MAXIMUM-ENTROPY SCATTERING MODELS FOR FINANCIAL TIME SERIES

Roberto Leonarduzzi[1], *Gaspar Rochette*[1], *Jean-Philippe Bouchaud*[2] *and Stéphane Mallat*[1,3]

[1] École Normale Supérieure, PSL, 75005, Paris, France
[2] Science & Finance, Capital Fund Management, 75009, Paris, France
[3] Collège de France, 75005 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. 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 new 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 due to causality effects. 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 prediction or to understand mechanisms that control the dynamics of financial markets. However, financial time series have the complex statistical properties described above, which current models fail to characterize completely [4, 5].

Financial models. Initial gaussian models, based on brownian motions, were unable to capture neither heavy-tails nor the complex temporal dependences (for instance, the alternation of periods of large and small changes known as "volatility clustering" [5], see Fig. 1).

Work supported by ERC InvariantClass 320959 and the CFM-ENS chair.

GARCH models were 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 limitations of GARCH [5, 9]. They specify the evolution of all moments at different time scales. Interscale relationships are captured, and imposing high-order moments leads to heavy-tailed distributions and volatility clustering [5, 10, 11]. However, multifractal quantities are based on high-order moments which are computed with estimators having a large variance. Accurate estimation of high-order moments often requires much more data than what is available. High-order moments also do not capture temporal asymmetries.

Modeling challenges. Accurate stochastic models for complex time-series are required to satisfy the following four properties [4, 5, 11–15]:

1. Compute estimators having a small variance,
2. Capture interactions between scales,
3. Recover intermittent temporal structure,
4. Specify temporal asymmetries.

Goals and contributions. We introduce a maximum entropy model based on moment estimators which satisfy the four properties above. We define estimators with a small variance by estimating a reduced number of second-order moments of a nonlinear contractive representation. We capture interscale relationships with wavelet transforms which separate signal variations at different scales. Intermittent temporal evolutions are captured by scattering coefficients, which are sufficient to model heavy tails and multifractal properties. Temporal asymmetries are specified by wavelet phase harmonics at different scales.

Numerical results show that this model captures heavy tails, multifractal properties and temporal asymmetries of synthetic multifractal random processes as well real financial time series, without estimating high-order moments.

978-1-5386-4658-8/18/$31.00 ©2019 IEEE

2. STATISTICAL PROPERTIES OF FINANCIAL TIME SERIES

Notation. Throughout the paper, $X(t)$ denotes a 1D random process, and x a realization of X with d time samples. Further, $\langle x(t) \rangle = d^{-1} \sum_t x(t)$ denotes the empirical average.
Financial returns. Let $c(t)$ be the price of an asset at time t. The absolute returns at scale δ are defined as $r(t) = c(t) - c(t-\delta)$. We shall 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 have 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 δ is not too large [4, 12]:

$$p(r) \sim r^{-\alpha}, \quad r \to \infty, \quad \text{where } 3 \leq \alpha \leq 5. \tag{1}$$

The probability of observing extreme returns is higher than what is predicted by a normal distribution. We denote $P(r)$ the corresponding cumulative distribution function (cdf).
Multifractal structure. Empirical analyses also show that returns are scale invariant, have nongaussian high-order statistics, and are everywhere irregular with intermittent changes of regularity [4, 5, 11]. Multifractal analysis summarizes these properties through a *scaling function* $\zeta(q)$, defined for all moments of order $q \in \mathbb{R}$ by

$$\langle |x(t+a) - x(t)|^q \rangle \sim a^{\zeta(q)}. \tag{2}$$

The function $\zeta(q)$ specifies the evolutions of moments across scales. It is related to heavy tails, long-range correlations and volatility clustering [5, 9, 16]. Numerically, the increments in (2) are replaced by more robust estimators such as wavelet leaders [17].
Multifractal random walk. Multifractal random walk (MRW) [9, 18] has been introduced to model financial time series. An MRW is defined as $X(t) = B(t)e^{\omega_\lambda(t)}$, where $B(t)$ is white gaussian noise and $\omega_\lambda(t)$ is a stationary gaussian process whose covariance has a slow decay up to an integral scale T: $\operatorname{cov}(\omega(t_1), \omega(t_2)) = \lambda^2 \log(T/|t_1 - t_2| + 1)$ if $|t_1 - t_2| \leq T$, and $\operatorname{cov}(\omega(t_1), \omega(t_2)) = 0$ otherwise. A MRW captures 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 (black 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, see [19]. The strongest stigma of such asymmetries in financial markets is the leverage effect, which reflects the tendency for increased volatility after decreases in price [20]. It can be defined as the correlation between price change and a measure of the square volatility [20]:

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

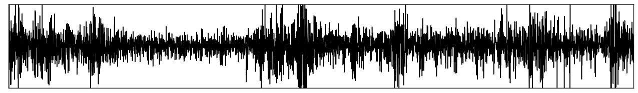

Fig. 1. S&P 500 daily absolute returns time series.

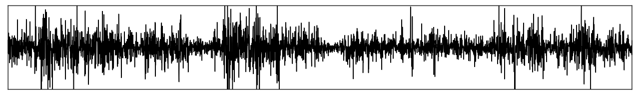

Fig. 2. Single realization of multifractal random walk.

The leverage $L(\tau)$ has been found to be 0 for $\tau < 0$, and a negative exponential for $\tau > 0$ [20].
Contribution. This paper defines models of stochastic processes that capture properties (1), (2) and (3) with low-variance estimators, hence not based on high-order moments.

3. MAXIMUM ENTROPY MODELS

Low-variance moment estimators. We build maximum entropy models that are conditioned on empirical moments computed from a single realization x of X. These moments are covariance coefficients of a contractive representation UX of X. We define a vector of time series $Ux(t) = \{U_\ell x(t)\}_\ell$ which is contractive for all signals x and x' in the sense that:

$$\sum_\ell \sum_t |U_\ell x(t) - U_\ell x'(t)|^2 \leq \sum_t |x(t) - x'(t)|^2, \tag{4}$$

i.e. U preserves or reduces distances.

The covariance of UX is estimated from a single realization x by computing the empirical means $\mu_\ell = \langle U_\ell x \rangle$ and the empirical covariances relatively to this mean

$$C_{\ell,\ell'} x = \langle (U_\ell x - \mu_\ell)\, (U_{\ell'} x - \mu_{\ell'})^* \rangle. \tag{5}$$

The contraction guaranties that UX reduces the variability of X and hence that these covariance estimators have a variance dominated by the variance of second order moments of X. We estimate a minimum subset of covariance coefficients $C_{\ell,\ell'} x$ for few indices $(\ell, \ell') \in \mathcal{C}$ to again reduce the estimation variance. Sections 4 and 5 explain how to choose of U and $\mathcal{C}$ while restoring the important properties of X.
Sampling microcanonical models. A microcanonical model is a maximum entropy model conditioned by empirical moments computed from a single realization x. For covariance moments, it is a uniform probability distribution over a set of signals $\tilde{x}$ such that $U\tilde{x}$ has nearly the same empirical covariance in as Ux over $\mathcal{C}$:

$$\mathcal{E}(x, \tilde{x}) = \sum_{(\ell,\ell') \in \mathcal{C}} |C_{\ell,\ell'} x - C_{\ell,\ell'} \tilde{x}|^2 \leq \epsilon\,. \tag{6}$$

We sample this probability distribution with a gradient descent algorithm whose properties are studied in [21]. It begins from x_0 sampled from a maximum entropy distribution, hence a gaussian white noise. The covariance error (6) is minimized by a gradient descent which updates x_0 until it reaches an $\tilde{x}$ which yields a covariance error below a small ϵ.

4. SCATTERING MOMENTS

4.1. Scattering representation

Wavelet transform. To analyze long-range-dependent processes such as financial time series, we separate variabilities at different scales with a wavelet transform. Let ψ 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 ψ at scales 2^j: $\psi_j(t) = 2^{-j}\psi(2^{-j}t)$ for $1 \leq j < J$. Low-frequency information is captured by a low-pass filter ψ_J at the maximum scale 2^J. The wavelet coefficients $x \star \psi_j(t)$ compute variations of x at different scales 2^j. [22].

Scattering moments. Nongaussian processes often have intermittent bursts of activity as in Figs. 1 and 2, which appear through wavelet coefficients of higher amplitude $|x \star \psi_j|$. The distribution of these bursts is characterized by computing the multiscale variations of $|x \star \psi_j|$, through convolutions with a new family of wavelets:

$$U_\ell x = |x \star \psi_j| \star \psi_{j'} \quad \text{with } \ell = (j, j'), \tag{7}$$

where $1 \leq j < j' \leq J$. For appropriate wavelets, it defines a contractive representation in the sense of (4) [22].

If $j' < J$ then wavelets have zero average so $\langle |x \star \psi_j| \star \psi_{j'} \rangle = 0$. The only non-zero empirical means correspond to $j' = J$ and can be written as $\mathbf{l^1}$ norms of wavelet coefficients

$$\mu_\ell = \langle |x \star \psi_j| \star \psi_J \rangle = \langle |x \star \psi_j| \rangle = \|x \star \psi_j\|_1. \tag{8}$$

A scattering transform only computes the diagonal covariance coefficients. Second order scattering coefficients can be rewritten as squared $\mathbf{l^2}$ norms

$$C_{\ell,\ell} x = \langle ||x \star \psi_j| \star \psi_{j'}|^2 \rangle = \||x \star \psi_j| \star \psi_{j'}\|_2^2. \tag{9}$$

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 that the irregular, intermittent temporal structure is satisfactorily captured by scattering moments. The bottom row of

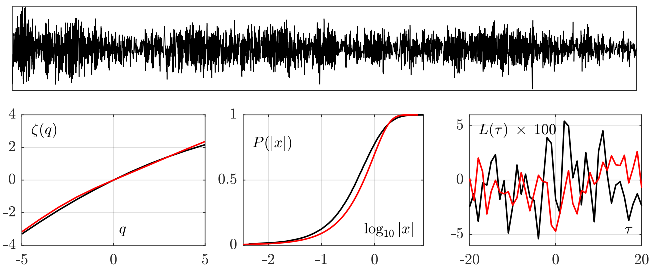

Fig. 3. **MRW: scattering reconstruction.** Top: example of reconstruction. Bottom, left to right: multifractal scaling function $\zeta(q)$, cdf $P(x)$, and leverage $L\tau$, for the original MRW (black) and reconstructions (red).

Fig. 3 shows the scaling function ζ, the cdf P and the leverage L of the original and reconstructed signals. The estimates for ζ 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. The high-order moments of multifractals are completely recovered using only second-order scattering moments, confirming observations in [23].

Since the white B and correlated ω_λ are independent, the leverage correlation of a MRW model is 0 and thus does not reflect the causality properties of financial time-series. The right bottom graph shows that the scattering model also has a 0 leverage, up to random variations of the estimator. Indeed, the scattering means (8) and covariances (9) does not incorporate any phase information. These coefficients remain the same if time is reversed by transforming $x(t)$ in $x(-t)$.

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 cdf in Fig. 4 (bottom left and right, respectively) shows that reconstructions correctly capture both statistical properties. However, the scattering model has no temporal asymmetry, as shown by the leverage L which is nearly zero, at the bottom right of Fig. 4, as opposed to the case of real financial series.

5. JOINT SCATTERING AND 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. The correlation $\langle x \star \psi_j \, x \star \psi_{j'} \rangle$ is nearly zero since the supports of the Fourier transforms $\hat{\psi}_j$ and $\hat{\psi}_{j'}$ barely overlap. Phase-harmonics are non-linear transformations of the phase, providing non-zero correlations across scales.

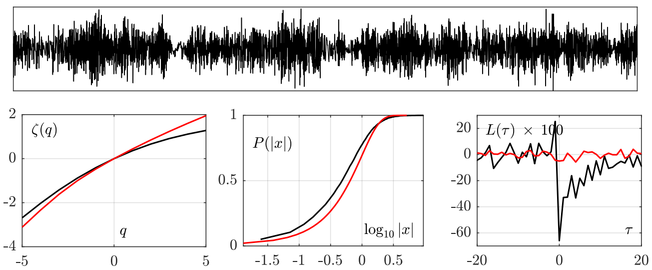

Fig. 4. **Financial time series: scattering reconstruction.** Top: example of reconstruction. Bottom, left to right: multifractal scaling function $\zeta(q)$, cdf $P(x)$, and leverage $L\tau$, for the original S&P 500 data (black) and reconstructions (red).

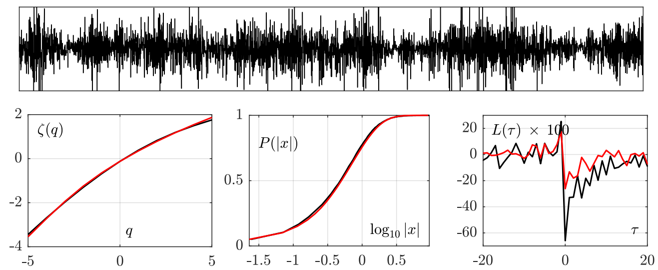

Fig. 5. **Financial time series: mixed reconstruction.** Top: example of reconstruction. Bottom, left to right: multifractal scaling function $\zeta(q)$, cdf $P(x)$, and leverage $L\tau$, for the original S&P 500 data (black) and reconstructions (red).

Wavelet phase harmonics. Let $\varphi(z)$ denote the phase of $z \in \mathbb{C}$. Wavelet phase harmonics are defined in [24] by

$$\forall k \in \mathbb{Z}, \quad [x \star \psi_j]^k = |x \star \psi_j| e^{ik\varphi(x \star \psi_j)}. \qquad (10)$$

Wavelet phase harmonics $[x \star \psi_j]^k$ thus have the same modulus as $x \star \psi_j$, but their phase is amplified by a factor k.

The Fourier transform of $x \star \psi_j$ is $\hat{x}\,\hat{\psi}_j$. Multiplying the phase by a factor k also multiplies all frequencies by k. This nonlinear transformation essentially dilates the Fourier support of $\hat{x}\,\hat{\psi}_j$ by a factor k but it does not modify the variation in time of the modulus $|x \star \psi_j(t)|$. It can thus be interpreted as a frequency transposition, as in a musical score, which transposes frequencies without affecting "rhythms" and "melodies".

Phase-harmonic moments. A phase-harmonic representation is defined by

$$U_\ell x = c_k\,[x \star \psi_j]^k \quad \text{with } \ell = (j,k), \quad 1 \leq j \leq J, \qquad (11)$$

where the multiplicative constants c_k are adjusted to satisfy the contractive property (4). For $k = 0$, the mean is the same as scattering means (8):

$$\mu_\ell = \langle |x \star \psi_j| \rangle = \|x \star \psi_j\|_1 \qquad (12)$$

and for $k \neq 0$ the mean is nearly zero because $\langle \psi_j \rangle = 0$. One can also verify that a wavelet phase harmonic covariance matrix is sparse [24]. Nonzero covariance coefficients capture scale interactions for $\ell = (k,j)$ and $\ell' = (k',j')$ when: i) $k = 1$ and $j' = k' + j$ to correlate variability at different scales, ii) $k = k' = 0$, to correlate low-frequency envelopes, and iii) $k = 0$ and $k' \in \{1,2,3\}$ to correlate envelopes with coarse scales. When k or k' are nonzero, these moments retain the phase information, and measure the effects of temporal asymmetries at different scales.

Joint scattering and phase harmonics models They are computed with a representation Ux which incorporates both the scattering coefficients (7) and the wavelet phase harmonics (10). The resulting covariance matrix is the union of the diagonal scattering covariance coefficients and the scale-interaction wavelet phase-harmonic covariance. Realizations $\tilde{x}$ of the resulting microcanonical models are computed from empirical scattering and wavelet phase-harmonic covariances.

5.2. 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 irregularity are correctly captured. Fig. 5 (bottom left and middle) further shows that the addition of phase-harmonic moments improves estimates of multifractal properties: the functions ζ and P for the original and replicates are almost indistinguishable.

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

This paper introduces 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 demonstrate that second-order scattering moments are sufficient to capture the high-order statistics of multifractal processes. We also showed that phase-harmonic correlations can reproduce temporal asymmetries such as the leverage effect. An extension to other statistical properties and to multivariate situations is under development.

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] Thomas Lux, “Stochastic behavioral asset-pricing models and the stylized facts,” in *Handbook of financial markets: Dynamics and evolution*, pp. 161–215. Elsevier, 2009.

[14] Rosario N Mantegna and H Eugene Stanley, *Introduction to econophysics: correlations and complexity in finance*, Cambridge university press, 1999.

[15] Jean-Philippe Bouchaud and Marc Potters, *Theory of financial risks*, vol. 4, Cambridge University Press, Cambridge From Statistical Physics to Risk . . . , 2000.

[16] 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.

[17] H. Wendt, P. Abry, and S. Jaffard, “Bootstrap for empirical multifractal analysis,” *IEEE Signal Proc. Mag.*, vol. 24, no. 4, pp. 38–48, 2007.

[18] E. Bacry, J. Delour, and J. F. Muzy, “Multifractal random walk,” *Physical Review E*, vol. 64, no. 2, pp. 026103, 2001.

[19] Gilles Zumbach, “Time reversal invariance in finance,” *Quantitative Finance*, vol. 9, no. 5, pp. 505–515, 2009.

[20] 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.

[21] Joan Bruna and Stéphane Mallat, “Multiscale Sparse Microcanonical Models,” 2019, under review. arXiv:1801.02013v.

[22] Stéphane Mallat, “Group Invariant Scattering,” *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.

[23] 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.

[24] Stéphane Mallat, Sixin Zhang, and Gaspar Rochette, “Phase Harmonics and Correlation Invariants in Convolutional Neural Networks,” 2019, under review. arXiv:1810.12136v1.