrics* 211.2, pp. 388–413.

Taylor, James W (2019). "Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution". In: *Journal of Business & Economic Statistics* 37.1, pp. 121–133.

Žikeš, Filip and Jozef Baruník (2015). "Semi-parametric conditional quantile models for financial returns and realized volatility". In: *Journal of Financial Econometrics* 14.1, pp. 185–226.

Thanks for your attention!