

# MAXIMUM-ENTROPY SCATTERING MODELS FOR FINANCIAL TIME SERIES

*Roberto Leonarduzzi<sup>1</sup>, Stéphane Mallat<sup>1,2</sup>, Jean-Phillipe Bouchaud<sup>3</sup> and Gaspar Rochette<sup>1</sup>*

<sup>1</sup> École Normale Supérieure, PSL, 75005, Paris, France

<sup>2</sup> Collège de France, France

<sup>3</sup> Science & Finance, Capital Fund Management, 75009, Paris, France

## ABSTRACT

Modeling time series with complex statistical properties such as heavy-tails, long-range dependence, and temporal asymmetries remains an open problem. In particular, financial time series exhibit such properties, and existing models suffer from serious limitations and often rely on high-order moments. We introduce a wavelet-based maximum entropy model for such random processes, based on novel scattering and phase-harmonic moments. We analyze the model's performance with a synthetic multifractal random process and real-world financial time series. We show that scattering moments capture heavy tails and multifractal properties without estimating high-order moments. Further, we show that additional phase-harmonic terms capture temporal asymmetries.

*Index Terms*— Maximum entropy models, scattering transform, wavelets, financial time series

## 1. INTRODUCTION

**Stochastic process modeling.** In a wide range of domains, such as internet traffic [1], turbulence [2] and medicine [3], data are characterized by complex statistical properties such as heavy-tailed distributions, long-range correlations, intermittent time evolutions, and temporal asymmetries. Despite the prevalence of such data, state-of-the-art stochastic models often struggle to represent all these types of behavior.

**Financial time series.** A case study of the above situation is provided by financial time series, consisting in the evolution of prices of different assets over time. The development of accurate models for such time series is a crucial task in finance, for applications such as the prediction of crises or the understanding of the mechanisms that control the dynamics of financial markets. However, financial time series are characterized by the complex statistical properties described above, which current models fail to characterize completely and efficiently [4, 5].

**Financial models.** Initial gaussian models based on brownian motions were unable to capture neither heavy-tails nor the complex temporal dependence that alternates periods of large

and small changes (known as “volatility clustering” [5], see Fig. 1).

The family of GARCH models was introduced to address these limitations [6, 7] through autoregressive components of volatility that capture the temporal dependence. Despite many improvements and variations [8], GARCH models still suffer from issues such as difficulties to represent the interactions between price changes at different time scales, or to capture heavy-tailed distributions [5].

Multifractal models were proposed to address these limitations of the GARCH family [5, 9]. They are based on prescribing the way in which all moments evolve at different time scales. Interscale relationships are naturally captured, and the imposition of high-order moments leads to heavy-tailed distributions and volatility clustering [5, 10, 11]. However, multifractal estimators are based on high-order moments and have a large variance. Also, they are not able to capture temporal asymmetries.

**Modeling challenges.** In light of this, the challenge in building a model for these complex types of data is to include the following features:

1. Estimators with small variance,
2. Interactions between scales,
3. Intermittent temporal structure,
4. Temporal asymmetries.

**Goals and contributions.** In this paper, we introduce a maximum entropy model with the features described above.

We use second-order moments of nonlinear contractive representations to bound the variance of estimators. Further, we capture interscale relationships through wavelet transforms, which provide a separation of scales. We use scattering coefficients to capture the complex, intermittent temporal structure related to heavy tails and multifractal properties. Finally, we introduce wavelet phase harmonics to measure temporal asymmetries and multiscale phase relationships.

Our results show that our model correctly captures the heavy tails, multifractal properties and temporal asymmetries of synthetic multifractal random processes and financial time series, without relying on high-order moments.

## 2. STATISTICAL PROPERTIES OF FINANCIAL TIME SERIES

**Notation.** Throughout the paper,  $X$  denotes a 1D random process, and  $x$  a realization of  $X$  with  $d$  samples. Further,  $\langle x(t) \rangle = d^{-1} \sum_t x(t)$  denotes the empirical average.

**Returns time series.** Let  $c(t)$  be the price of an asset at time  $t$ . The absolute returns at scale  $\delta$  are defined as  $r(t) = c(t) - c(t - \delta)$ . Here we will make use of the daily returns (i.e.  $\delta = 1$  day) of the S&P 500 index between the years 2000 and 2018, illustrated in Fig. 1. Returns are characterized by complex statistical properties including heavy tails, multifractal structure, and temporal asymmetries [4, 5].

**Heavy tails.** It has been empirically shown that the probability distribution of returns  $r$  is nongaussian and decays as a power law when  $\delta$  is not too large:

$$P(r) \sim r^{-\alpha}, \quad (1)$$

where  $3 \leq \alpha \leq 5$  [4, 12]. Heavy tails mean that the probability of observing extreme returns is higher than what is predicted by a normal distribution.

**Multifractal structure.** Empirical analyses also show that returns are scale invariant, have nontrivial high-order statistics, and are everywhere irregular with intermittent changes of regularity [4, 5, 11]. Multifractal analysis summarizes these properties through the so-called *scaling function*  $\zeta(q)$ , defined through the relation

$$\langle |x(t+a) - x(t)|^q \rangle \sim a^{\zeta(q)}, \quad q \in \mathbb{R}, \quad (2)$$

The function  $\zeta(q)$  quantifies how the moments of different orders evolve with the scales, and is related to heavy tails, long-range correlations and volatility clustering [5, 9, 13]. In practice, the increments in (2) are replaced by more robust quantities such as wavelet leaders [14].

**Multifractal random walk.** Multifractal random walk (MRW) [9, 15] has been proposed to model financial time series. It is defined as  $X(t) = B(t)e^{\omega_\lambda(t)}$ , where  $B(t)$  is a brownian motion and  $\omega_\lambda(t)$  is a gaussian process with covariance  $\text{cov}(w(t_1), w(t_2)) = \lambda^2 \log(T|t_1 - t_2| + 1)$  if  $|t_1 - t_2| \leq T$ , and  $\text{cov}(w(t_1), w(t_2)) = 0$  otherwise. MRW captures properties such as heavy tails and multifractal structure; its scaling function is  $\zeta(q) = (\lambda^2 + 1/2)q - \lambda^2 q^2/2$ . A sample realization of MRW is shown in Fig. 2, and its statistical properties are shown in Fig. 3 (blue lines).

**Temporal asymmetries.** Temporal asymmetries in financial time series arise from causality relationships, whereby economic actors anticipate events in the future using information from the past. One measure for temporal asymmetries is the leverage effect, which reflects the tendency for increased volatility after decreases in price [16]. It can be defined as the correlation between price change and a measure of the square volatility [16]:

$$L(\tau) = \frac{\langle r^2(t+\tau)r(t) \rangle}{\langle r_\delta^2(t) \rangle^2}. \quad (3)$$

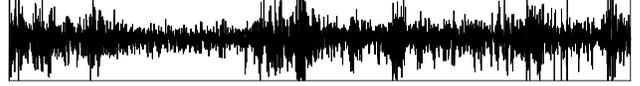


Fig. 1. S&P 500 daily absolute returns time series.

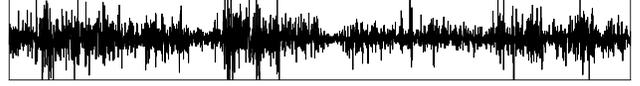


Fig. 2. Single realization of multifractal random walk.

The leverage  $L(\tau)$  has been found to be 0 for  $\tau < 0$ , and a negative exponential for  $\tau > 0$  [16].

**Goal.** The goal of this contribution is to build models of stochastic processes that capture properties (1), (2) and (3), while avoiding the pitfalls of the estimation of high-order moments.

## 3. MAXIMUM ENTROPY MODELS

**Maximum entropy microcanonical models.** We build maximum entropy models for  $X$  that are conditioned on their empirical moments. We use an informative representation  $\{U_\ell x : \ell \in \mathcal{L}\}$  that captures the properties of  $X$ . We assume that the empirical moments  $\langle U_\ell x \rangle$  concentrate with high probability around their expected values  $\mathbb{E}(U_\ell X)$  when the number of samples  $d$  is large enough. Given a fixed realization  $x$ , a microcanonical model will estimate a probability distribution  $P$  such that samples  $\tilde{x}$  drawn from  $P$  satisfy  $\langle U_\ell \tilde{x} \rangle \approx \langle U_\ell x \rangle \forall \ell \in \mathcal{L}$ . Under mild assumptions, the maximum entropy distribution  $P$  is uniform [17].

**Choice of representation.** Their key step for the success of a maximum entropy models is to choose a representation  $\{U_\ell\}_\ell$  that captures the properties of  $X$ , and that can be computed robustly. To keep the variance of estimates under control, we measure the second-order variability of nonlinear representations that are Lipschitz continuous, i.e. such that

$$\|(U_\ell x - U_\ell x')\| \leq K \|x - x'\|. \quad (4)$$

The Lipschitz condition (4) ensures that the variance of second-order moments of  $U_\ell x$  is bounded. In Sections 4 and 5 we propose explicit wavelet-based representations  $U_\ell$  that capture different aspects of  $X$ .

**Loss function.** Let  $\mu_\ell = \langle U_\ell x \rangle$  be the empirical mean of  $U_\ell$ , and let  $\text{cov}(U_\ell x, U_{\ell'} x) = \langle (U_\ell x - \mu_\ell, U_{\ell'} x - \mu_{\ell'}) \rangle$  be the empirical covariance with respect to mean  $\mu_\ell$ . The error between a generated sample  $\tilde{x}$  and the target  $x$  is measured by a the loss function

$$\mathcal{E}(x, \tilde{x}) = \sum_{(\ell, \ell') \in \mathcal{C}} (\text{cov}(U_\ell x, U_{\ell'} x) - \text{cov}(U_\ell \tilde{x}, U_{\ell'} \tilde{x}))^2 \quad (5)$$

The set  $\mathcal{C}$  contains the indices of the informative elements of the covariance that are used.

**Gradient descent algorithm.** Samples  $\tilde{x}$  can be efficiently drawn from the microcanonical model with a gradient descent algorithm [17]. From an initial white Gaussian noise  $x_0$ , the algorithm computes approximations  $x_n$  through the iteration

$$x_n = x_{n-1} - \alpha \nabla \mathcal{E}(x, x_{n-1}). \quad (6)$$

The loss  $\mathcal{E}$  ensures that  $\langle U_\ell x_n \rangle \rightarrow \langle U_\ell x \rangle$ . The convergence properties of this algorithm are studied in [17].

## 4. SCATTERING MOMENTS

### 4.1. Scattering representation

**Wavelet transform.** To analyze long-range-dependent processes such as financial time series, it is crucial to separate the variability at different scales. To that end, we use a wavelet transform. Let  $\psi$  be a mother wavelet, a band-pass filter with  $\int \psi(t) dt = 0$  that is well localized in both time and frequency. A dyadic wavelet filter bank is obtained by scaling  $\psi$  at scales  $2^j$ :  $\psi_j(t) = 2^{-j} \psi(2^{-j}t)$  for  $1 \leq j < J$ . Low-frequency information not captured by wavelets is recovered by a low-pass filter  $\psi_J$  at scale  $2^J$ . Wavelet coefficients of  $x$  are obtained through convolutions  $x \star \psi_j$  [18].

**Scattering moments.** Nongaussian processes such as time series are characterized by well-localized sharp transitions (see Figs. 1 and 2) that yield sparse wavelet representations. These are captured by first-order coefficients  $\langle |x \star \psi_{j_1}| \star \psi_{j_2} \rangle \approx \langle |x \star \psi_{j_1}| \rangle$ .

First-order moments fail to capture the complex temporal evolution of wavelet coefficients, which is lost in the average. This information is recovered by second-order moments that measure the variability of an iterated wavelet transform  $\langle |x \star \psi_{j_1}| \star \psi_{j_2} \rangle$  [18, 19]. Thus, scattering moments are

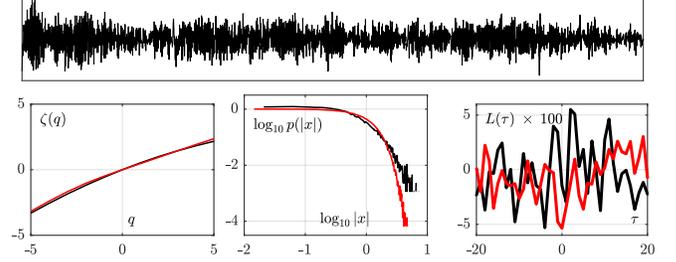
$$U_\ell^S x = |x \star \psi_{j_1}| \star \psi_{j_2} \quad \text{with } \ell = (j_1, j_2), \quad (7)$$

where  $1 \leq j_1 < j_2 \leq J$ . Note that  $U_\ell^S$  satisfies (4) with  $K = 1$  [18]. Only diagonal covariances  $\text{cov}(U_\ell^S x, U_{\ell'}^S x)$  are informative, and thus  $\mathcal{C}^S = \{\ell, \ell' : \ell = \ell'\}$ .

### 4.2. Numerical results

**Setup.** The daily S&P 500 time series used for simulations has  $d = 2^{12}$  samples. 100 realizations of MRW were synthesized with length  $d = 2^{12}$ , and parameters  $H = 0.5$ ,  $\lambda = \sqrt{0.05}$  and  $T = d$ . Morlet wavelets were used, with 1 voice per octave and  $J = 9$  octaves. The L-BFGS-M algorithm was used to perform the gradient descent, with a tolerance  $\epsilon = 10^{-10}$ . Results show averages over 100 reconstructions from S&P 500 time series, and 1 reconstruction for each realization of MRW.

**Multifractal random walk.** Figure 3 (top) shows a reconstructed realization of MRW. A comparison with Fig. 2 shows



**Fig. 3. MRW: scattering reconstruction.** Top: example of reconstruction. Bottom, left to right: multifractal scaling function  $\zeta(q)$ , histogram  $p(x)$ , and leverage correlation  $L\tau$ , for the original MRW data (black) and reconstructions (red).

that the irregular, intermittent temporal structure is satisfactorily captured by scattering moments. Further, Fig. 3 (bottom row) shows the scaling function  $\zeta$ , the histograms  $p$  and the leverage  $L$  of the original and reconstructed signals. The estimates for  $\zeta$  and  $p$  computed from reconstructions and originals are remarkably close, suggesting that a second-order scattering representation captures well the heavy-tailed, multifractal nature of MRW.

Since the components  $B$  and  $\omega_\lambda$  in MRW are independent, its leverage correlation is 0 and provides no information.

Let us emphasize that even though multifractal properties are defined through high-order moments (see (2)), they are completely recovered using only second-order scattering moments, confirming observations in [19].

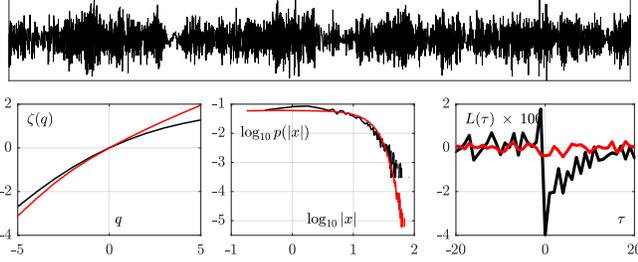
**Financial time series.** Figure 4 (top) shows a reconstructed realization of the S&P 500 daily returns. Comparison with Fig. 1 again shows that the temporal intermittency and general shape are correctly reflected. Inspection of the multifractal properties and histograms in Fig. 4 (bottom left and right, respectively) shows that reconstructions correctly capture both statistical properties.

However, Fig. 4 (bottom right) shows that the scattering representation completely fails to recover temporal asymmetries quantified by the leverage  $L$ : the estimated  $L$  is null for all lags  $\tau$ . Considering that scattering moments measure  $\langle U_\ell^S x, U_{\ell'}^S x \rangle = \langle ||x \star \psi_{j_1}| \star \psi_{j_2}|^2 \rangle$ , it becomes clear that they are unable to capture temporal asymmetries between  $x(t)$  and  $x(-t)$  because the modulus loses all phase information.

## 5. PHASE HARMONIC MOMENTS

### 5.1. Phase-harmonic representation

**Multiscale phase correlations.** To capture temporal asymmetries, it is necessary to use quantities that preserve their phase, and to measure their interactions at different scales. Measuring the correlation  $\langle x \star \psi_j, x \star \psi_{j'} \rangle$  does not work: the Parseval formula reveals that it is small since the supports of  $\hat{\psi}_j$  and  $\hat{\psi}_{j'}$  barely overlap. These issues can be overcome by



**Fig. 4. Financial time series: scattering reconstruction.** Top: example of reconstruction. Bottom, left to right: multifractal scaling function  $\zeta(q)$ , histogram  $p(x)$ , and leverage correlation  $L\tau$ , for the original S&P 500 data (black) and reconstructions (red).

resorting to phase harmonics, defined below [20].

**Wavelet phase harmonics.** Let  $\varphi(z)$  denote the phase of  $z \in \mathbb{C}$ . Wavelet phase harmonics are defined as [20]

$$\forall k \in \mathbb{Z}, \quad [x \star \psi_j]^k = |x \star \psi_j| e^{ik\varphi(x \star \psi_j)}. \quad (8)$$

Wavelet phase harmonics  $[x \star \psi_j]^k$  have thus the same modulus than  $x \star \psi_j$ , but their phase has been accelerated by a factor  $k$ .

The Fourier transform of  $x \star \psi_j$  is centered at frequency  $\omega_0 2^{-j}$  and has a support of size  $\beta_0 2^{-j}$ , where  $\omega_0$  and  $\beta_0$  are the central frequency and bandwidth of the mother wavelet. The exponent  $k$  of  $[x \star \psi_j]^k$  accelerates  $k$  times the phase of  $x \star \psi_j$ , and thus the Fourier transform of  $[x \star \psi_j]^k$  is centered at frequency  $k\omega_0 2^{-j}$ . In the particular case  $k = 0$ ,  $[x \star \psi_j]^0 = |x \star \psi_j|$  and the Fourier transform is centered at frequency 0. Assuming that the supports of the Fourier transforms of  $|x \star \psi_j|$  and  $\varphi(x \star \psi_j)$  are  $\sim \beta_0 2^{-j}$ , then the exponent  $k$  performs  $k$  convolutions in the Fourier domain, expanding the support of the Fourier transform of  $[x \star \psi_j]^k$  to  $\sim k\beta_0 2^{-j}$ .

**Phase-harmonic moments.** Phase-harmonic moments are defined as

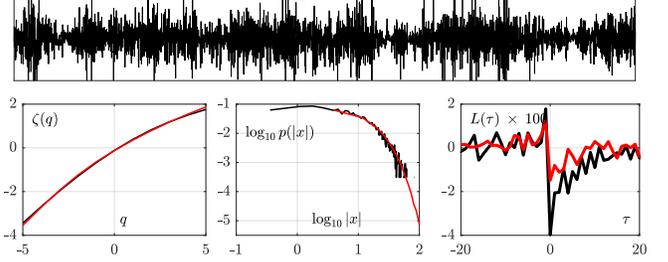
$$U_\ell^P x = [x \star \psi_j]^k \quad \text{with } \ell = (j, k), \quad 1 \leq j_1 \leq J. \quad (9)$$

Note that [20] shows that  $U_\ell^S$  satisfies Lipschitz condition (4).

Correlations  $\langle U_\ell^P x, U_{\ell'}^P x \rangle$  are redundant and highly sparse. The set  $\mathcal{C}^P$  defined by the following conditions that yield nonzero, nonredundant and informative correlations: i)  $j$  arbitrary,  $k > 1$ ,  $k' = 1$  and  $j' = kj$  to correlate variability at different scales, ii)  $j > j'$ ,  $k = k' = 0$ , to correlate low-frequency envelopes, and iii)  $j > j'$  arbitrary,  $k = 0$ ,  $k' \in \{1, 2, 3\}$  to correlate envelopes with coarse scales. When  $k$  or  $k'$  are nonzero, these moments retain all phase information, and are thus able to measure the effects of temporal asymmetries at different scales.

## 5.2. Combined microcanonical model

A phase-harmonic microcanonical model is obtained using representation (9) in (5) to define the phase-harmonic loss



**Fig. 5. Financial time series: mixed reconstruction.** Top: example of reconstruction. Bottom, left to right: multifractal scaling function  $\zeta(q)$ , histogram  $p(x)$ , and leverage correlation  $L\tau$ , for the original S&P 500 data (black) and reconstructions (red).

$\mathcal{E}^P$ . Finally, the mixed loss that combines phase-harmonic and scattering moments is defined as

$$\mathcal{E}(x, \tilde{x}) = \frac{\mathcal{E}_S(x, \tilde{x})}{\mathcal{E}_S(x, 0)} + \frac{\mathcal{E}_P(x, \tilde{x})}{\mathcal{E}_P(x, 0)} \quad (10)$$

## 5.3. Numerical results

**Financial time series.** Figure 5 (top) shows a reconstructed realization of S&P daily returns using the combined microcanonical model. The temporal intermittency and general irregularity are correctly captured. Fig. 5 (bottom right) further shows that the addition of phase-harmonic moments improves estimates of multifractal properties: the functions  $\zeta$  for the original and replicates are almost indistinguishable. The histograms in Fig. 5 (bottom middle) also reproduce the originals almost exactly.

Figure 5 (bottom right) shows the main benefit of the addition of phase-harmonic moments: the leverage effect is correctly captured. The leverage  $L(\tau)$  is found to be 0 for  $\tau < 0$ , and negative for  $\tau > 0$ . Further, despite a clear bias, the dynamics of the recovered  $L(\tau)$  for  $\tau > 0$  resembles an exponential with a similar time constant as the original. These results clearly suggest that phase-harmonic moments succeed in capturing temporal asymmetries in the data.

## 6. CONCLUSIONS

In this paper, we proposed a stochastic model to represent intermittent time series with nongaussian heavy-tailed distributions, long-range correlations, and temporal asymmetries. We showed that it provides a good model for financial time series such as the S&P 500 daily returns. Our results remarkably show that second-order scattering moments are enough to capture the high-order statistics typical of multifractal processes. Further, our results show that phase-harmonic moments correctly capture the phase relationships lost by scattering moments, and are able to reproduce temporal asymmetries such as the leverage effect.

## 7. REFERENCES

- [1] Walter Willinger, Vern Paxson, and Murad S Taqqu, “Self-similarity and heavy tails: Structural modeling of network traffic,” *A practical guide to heavy tails: statistical techniques and applications*, vol. 23, pp. 27–53, 1998.
- [2] B. B. Mandelbrot, “Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier,” *J. Fluid Mech.*, vol. 62, pp. 331–358, 1974.
- [3] Ary L Goldberger, Luis AN Amaral, Jeffrey M Hausdorff, Plamen Ch Ivanov, C-K Peng, and H Eugene Stanley, “Fractal dynamics in physiology: alterations with disease and aging,” *Proceedings of the national academy of sciences*, vol. 99, no. suppl 1, pp. 2466–2472, 2002.
- [4] Lisa Borland, Jean-Philippe Bouchaud, J.F. Muzy, and Gilles Zumbach, “The dynamics of financial markets – mandelbrot’s multifractal cascades, and beyond,” *Wilmott Magazine*, pp. 1–25, March 2005.
- [5] Benoit Mandelbrot, Adlai Fisher, and Laurent Calvet, “A Multifractal Model of Asset Returns,” *Cowles Foundation discussion paper No. 1164*, 1997.
- [6] Robert F Engle, “Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation,” *Econometrica: Journal of the Econometric Society*, pp. 987–1007, 1982.
- [7] Tim Bollerslev, “Generalized autoregressive conditional heteroskedasticity,” *Journal of econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [8] Rémy Chicheportiche and Jean-Philippe Bouchaud, “The fine-structure of volatility feedback i: Multi-scale self-reflexivity,” *Physica A: Statistical Mechanics and its Applications*, vol. 410, pp. 174–195, 2014.
- [9] J.F. Muzy, J. Delour, and E. Bacry, “Modelling fluctuations of financial time series: From cascade process to stochastic volatility model,” *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 17, no. 3, pp. 537–548, 2000.
- [10] J. F. Muzy, E. Bacry, and A. Kozhemyak, “Extreme values and fat tails of multifractal fluctuations,” *Physical Review E*, vol. 73, no. 6, pp. 066114, 2006.
- [11] Rama Cont, “Volatility clustering in financial markets: empirical facts and agent-based models,” in *Long memory in economics*, pp. 289–309. Springer, 2007.
- [12] Rama Cont, “Empirical properties of asset returns: stylized facts and statistical issues,” *Quantitative Finance*, vol. 1, no. 2, pp. 223–236, 2001.
- [13] R. H. Riedi, “Multifractal processes,” in *Theory and applications of long range dependence*, P. Doukhan, G. Oppenheim, and M.S. Taqqu, Eds. 2003, pp. 625–717, Birkhäuser.
- [14] H. Wendt, P. Abry, and S. Jaffard, “Bootstrap for empirical multifractal analysis,” *IEEE Signal Proc. Mag.*, vol. 24, no. 4, pp. 38–48, 2007.
- [15] E. Bacry, J. Delour, and J. F. Muzy, “Multifractal random walk,” *Physical Review E*, vol. 64, no. 2, pp. 026103, 2001.
- [16] Jean-Philippe Bouchaud, Andrew Matacz, and Marc Potters, “Leverage Effect in Financial Markets: The Retarded Volatility Model,” *Physical Review Letters*, vol. 87, no. 22, pp. 228701, 2001.
- [17] Joan Bruna and Stéphane Mallat, “Multiscale Sparse Microcanonical Models,” 2018, Under review.
- [18] Stéphane Mallat, “Group Invariant Scattering,” *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.
- [19] Joan Bruna, Stéphane Mallat, Emmanuel Bacry, and Jean-François Muzy, “Intermittent process analysis with scattering moments,” *The Annals of Statistics*, vol. 43, no. 1, pp. 323–351, 2015-02.
- [20] Stéphane Mallat, Sixin Zhang, and Gaspar Rochette, “Multiscale Correlations with Wavelet Phase Harmonics,” 2018, In preparation.