Information Measures for Long-Range Correlated Time Series

Pietro Murialdo

Politecnico di Torino

Outline

Much effort has been spent on the study of complex interactions in financial markets by means of information measures. Such theoretical measures origin in statistical physics and later found broad application in other fields where several heterogeneous interactions give rise to phenomena that cannot be studied according to linear dynamics. Socio-economic complex systems are characterized by patterns emerging from their seemingly random structure. These emergent dynamics are the results of the information embedded in the patterns, whose quantification can shed lights on the fundamental phenomena that origin them. An information measure, the Shannon entropy, $S(\tau)$, was proposed with the aim of quantifying the degree of uncertainty of strings of elementary random events in terms of their probabilities. Therefore, given a string of length N, a particular realization of the string can be described in terms of its length, τ, $\tau \ll N$, and can be associated to a probability, $p_i(\tau)$, which for stationary processes does not depend on its location in the sequence. The Shannon entropy is then calculated in terms of the expected value of all possible realizations, $S(\tau) = \sum_i p_i(\tau) \log p_i(\tau)$.

Long-range correlation of financial time series is studied with the objective of quantifying characteristic features emerging from the complex interactions underlying price formation. Financial markets and artificial time series were used to perform the analysis. More specifically, tick-by-tick prices of NASDAQ, DJIA and S&P500 from Jan 1^{st} to Dec 31^{st} 2018 and widely adopted artificial stochastic processes, such as FBM, ARFIMA and GARCH, were employed. The results are obtained with a *Cluster Entropy* method, which is used to quantify the following features:

Dynamics: price series have been investigated over consecutive temporal horizon to reveal a systematic dependence of the asset on the time horizon. To this purpose, given a temporal horizon, M, a *Market Dynamic Index*, $I_p(M)$, has been developed to quantify in a single figure the dynamics of asset prices over varying temporal horizons.

Heterogeneity: volatility series of asset prices have been investigated over consecutive temporal horizons to understand differences between different assets during time. In this case, given a temporal horizon, M and a volatility window, T, a *Market Heterogeneity Index*, $I_v(M)$, has been defined to quantify the heterogeneity of different assets and to yield the weights of a multi-period portfolio, as a complement to Markowitz and Sharpe traditional approaches.

Cluster Entropy Method

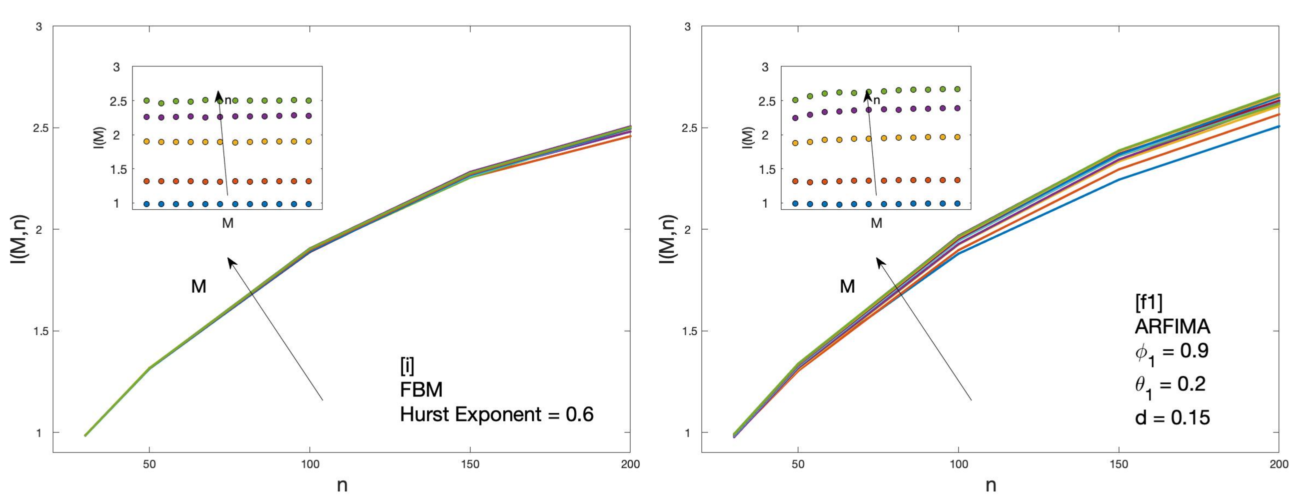
A financial time series, $y(t)$, is partitioned in clusters according to the intersections between the time series, $y(t)$, itself and a moving average time series, $\tilde{y}_n(t) = \frac{1}{n}\sum_{k=0}^{k=n-1} y(t-k)$, with n the moving average window.

Consecutive intersections yield a partition of the phase space into a series of *clusters*. Each cluster is defined as the portion of the time series, $y(t)$, between two consecutive intersections of $y(t)$ and $\tilde{y}_n(t)$, and has length (duration), τ, equal to $\tau \equiv ||t_j - t_{j-1}||$, where t_j and t_{j-1} uniquely identify two consecutive intersections.

For each moving average window, n, a probability mass function, $P(\tau, n)$, which associates the cluster duration to its frequency is obtained. The resulting probability mass function takes the following form: $P(\tau, n) \sim \tau^{-D}\mathcal{F}(\tau, n)$, where D indicates the fractal dimension, which can also be expressed in terms of the Hurst Exponent as $D = 2 - H$. Since the Hurst exponent ranges between $0 < H < 1$, the fractal dimension ranges between $1 < D < 2$. The fractal nature of the probability distribution implies that clusters are organized in a self-similar way along the time series, resulting in long-range correlation between clusters far away from each other. The term $\mathcal{F}(\tau, n)$ takes the form of $e^{\tau/n}$ to account for the drop-off of the power-law behavior and the onset of the exponential decay when $\tau \leq n$, due to the finiteness of n.

The Shannon entropy evaluated with the probability distribution $P(\tau, n)$ takes the form $S(\tau, n) = S_0 + \log \tau^D + \tau/n$, where the first term on the right is a constant, and the second and third term arise respectively for power-law and exponentially correlated clusters.

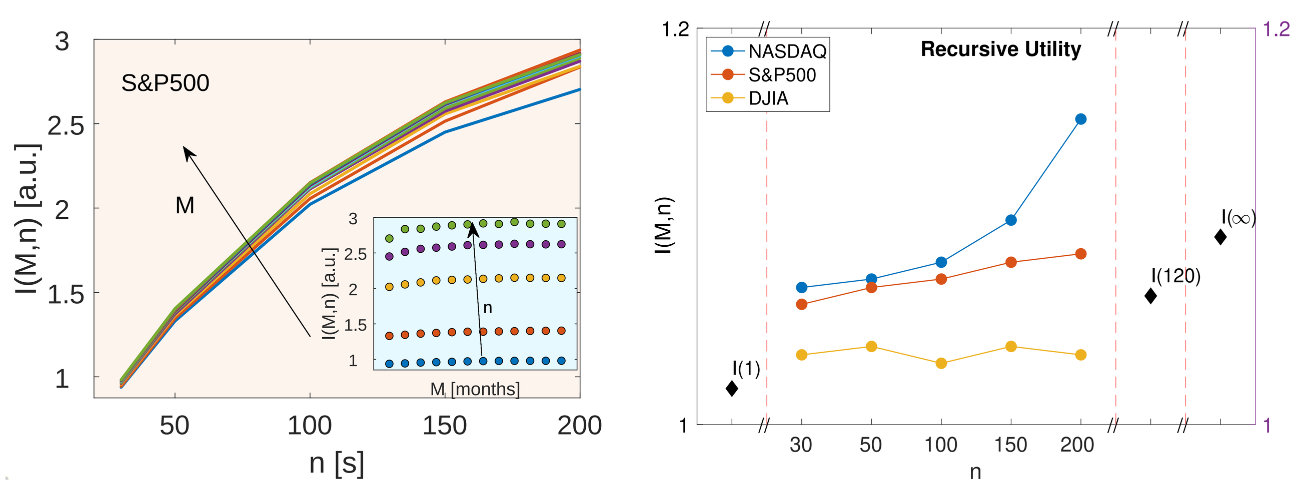
Real-World and Artificial Markets



* Widely adopted stochastic processes, as Fractional Brownian Motion (FBM), Auto-Regressive Fractionally Integrated Moving Average (ARFIMA), Generalized Auto-Regressive Conditional Heteroskedastic (GARCH) processes, have been used to generate an extensive number of time series with a broad range of values of the Hurst exponent H and of the parameters ϕ, θ, d.
* To quantify in a single number the horizon dependence and long-range correlation, the Market Dynamic Index, defined as $I(M, n) = \sum_\tau S(\tau, n)$, is used.
* The analysis via cluster entropy, $S(\tau, n)$, of the above mentioned artificial series has evidenced significant market and horizon dependence in the presence of long-range correlation. Hence, the characteristic horizon dependence exhibited by the cluster entropy is related to long-range positive correlation.

[1] Murialdo, P., Ponta, L., and Carbone. A., "Long-range dependence in financial markets: A moving average cluster entropy approach." *Entropy* 22, no. 6 (2020): 634.

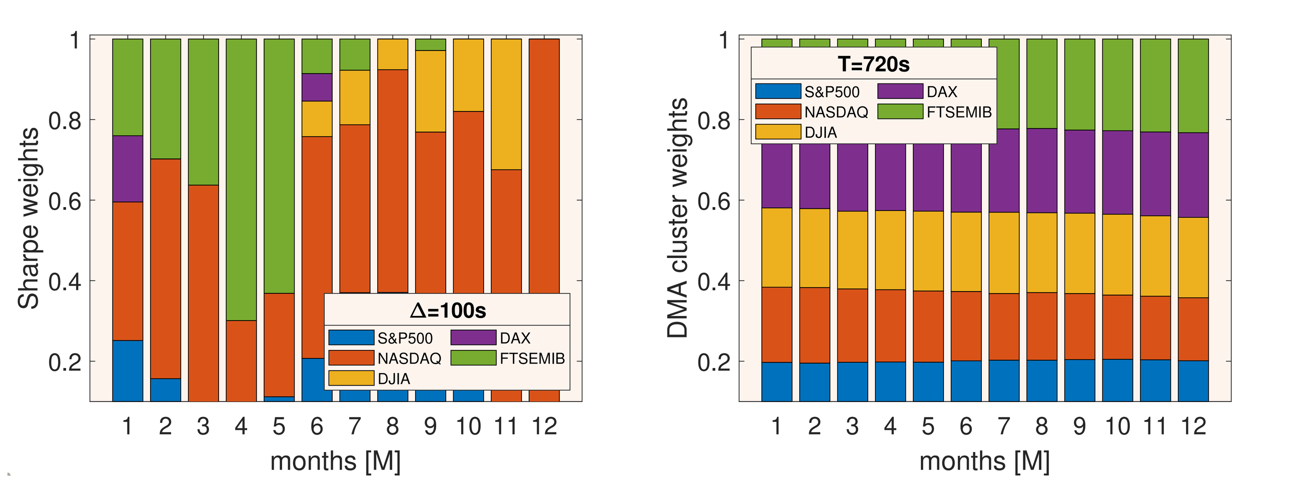
Real-World and Representative Agent Models



* The analysis has been extended to quantify market price heterogeneity and dynamics over different temporal horizons in real world assets, tick-by-tick prices of NASDAQ, DJIA and S&P500 from Jan 1^{st} to Dec 31^{st} 2018 with length $N = 6982017$, $N = 5749145$ and $N = 6142443$, respectively, via the cluster entropy method.
* A direct comparison is made between the horizon dependence obtained respectively by the moving average cluster entropy and the Kullback–Leibler relative entropy with representative agent models of price evolution.
* The comparison with the Kullback–Leibler entropy showed that the approaches yield consistent results. In particular, the Market Dynamic Index, $I(M, n)$, with recursive utility more faithfully reproduces the dynamics of real world data compared with difference habit and power utility representative agent models, which respectively under-estimate and over-estimate the increasing trends of the entropy production rate.

[2] Ponta, L., Murialdo, P., and Carbone A., "Information measure for long-range correlated time series: Quantifying horizon dependence in financial markets." *Physica A: Statistical Mechanics and its Applications* 570 (2021): 125777.

Multi-Period Portfolios



* On the account of the above theoretical findings, an approach to optimal portfolio estimation based on the cluster entropy approach is put forward.
* To obtain the portfolio weights for volatility series of each market, i, the Market Dynamic Index, $I(M, n)$, is averaged over the moving average windows, n, resulting in a cumulative figure of diversity, $I(M) = \sum_n I(M, n)$. Given a volatility window T, and the number of assets $\mathcal{N}$, the portfolio weights, are equal to $w_{i,T} = I_i(M)/\sum_k^{\mathcal{N}} I_k(M)$, for each market i.
* Contrary to the Sharpe theory, the cluster entropy approach provides stable and diverse portfolio weights, as can be seen in the figures. At short volatility windows T, weights take values close to the equally distributed $1/\mathcal{N}$ portfolio, consistently with investment strategies where volatility (risk) does not give a prominent contribution. As T increases, volatility plays an increasing role in the weights estimation, which deviate from the uniform distribution more significantly.

[3]Murialdo, P., Ponta, L., and Carbone. A., "Inferring Multi-Period Optimal Portfolios via Detrending Moving Average Cluster Entropy." *EPL* 133 (2021): 60004.