Scattering Statistics of Generalized Spatial Poisson Point Processes

Michael Perlmutter PERLMUT6@MSU.EDU
Jieqian He HEJIEQIA@MSU.EDU
Department of Computational Mathematics, Science & Engineering
Michigan State University
East Lansing, Michigan, USA

Matthew Hirn MHIRN@MSU.EDU
Department of Computational Mathematics, Science & Engineering
Department of Mathematics
Michigan State University
East Lansing, Michigan, USA

Abstract

We present a machine learning model for the analysis of randomly generated discrete signals, which we model as the points of a homogeneous or inhomogeneous, compound Poisson point process. Like the wavelet scattering transform introduced by S. Mallat, our construction is a mathematical model of convolutional neural networks and is naturally invariant to translations and reflections. Our model replaces wavelets with Gabor-type measurements and therefore decouples the roles of scale and frequency. We show that, with suitably chosen nonlinearities, our measurements distinguish Poisson point processes from common self-similar processes, and separate different types of Poisson point processes based on the first and second moments of the arrival intensity $\lambda(t)$, as well as the absolute moments of the charges associated to each point.

Keywords: Scattering transform, Poisson point process, convolutional neural network

1. Introduction

Convolutional neural networks (CNNs) have obtained impressive results for a number of learning tasks in which the underlying signal data can be modelled as a stochastic process, including texture discrimination (Sifre and Mallat, 2013), texture synthesis (Gatys et al., 2015; Antognini et al., 2018), time series analysis (e.g., finance) (Binkowski et al., 2018), and wireless networks (Brochard et al., 2018). However, precise theoretical understanding of these results is lacking. In many scenarios it is natural to model the signal data as the points of a (potentially complex) spatial point process. Furthermore, there are numerous other fields, including stochastic geometry (Haenggi et al., 2009), forestry (Genet et al., 2014), geoscience (Schoenberg, 2016) and genetics (Fromion et al., 2013), in which spatial point processes are used to model the underlying generating process of certain phenomena (e.g., earthquakes). Motivated by these existing empirical results, as well as the potential for numerous others in yet untapped research, we consider the capacity of CNNs to capture the statistical properties of spatial point processes.

In order to facilitate provable statistical guarantees, we leverage a CNN architecture similar to the wavelet scattering transform of Mallat (2012). The wavelet scattering transform is a mathematical model for CNNs that is provably invariant to local (or global) translations of the input signal and is also Lipschitz stable to the actions of diffeomorphisms on the input. It has been used to achieve near state of the art results in the fields of audio signal processing (Andén and Mallat, 2011, 2014; Wolf et al., 2014, 2015; Andén et al., 2018), computer vision (Bruna and Mallat, 2011, 2013; Sifre and Mallat, 2012, 2013, 2014; Oyallon and Mallat, 2015), and quantum chemistry (Hirn et al., 2017; Eickenberg et al., 2017, 2018; Brumwell et al., 2018), amongst others. It consists of an alternating cascade of linear wavelet transforms and complex modulus nonlinearities. In this paper, we examine a generalized scattering transform that utilizes a broader class of filters, but which still includes wavelets as a special case. Our main focus is on scattering architectures constructed with filters that have small spatial support as is the case in most traditional CNNs.

Expected wavelet scattering moments for stochastic processes with stationary increments were introduced in Bruna et al. (2015), where it is shown that such moments capture important statistical information of one-dimensional Poisson processes, fractional Brownian motion, α-stable Lévy processes, and a number of other stochastic processes. In this paper, we extend the notion of scattering moments to our generalized architecture, and in the process of doing so, we recover many of the important small scale results in Bruna et al. (2015). However, the main contributions contained here consist of new results for more general spatial point processes, including inhomogeneous Poisson point processes, which are not stationary and do not have stationary increments. The collection of expected scattering moments is a non-parametric model for these processes, which we prove captures important summary statistics of inhomogeneous, compound spatial Poisson point processes.

The remainder of this paper is organized as follows. Expected scattering moments are introduced in Section 2. Sections 3 and 4 analyze the first-order and second-order scattering moments of inhomogeneous, compound spatial Poisson point processes. Section 5 compares the scattering moments of one-dimensional Poisson processes to two self-similar processes, fractional Brownian motions and the α-stable process. Section 6 presents stylized numerical examples to highlight certain aspects of the presented theory. A short conclusion is given in Section 7. All proofs are in the appendices, in addition to details on the numerical work.

2. Expected Scattering Moments for Random Signed Measures

Let $\psi \in \mathbf{L}^2(\mathbb{R})$ be a compactly supported mother wavelet with dilations $\psi_j(t) = 2^{-j}\psi(2^{-j}t)$, let and $X(t), t \in \mathbb{R}$, be a stochastic process with stationary increments defined on the real line. In Bruna et al. (2015), first-order wavelet scattering moments are defined as $SX(j) = \mathbb{E}[|\psi_j * X|]$, where the expectation does not depend on t since if $X(t)$ has stationary increments, then $X * \psi_j(t)$ is stationary so long as ψ_j is a wavelet. Much of the mathematical analysis of wavelet scattering moments relies on the fact that they can be rewritten as $SX(j) = \mathbb{E}[|\overline{\psi}_j * dX|]$, where $\overline{\psi}_j$ is the primitive of ψ_j, i.e., $d\overline{\psi}_j = \psi_j$. This reformulation motivates us to define scattering moments as the integration of a filter, which is not necessarily a wavelet, against a random signed measure $Y(dt)$.

To that end, let $w \in \mathbf{L}^2(\mathbb{R}^d)$ be a continuous window function with support contained in the unit cube $[0,1]^d$. Denote by $w_s(t) = w\left(\frac{t}{s}\right)$ the dilation of w, supported on the cube $Q_s = [0,s]^d$, and set $g_\gamma(t)$ to be the Gabor-type filter with scale s and central frequency $\xi \in \mathbb{R}^d$,

$$g_\gamma(t) = w_s(t)e^{i\xi\cdot t}, \quad \gamma = (s,\xi), \ \ t \in \mathbb{R}^d\,. \tag{1}$$

Note that with an appropriately chosen window function w, (1) includes dyadic wavelet families in the case that one selects $s = 2^j$ and $|\xi| = C/s$. However, it also includes many other families of filters, including Gabor filters used in the windowed Fourier transform.

For a random signed measure $Y(dt)$ we define the first-order $\mathbf{L}^p$ scattering moments, $1 \leq p < \infty$, at location t as

$$S_{\gamma,p}Y(t) := \mathbb{E}[|g_\gamma * Y(t)|^p] := \mathbb{E}\left[\left|\int_{\mathbb{R}^d} g_\gamma(t-u)\,Y(du)\right|^p\right]. \tag{2}$$

Note there is no assumption on the stationarity of $Y(du)$, which is why these scattering moments a priori depend on t. We define invariant (i.e., location independent) first-order scattering coefficients of Y by

$$SY(\gamma,p) = \lim_{R\to\infty} \frac{1}{(2R)^d} \int_{|t_i|\leq R} \mathbb{E}[|g_\gamma * Y(t)|^p]\,dt\,, \tag{3}$$

if the limit on the right hand side exists.

We call Y a periodic measure if there exists $T > 0$ such that for any Borel set B, the family of sets $B + Te_i = \{b + Te_i : b \in B\}$ satisfies

$$Y(B) \overset{d}{=} Y(B + Te_i), \quad \forall\, 1 \leq i \leq d\,,$$

where $\{e_i\}_{i\leq d}$ is the standard orthonormal basis for $\mathbb{R}^d$. In this case one can verify, by approximating g_γ with simple functions, that $(g_\gamma * Y)(t + Te_i) \overset{d}{=} (g_\gamma * Y)(t)$, and therefore

$$S_{\gamma,p}Y(t + Te_i) = S_{\gamma,p}Y(t), \quad \forall\, t \in \mathbb{R}^d\,.$$

Thus the limit in (3) exists, and

$$SY(\gamma,p) = \frac{1}{T^d} \int_{Q_T} \mathbb{E}[|g_\gamma * Y(t)|^p]\,dt\,. \tag{4}$$

Note that in the special case when the distribution of $Y(B)$ depends only on the Lebesgue measure of B, then $S_{\gamma,p}Y(t)$ is independent of t and the above limit (3) exists with $SY(\gamma,p) = S_{\gamma,p}Y(t)$ for any $t \in \mathbb{R}^d$.

First-order scattering moments compute summary statistics of the measure Y based upon its responses against the filters g_γ. Higher-order summary statistics can be obtained by computing first-order scattering moments for larger powers p, or by cascading lower-order modulus nonlinearities as in a CNN. This leads us to define second-order scattering moments by

$$S_{\gamma,p,\gamma',p'}Y(t) = \mathbb{E}\left[||g_\gamma * Y|^p * g_{\gamma'}(t)|^{p'}\right]\,.$$

First-order invariant scattering moments collapse additional information by aggregating the variations of the random measure Y, which removes information related to the intermittency of Y. Second-order invariant scattering moments augment first-order scattering moments by iterating on the cascade of linear filtering operations and nonlinear $|\cdot|^p$ operators, thus recovering some of this lost information. They are defined (assuming the limit on the right exists) by

$$SY(\gamma,p,\gamma',p') = \lim_{R\to\infty} \frac{1}{(2R)^d} \int_{|t_i|\leq R} \mathbb{E}\left[\left||g_\gamma * Y|^p * g_{\gamma'}(t)\right|^{p'}\right]\,dt\,.$$

The collection of (invariant) scattering moments is a set of non-parametric statistical measurements of the random measure Y. In the following sections, we analyze these moments for arbitrary frequencies ξ and small scales s, thus allowing the filters g_γ to serve as a model for the learned filters in CNNs. In particular, we will analyze the asymptotic behavior of the scattering moments as the scale parameter s decreases to zero.

3. First-Order Scattering Moments of Generalized Poisson Processes

We consider the case where $Y(dt)$ is an inhomogeneous, compound spatial Poisson point process. Such processes generalize ordinary Poisson point processes by incorporating variable charges (heights) at the points of the process and a non-uniform intensity for the locations of the points. They thus provide a flexible family of point processes that can be used to model many different phenomena. In this section we consider first-order scattering moments of these generalized Poisson processes. In Sec. 3.1 we provide a review of such processes, and in Sec. 3.2 we show that first-order scattering moments capture a significant amount of statistical information related these processes, particularly when using very localized filters.

3.1. Inhomogeneous, Compound Spatial Poisson Point Processes

Let $\lambda(t)$ be a continuous function on $\mathbb{R}^d$ such that

$$0 < \lambda_{\min} := \inf_t \lambda(t) \leq \|\lambda\|_\infty < \infty\,, \tag{5}$$

and let $N(dt)$ be an inhomogeneous Poisson point process with intensity function $\lambda(t)$. That is,

$$N(dt) = \sum_{j=1}^{\infty} \delta_{t_j}(dt)$$

is a random measure, concentrated on a countable set of points $\{t_j\}_{j=1}^\infty$, such that for all Borel sets $B \subset \mathbb{R}^d$, the number of points of N in B, denoted $N(B)$, is a Poisson random variable with parameter

$$\Lambda(B) = \int_B \lambda(t)\, dt\,, \tag{6}$$

i.e.,

$$\mathbb{P}[N(B) = n] = e^{-\Lambda(B)} \frac{(\Lambda(B))^n}{n!}\,,$$

and $N(B)$ is independent of $N(B')$ for all sets B' that do not intersect B. Now let $(A_j)_{j=1}^\infty$ be a sequence of i.i.d. random variables independent of N, and let $Y(dt)$ be the random signed measure that gives charge A_j to each point t_j of N, i.e.,

$$Y(dt) = \sum_{j=1}^{\infty} A_j \delta_{t_j}(dt)\,. \tag{7}$$

We refer to $Y(dt)$ as an inhomogeneous, compound Poisson point process. For a Borel set $B \subset \mathbb{R}^d$, $Y(B)$ has a compound Poisson distribution and we will (in a slight abuse of notation) write

$$Y(B) = \sum_{j=1}^{N(B)} A_j\,.$$

In many of our proofs, it will be convenient to consider the random measure $|Y|^p(dt)$ defined formally by

$$|Y|^p(dt) := \sum_{j=1}^{\infty} |A_j|^p \delta_{t_j}(dt).$$

For a further overview of these processes, and closely related marked point processes, we refer the reader to Section 6.4 of Daley and Vere-Jones (2003).

3.2. First-order Scattering Asymptotics

Computing the convolution of g_γ with $Y(dt)$ gives

$$(g_\gamma * Y)(t) = \int_{\mathbb{R}^d} g_\gamma(t-u)\, Y(du) = \sum_{j=1}^{\infty} A_j g_\gamma(t - t_j)\,,$$

which can be interpreted as a waveform g_γ emitting from each location t_j. Invariant scattering moments aggregate the random interference patterns in $|g_\gamma * Y|$. The results below show that the expectation of these interferences, for small scale waveforms g_γ, encode important statistical information related to the point process.

For notational convenience, we let

$$\Lambda_s(t) := \Lambda\left([t-s,t]^d\right) = \int_{[t-s,t]^d} \lambda(u)\, du$$

denote the expected number of points of N in the support of $g_\gamma(t - \cdot)$. If $\lambda(t)$ is a periodic function in each coordinate with period T, then $\Lambda_s(t) = \Lambda_s(t + Te_i)$ for $1 \le i \le d$ and therefore, the invariant scattering coefficients of Y may be defined as in (4).

Theorem 1 *Let $1 \le p < \infty$ and suppose that $Y(dt)$ is an inhomogeneous, compound Poisson point process as defined above, where $(A_j)_{j=1}^{\infty}$ is an i.i.d. sequence of random variables, $\mathbb{E}[|A_1|^p] < \infty$ and $\lambda(t)$ is a continuous intensity function satisfying (5). Then for every $t \in \mathbb{R}^d$, every $\gamma = (s, \xi)$ such that $s^d\|\lambda\|_\infty < 1$, and for every $m \ge 1$.*

$$S_{\gamma,p}Y(t) = \sum_{k=1}^{m} e^{-\Lambda_s(t)} \frac{(\Lambda_s(t))^k}{k!} \mathbb{E}\left[\left|\sum_{j=1}^{k} A_j w(V_j) e^{is\xi\cdot V_j}\right|^p\right] + \epsilon(m,s,\xi,t)\,, \tag{8}$$

where the error term $\epsilon(m, s, \xi, t)$ satisfies

$$|\epsilon(m,s,\xi,t)| \le C_{m,p} \frac{\|\lambda\|_\infty}{\lambda_{\min}} \|w\|_p^p \mathbb{E}[|A_1|^p] \|\lambda\|_\infty^{m+1} s^{d(m+1)} \tag{9}$$

and $V_1, V_2, \ldots$ is an i.i.d. sequence of random variables, independent of the A_j, taking values in the unit cube Q_1 and with density

$$p_V(v) = \frac{s^d}{\Lambda_s(t)} \lambda(t - vs)\,, \quad v \in Q_1\,.$$

The main idea of the proof of Theorem 1 is to condition on $N\left([t-s,t]^d\right)$, which is the number of points in the support of g_γ, and to use the fact that

$$\mathbb{P}\left[N\left([t-s,t]^d\right) > m\right] = O\left(\left(s^d\|\lambda\|_\infty\right)^{m+1}\right), \quad \forall\, s^d\|\lambda\|_\infty < 1\,.$$

Theorem 1 shows that even at small scales the scattering moments $S_{\gamma,p}Y(t)$ depend upon higher-order information related to the distribution of the points, encapsulated by the term $(\Lambda_s(t))^k$, regardless of the scattering moment p. However, the influence of the higher-order terms diminishes rapidly as the scale of the filter shrinks, which is indicated by the bound (9) on the error function. Theorem 1 also shows that p^{th} scattering moments depend on the p^{th} moments of the charges, $(A_j)_{j=1}^\infty$. The next result uses Theorem 1 to examine the behavior of scattering moments for small filters in the asymptotic regime as the scale $s \to 0$.

Theorem 2 *Let $1 \leq p < \infty$, and suppose that $Y(dt)$ is an inhomogeneous, compound Poisson point process satisfying the same assumptions as in Theorem 1. Let $\gamma_k = (s_k, \xi_k)$ be a sequence of scale and frequency pairs such that $\lim_{k\to\infty} s_k = 0$. Then*

$$\lim_{k\to\infty} \frac{S_{\gamma_k,p}Y(t)}{s_k^d} = \lambda(t)\mathbb{E}[|A_1|^p]\|w\|_p^p\,. \tag{10}$$

Furthermore, if $\lambda(t)$ is periodic with period T along each coordinate, then

$$\lim_{k\to\infty} \frac{SY(\gamma_k,p)}{s_k^d} = m_1(\lambda)\mathbb{E}[|A_1|^p]\|w\|_p^p, \quad \textit{where} \quad m_1(\lambda) = \frac{1}{T^d}\int_{Q_T} \lambda(t)\,dt\,. \tag{11}$$

Theorem 2 is proved via asymptotic analysis of the $m = 1$ case of Theorem 1. The key to the proof, which is similar to the technique used to prove Theorem 2.1 of Bruna et al. (2015), is that in a small cube $[t-s,t]^d$ there is at most one point of N with overwhelming probability. Therefore, when s is very small, with very high probability, $|g_\gamma * Y|^p(t) = \left(|g_\gamma|^p * |Y|^p\right)(t)$.

This theorem shows that for small scales the scattering moments $S_{\gamma,p}Y(t)$ encode the intensity function $\lambda(t)$, up to factors depending upon the summary statistics of the charges $(A_j)_{j=1}^\infty$ and the window w. Recall that $\Lambda(B)$, defined in (6), determines the concentration of events within the set B. Thus even a one-layer location dependent scattering network yields considerable information regarding the underlying data generation, at least in the case of inhomogeneous Poisson processes. However, it is often the case, e.g., Bruna and Mallat (2018), that invariant statistics are utilized. In this case (11) shows that invariant scattering statistics mix the mean of $\lambda(t)$ and the p^{th} moment of the charge magnitudes. However, we can decouple these statistics as we now explain.

As a special case, Theorem 2 proves that for non-compound inhomogeneous Poisson processes (i.e., $A_j = 1$ for all $j \geq 1$), small scale scattering moments recover $\lambda(t)$ or $m_1(\lambda)$, depending on whether one computes invariant or time-dependent scattering moments. For compound processes, we can add an additional nonlinearity, namely the signum function sgn, which when applied to the Poisson point process in (7) yields,

$$\overline{Y}(dt) = \operatorname{sgn}[Y(dt)] = \sum_{j=1}^\infty \delta_{t_j}(dt)\,.$$

Thus computing $S\overline{Y}(\gamma, p)$ and the ratio $SY(\gamma, p)/S\overline{Y}(\gamma, p)$ at small scales decouples the mean of $\lambda(t)$ from the p^{th} moment of $|A_1|$. We remark that the signum function is a simple perceptron and is closely related to the sigmoid nonlinearity, which is used in many neural networks. We further remark that the computation of $S\overline{Y}$ constitutes a small two-layer network, consisting of the nonlinear sgn function, the linear filtering by the collection of filters g_γ, the nonlinear p^{th} modulus $|\cdot|^p$, and the linear integration operator.

If $Y(dt)$ is a homogeneous Poisson process, then $\lambda(t)$ is constant, meaning that (10) and (11) are equivalent. In the case of ordinary (non-compound) Poisson processes, Theorem 2 recovers the constant intensity. For periodic $\lambda(t)$ and invariant scattering moments, the effect of higher-order moments of $\lambda(t)$ can be partially isolated by considering higher-order expansions (e.g., $m > 1$) in (8). The next theorem considers second-order expansions and illustrates their dependence on the second moment of $\lambda(t)$.

Theorem 3 *Let $1 \le p < \infty$, and suppose $Y(dt)$ is an inhomogeneous, compound Poisson point process satisfying the same assumptions as in Theorem 1. If $\lambda(t)$ is periodic with period T in each coordinate, and if $(\gamma_k)_{k\ge1} = (s_k, \xi_k)_{k\ge1}$, is a sequence of scale and frequency pairs such that $\lim_{k\to\infty} s_k = 0$ and $\lim_{k\to\infty} s_k\xi_k = L \in \mathbb{R}^d$, then*

$$\lim_{k\to\infty}\left(\frac{SY(\gamma_k, p)}{s_k^{2d}\mathbb{E}[|A_1|^p]\mathbb{E}[|V_k|^p]} - \frac{1}{T^d}\int_{Q_T}\frac{\Lambda_{s_k}(t)}{s_k^{2d}}\,dt\right) = m_2(\lambda)\left(\frac{\mathbb{E}\left[\left|A_1 w(U_1)e^{iL\cdot U_1} + A_2 w(U_2)e^{iL\cdot U_2}\right|^p\right]}{2\|w\|_p^p\mathbb{E}[|A_1|^p]}\right), \quad (12)$$

where $m_2(\lambda) = T^{-d}\int_{Q_T}\lambda(t)^2\,dt$; U_1, U_2 are independent uniform random variables on Q_1; and $(V_k)_{k\ge1}$ is a sequence of random variables independent of the A_j taking values in the unit cube Q_1 and with respective densities,

$$p_{V_k}(v) = \frac{s_k^d}{\Lambda_{s_k}(t)}\lambda(t - vs_k)\,, \quad v \in Q_1\,.$$

We first remark that the scale normalization on the left hand side of (12) is s^{-2d}, compared to a normalization of s^{-d} in Theorem 2. Thus even though (12) is written as a small scale limit, intuitively Theorem 3 is capturing information at moderately small scales that are larger than the scales considered in Theorem 2. This is further indicated by the term multiplied against $m_2(\lambda)$ on the right hand side of (12), which depends on two points of the process (as indicated by the presence of two charges A_1 and A_2).

Unlike Theorem 2, which gives a way to compute $m_1(\lambda)$, Theorem 3 does not allow one to compute $m_2(\lambda)$ since it would require knowledge of $\Lambda_{s_k}(t)$ in addition to the distribution from which the charges $(A_j)_{j=1}^\infty$ are drawn. However, Theorem 3 does show that at moderately small scales the invariant scattering coefficients depend non-trivially on the second moment of $\lambda(t)$. This behavior at moderately small scales can be used to distinguish between, for example, an inhomogeneous Poisson point process with intensity function $\lambda(t)$ and a homogeneous Poisson point process with constant intensity $\lambda_0 = m_1(\lambda)$, whereas Theorem 2 indicates that at very small scales the two processes will have the same invariant scattering moments.

4. Second-Order Scattering Moments of Generalized Poisson Processes

We prove that second-order scattering moments, in the small scale regime, encode higher-order moment information about the charges $(A_j)_{j=1}^\infty$.

Theorem 4 *Let $1 \leq p, p' < \infty$ and $q = pp'$. Suppose that $Y(dt)$ is an inhomogeneous Poisson point process satisfying the same assumptions as in Theorem 1 as well as the additional assumption that $\mathbb{E}|A_1|^q < \infty$. Let $\gamma_k = (s_k, \xi_k)$ and $\gamma'_k = (s'_k, \xi'_k)$ be two sequences of scale and frequency pairs such that $s'_k = cs_k$ for some fixed constant $c > 0$ and $\lim_{k\to\infty} s_k\xi_k = L \in \mathbb{R}^d$. Then,*

$$\lim_{k\to\infty} \frac{S_{\gamma_k,p,\gamma'_k,p'}Y(t)}{s_k^{d(p'+1)}} = K\lambda(t)\mathbb{E}[|A_1|^q] \,, \tag{13}$$

where

$$K \coloneqq \left\|g_{c,L/c} * |g_{1,0}|^p\right\|_{p'}^{p'} ,$$

is a constant depending on p, p', c, L, and w. Furthermore, if $\lambda(t)$ is periodic with period T along each coordinate, then

$$\lim_{k\to\infty} \frac{SY(\gamma_k, p, \gamma'_k, p')}{s_k^{d(p'+1)}} = Km_1(\lambda)\mathbb{E}[|A_1|^q] \,. \tag{14}$$

Note that the scaling factor $s^{-d(p'+1)}$ depends on p' but not p. Intuitively this corresponds to the behavior $\||g_{\gamma_k}|^p * g_{\gamma'_k}\|_{p'}^{p'} \approx s_k^{d(p'+1)}$ as $s_k \to 0$. Theorem 4 proves that second-order scattering moments capture higher-order moments of the charges $(A_j)_{j=1}^\infty$ via two pairs of lower-order filtering and modulus operators. If $p, p' > 1$, then $q = pp'$ will be larger than either p or p' and the result above will give us information about the higher order moment $\mathbb{E}|A_1|^q$.

It is also useful to consider the $p = 1$ case. Indeed, in Sec. 5 below it is shown that first-order invariant scattering moments can distinguish Poisson point processes from fractional Brownian motion and α-stable processes, if $p = 1$, but may fail to do so for larger values of p. However, Theorem 2 shows that first-order invariant scattering moments for $p = 1$ will not be able to distinguish between the various different types of Poisson point processes with a one-layer network at very small scales. Theorem 4 shows that a second-order calculation that augments the first-order calculation with $p = 1$ and $p' > 1$, will capture a higher-order moment of the charges $(A_j)_{j=1}^\infty$.

5. Poisson Point Processes Compared to Self Similar Processes

For one-dimensional processes (i.e., $d = 1$), we show that first-order invariant scattering moments can distinguish between inhomogeneous, compound Poisson point processes and certain self-similar processes. In particular, we show that if $X(t)$ is either an α-stable process or a fractional Brownian motion (fBM), then the corresponding first-order scattering moments will have different asymptotic behavior for infinitesimal scales than in the case of a Poisson point process. Similar results were initially reported in Bruna et al. (2015); here we generalize those results to the non-wavelet filters g_γ defined in (1) and for general p^{th} scattering moments, and further clarify their usefulness in the context of the new results presented in Sec. 3 and Sec. 4. As in Bruna et al. (2015), the proof will be based on the scaling relationships of these processes and therefore will not be able to distinguish between α-stable processes and fBM[1]. The key will be proving a lemma that says if a stochastic process X has a scaling relation, then that scaling relation is inherited by integrals of deterministic functions against dX.

More precisely, for a stochastic process $X(t)$, $t \in \mathbb{R}$, we consider the convolution of the filter g_γ with the noise dX defined by

$$g_\gamma * dX(t) = \int_{\mathbb{R}} g_\gamma(t-u)\, dX(u)\,,$$

and define (in a slight abuse of notation) the first-order scattering moments at time t by

$$S_{\gamma,p}X(t) = \mathbb{E}[|g_\gamma * dX(t)|^p]\,. \tag{15}$$

In the case where $X(t)$ is a compound, inhomogeneous Poisson (counting) process, $Y = dX$ will be a compound Poisson random measure and the scattering moments defined in (15) will coincide with the first-order scattering moments defined in (2).

The following two theorems analyze the small scale first-order scattering moments when X is either an α-stable process, for $1 < \alpha \leq 2$, or fractional Brownian motion. Thus dX will be stable Lévy noise or fractional Gaussian noise, respectively. These results show that the asymptotic decay of the corresponding scattering moments is guaranteed to differ from Poisson point processes, in the case $p = 1$. We also note that both α-stable processes and fBM have stationary increments; therefore the scattering moments do not depend on time and

$$S_{\gamma,p}X(t) = SX(\gamma,p) = \lim_{R\to\infty} \frac{1}{2R}\int_{|u|\leq R} \mathbb{E}[|g_\gamma * dX(u)|^p]\, du\,, \quad \forall\, t \in \mathbb{R}\,.$$

Theorem 5 *Let $1 \leq p < \infty$ and suppose $X(t)$ is a symmetric α-stable process for some $p < \alpha \leq 2$. Let $\gamma_k = (s_k, \xi_k)$ be a sequence of scale and frequency pairs such that $\lim_{k\to\infty} s_k = 0$ and $\lim_{k\to\infty} s_k\xi_k = L \in \mathbb{R}$. Then,*

$$\lim_{k\to\infty} \frac{SX(\gamma_k,p)}{s_k^{p/\alpha}} = \mathbb{E}\left[\left|\int_0^1 w(u)e^{iLu}\, dX(u)\right|^p\right]\,.$$

Theorem 6 *Let $1 \leq p < \infty$, suppose $X(t)$ is a fractional Brownian motion with Hurst parameter $0 < H < 1$. Assume that the window function w has bounded variation on $[0,1]$, and let $\gamma_k = (s_k, \xi_k)$ be a sequence of scale and frequency pairs such that $\lim_{k\to\infty} s_k = 0$ and $\lim_{k\to\infty} s_k\xi_k = L \in \mathbb{R}$. Then,*

$$\lim_{k\to\infty} \frac{SX(\gamma_k,p)}{s_k^{pH}} = \mathbb{E}\left[\left|\int_0^1 w(u)e^{iLu}\, dX(u)\right|^p\right]\,.$$

The key to proving Theorem 5 and Theorem 6 is the following lemma stated in Appendix E, which shows that X_t is a self-similar process, then, then stochastic integrals against dX satisfy an identity corresponding to the scaling relation of X_t.

1. We note that Bruna et al. (2015) proves that second-order scattering moments defined with wavelet filters do distinguish between α-stable processes and fBM, but we do not pursue this direction in this paper as we are concerned primarily with point processes.

Together, these two theorems indicate that first-order invariant scattering moments distinguish inhomogeneous, compound Poisson processes from both α-stable processes and fractional Brownian motion except in the cases where $p = \alpha$ or $p = 1/H$. In particular, if X is a Brownian motion, then SX will distinguish X from a Poisson point process except in the case that $p = 2$. For this reason, it appears that $p = 1$ is the best choice of the parameter p for the purposes of distinguishing a Poisson point process from a self-similar process. In the case of a multi-layer network, it is advisable to set $p = 1$. Larger values of p' in the second layer can then allow us to determine the higher moments of the arrival heights $(A_j)_{j=1}^{\infty}$.

6. Numerical Illustrations

We carry out several experiments to numerically validate the previously stated results and to illustrate their capacity for distinguishing between different types of random processes. In all of the experirments below, we will hold the frequency ξ constant while we let the scale s decrease to zero.

Homogeneous, compound Poisson point processes with the same intensities: We generated three different types of homogeneous compound Poisson point processes, all with the same intensity $\lambda(t) \equiv \lambda_0 = 0.01$. The three point processes are Y_1 (ordinary), Y_2 (Gaussian), and Y_3 (Rademacher), where the charges are sampled according to $(A_{1,j})_{j=1}^{\infty} \equiv 1$, $(A_{2,j})_{j=1}^{\infty} \sim \mathcal{N}(0, \sqrt{\pi/2})$, and $(A_{3,j})_{j=1}^{\infty} \sim$ Rademacher distribution (i.e., ± 1 with equal probability). The charges of the three signals have the same first moment $\mathbb{E}[|A_{i,j}|] = 1$ and different second moment with $\mathbb{E}[|A_{1,j}|^2] = \mathbb{E}[|A_{3,j}|^2] = 1$ and $\mathbb{E}[|A_{2,j}|^2] = \frac{\pi}{2}$. Theorem 2 thus predicts that $p = 1$ invariant first-order scattering moments will not be able to distinguish between the three processes, but $p = 2$ invariant first-order scattering moments will distinguish the Gaussian Poisson point process from the other two. Figure 1 illustrates this point by plotting the normalized invariant scattering moments for $p = 1$ and $p = 2$.

Homogeneous, compound Poisson point processes with different intensities and charges: We consider two homogeneous, compound Poisson point processes with different intensities and different charge distributions, but which nevertheless have the same first-order invariant scattering moments with $p = 1$ due to the mixing of intensity and charge information in (11). The first compound Poisson point process has constant intensity $\lambda_1 = 0.01$ and charges $A_{1,j} \sim \mathcal{N}(0, 1)$, whereas the second has intensity $\lambda_2 = 0.01/\sqrt{2}$ and $A_{2,j} \sim \mathcal{N}(0, 2)$. In this way, $\lambda_1\mathbb{E}[|A_{1,j}|] = \lambda_2\mathbb{E}[|A_{2,j}|] = 0.01 \cdot \sqrt{2/\pi} \approx 0.008$, but $\lambda_1\mathbb{E}[|A_{1,j}|^2] = 0.01$ and $\lambda_2\mathbb{E}[|A_{2,j}|^2] = 0.01 \cdot \sqrt{2} \approx 0.014$. Figure 2 plots the normalized invariant scattering moments for $p = 1$ and $p = 2$.

Inhomogeneous, non-compound Poisson point processes: We also consider inhomogeneous Poisson point processes. We use the intensity function $\lambda(t) = 0.01(1 + 0.5\sin(2\pi t/N))$ to generate inhomogeneous process. To estimate $S_{\gamma,p}Y(t)$, we average the modulus of the scattering transform at time t over 1000 realizations. Figure 3 plots the scattering moments for inhomogeneous process at different time.

Homogeneous, non-compound Poisson point process and self similar process: We consider Brownian motion with Hurst parameter $H = \frac{1}{2}$ and compare it with Poisson point process with intensity $\lambda = 0.01$ and charges $(A)_{j=1}^{\infty} \equiv 10$. Figure 4 shows that the 2nd moments cannot distinguish between Brownian motion and Poisson point process while the 1st moments can.

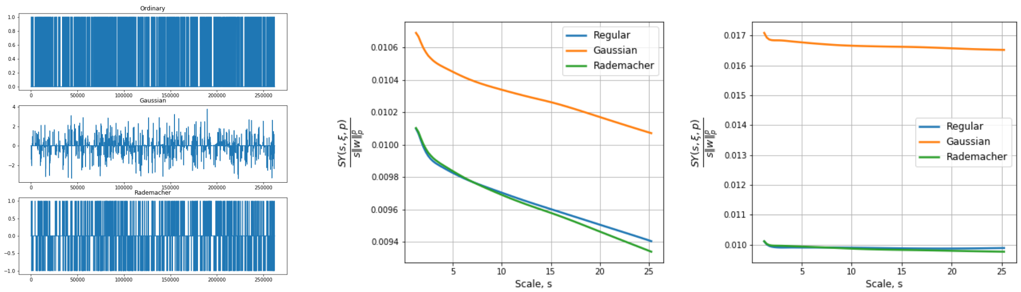

Figure 1: First-order invariant scattering moments for three types of homogeneous compound Poisson point processes with the same intensity λ_0. **Left:** *Top:* ordinary Poisson point process. *Middle:* Gaussian compound Poisson point process with normally distributed charges. *Bottom:* Rademacher compound Poisson point process with charges drawn from the Rademacher distribution. **Middle:** Normalized invariant scattering moments $SY(s,\xi,1)/s\|w\|_1$ (i.e., $p = 1$), which all converge to 0.01 as $s \to 0$ (up to numerical errors) since $\lambda_0 \mathbb{E}[|A_1|]$ is the same for all three point processes. **Right:** Normalized invariant scattering moments $SY(s,\xi,2)/s\|w\|_2^2$ (i.e., $p = 2$). In this case the ordinary Poisson point process and the Rademacher Poisson point process still converge to the same value as $s \to 0$ since $\mathbb{E}[|A_1|^2] = 1$ for both of them. However, the Gaussian Poisson point process converges to a different value since $\mathbb{E}[|A_1|^2] = \pi/2$ for this process.

7. Conclusion

We have constructed Gabor-filter scattering transforms for random measures on $\mathbb{R}^d$, and stochastic processes on $\mathbb{R}$. Our construction is closely related to Bruna et al. (2015), but extends their work in several important ways. First, while our Gabor-type filters include dyadic wavelets as a special case, they also include many other families of filters. We also do not assume that the random measure Y is stationary, and consider compound, possibly inhomogeneous, Poisson random measures on $\mathbb{R}^d$, in addition to ordinary Poisson processes on $\mathbb{R}$. We do note however, that Bruna et al. (2015) provides a detailed analysis of self-similar processes and multifractal random measures, whereas we have primarily focused on models of random sparse signals. We believe the results presented here open up several avenues of future research. Firstly, we have assumed throughout most of this paper that the points of our random measures were distributed according to a possibly inhomogenous Poisson process. It would be interesting to discover if our measurements can distinguish these signals from other point processes. Secondly, it would be interesting to explore the use of these measurements for a variety of machine learning tasks such as synthesizing new signals.

Acknowledgments

J.H. and M.H. are partially supported by NSF grant #1620216. M.H. is also partially supported by Alfred P. Sloan Fellowship #FG-2016-6607 and DARPA YFA #D16AP00117.

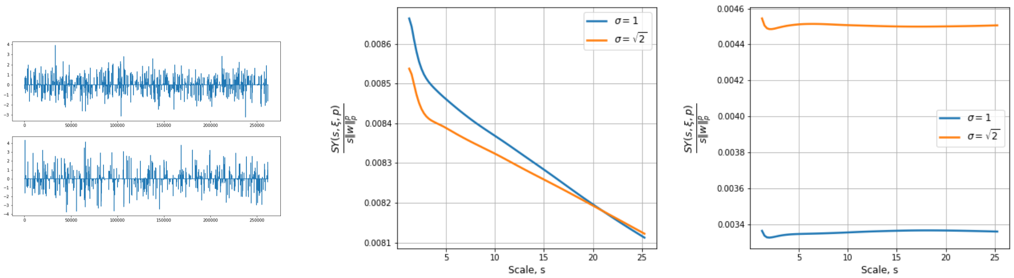

Figure 2: First-order invariant scattering moments for two homogeneous, Gaussian compound Poisson point processes with different intensity and variance. **Left:** *Top:* Homogeneous compound Poisson point process with intensity $\lambda_1 = 0.01$ and charges $A_{1,j} \sim \mathcal{N}(0,1)$. *Bottom:* Homogeneous compound Poisson point process with intensity $\lambda_2 = {}^{0.01}/\sqrt{2}$ and charges $A_{2,j} \sim \mathcal{N}(0,2)$. The two point processes are difficult to distinguish, visually. **Middle:** Normalized invariant scattering moments ${}^{SY(s,\xi,1)}/{s\|w\|_1}$ (i.e., $p = 1$), which both converge to approximately 0.08 up to numerical error, thus indicating that these moments cannot distinguish the two processes. **Right:** Normalized invariant scattering moments ${}^{SY(s,\xi,2)}/{s\|w\|_2^2}$ (i.e., $p = 2$). The two process are distinguished as $s \to 0$ since the values $\lambda_1 \mathbb{E}[|A_{1,j}|^2] = 0.01$ and $\lambda_2 \mathbb{E}[|A_{2,j}|^2] \approx 0.014$ differ by a significant margin.

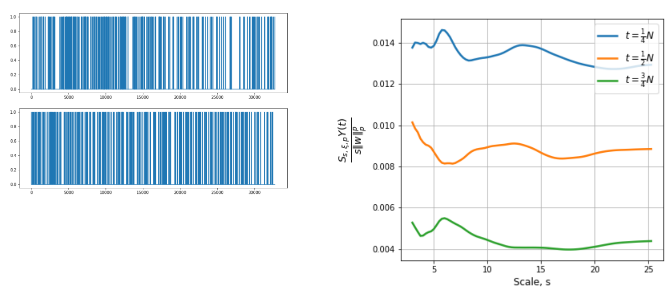

Figure 3: First-order invariant scattering moments for inhomogeneous non-compound Poisson point processes. **Left:** Inhomogeneous non-compound Poisson point process with intensity $\lambda(t) = 0.01(1 + 0.5\sin({}^{2\pi t}/{N}))$. **Right:** Scattering moments ${}^{S_{\gamma,p}Y(t)}/{s\|w\|_p^p}$ for inhomegeneous non-compound Poisson point process at $t_1 = {}^{N}/{4}$, $t_2 = {}^{N}/{2}$, $t_3 = {}^{3N}/{4}$. Note that $\lambda(t_1) = 0.015$, $\lambda(t_2) = 0.01$, $\lambda(t_3) = 0.005$. The plots show that for inhomogeneous process, scattering coefficients at time t converges to the intensity at that time.

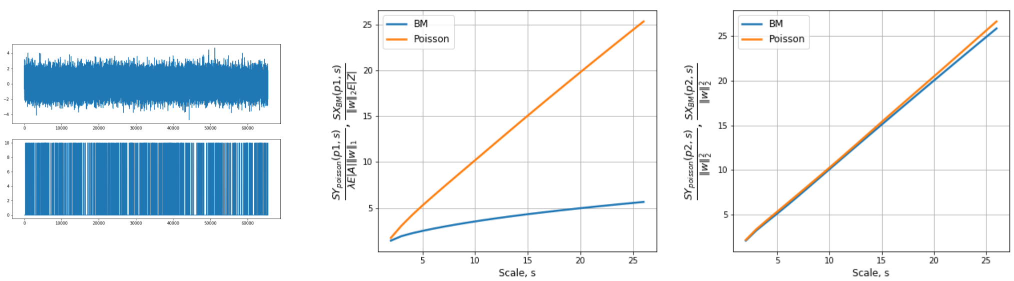

Figure 4: First-order invariant scattering moments for Brownian motion and Poisson point process. **Left:** *Top:* Brownian motion with Hurst parameter $H = 1/2$. *Bottom:* Ordinary Poisson point process. **Middle:** Normalized scattering moments for Brownian Motion ($SX_{BM}(s,\xi,1)/\|w\|_2 E|Z|$) and Poisson point process ($SY_{poisson}(s,\xi,1)/\lambda E|A|\|w\|_1$) at $p = 1$. This shows the normalized scattering to $\sqrt{s}$ for Brownian motion converge while to s for Poisson process, indicating the 1st moment can distinguish Brownian motion and Poisson point process. **Right:** Normalized scattering moments for Brownian Motion ($SX_{BM}(s,\xi,2)/\|w\|_2^2$) and Poisson point process ($SY_{poisson}(s,\xi,2)/\|w\|_2^2$) at $p = 2$. Both normalized scattering moments converge to s, so the 2nd moment scattering cannot distinguish the two processes.

References

Joakim Andén and Stéphane Mallat. Multiscale scattering for audio classification. In *Proceedings of the ISMIR 2011 conference*, pages 657–662, 2011.

Joakim Andén and Stéphane Mallat. Deep scattering spectrum. *IEEE Transactions on Signal Processing*, 62(16):4114–4128, August 2014.

Joakim Andén, Vincent Lostanlen, and Stéphane Mallat. Classification with joint time-frequency scattering. arXiv:1807.08869, 2018.

Joseph Antognini, Matt Hoffman, and Ron J. Weiss. Synthesizing diverse, high-quality audio textures. arXiv:1806.08002, 2018.

Rodrigo Bañuelos and Gang Wang. Sharp inequalities for martingales with applications to the beurling-ahlfors and riesz transforms. *Duke Math. J.*, 80(3):575–600, 12 1995.

Mikolaj Binkowski, Gautier Marti, and Philippe Donnat. Autoregressive convolutional neural networks for asynchronous time series. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 580–589, Stockholmsmässan, Stockholm Sweden, 10–15 Jul 2018. PMLR.

Antoine Brochard, Bartłomiej Błaszczyszyn, Stéphane Mallat, and Sixin Zhang. Statistical learning of geometric characteristics of wireless networks. arXiv:1812.08265, 2018.

Xavier Brumwell, Paul Sinz, Kwang Jin Kim, Yue Qi, and Matthew Hirn. Steerable wavelet scattering for 3D atomic systems with application to Li-Si energy prediction. In *NeurIPS Workshop on Machine Learning for Molecules and Materials*, 2018.

Joan Bruna and Stéphane Mallat. Classification with scattering operators. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1561–1566, 2011.

Joan Bruna and Stéphane Mallat. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1872–1886, August 2013.

Joan Bruna and Stéphane Mallat. Multiscale sparse microcanonical models. arXiv:1801.02013, 2018.

Joan Bruna, Stéphane Mallat, Emmanuel Bacry, and Jean-Francois Muzy. Intermittent process analysis with scattering moments. *Annals of Statistics*, 43(1):323 – 351, 2015.

Donald L. Burkholder. Sharp inequalities for martingales and stochastic integrals. In *Colloque Paul Lévy sur les processus stochastiques*, number 157-158 in Astérisque, pages 75–94. Société mathématique de France, 1988.

E. (Erhan) Cinlar. *Introduction to stochastic processes*. Prentice-Hall, Englewood Cliffs, N.J., 1975.

D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. ISBN 0-387-95541-0. Elementary theory and methods.

Michael Eickenberg, Georgios Exarchakis, Matthew Hirn, and Stéphane Mallat. Solid harmonic wavelet scattering: Predicting quantum molecular energy from invariant descriptors of 3D electronic densities. In *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, pages 6540–6549, 2017.

Michael Eickenberg, Georgios Exarchakis, Matthew Hirn, Stéphane Mallat, and Louis Thiry. Solid harmonic wavelet scattering for predictions of molecule properties. *Journal of Chemical Physics*, 148:241732, 2018.

Peter K Friz and Martin Hairer. *A course on rough paths: with an introduction to regularity structures*. Springer, 2014.

V. Fromion, E. Leoncini, and P. Robert. Stochastic gene expression in cells: A point process approach. *SIAM Journal on Applied Mathematics*, 73(1):195–211, 2013.

A. M. Garsia, E. Rodemich, and H. Rumsey Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana University Mathematics Journal*, 20(6):565–578, 1970.

Leon Gatys, Alexander S Ecker, and Matthias Bethge. Texture synthesis using convolutional neural networks. In *Advances in Neural Information Processing Systems 28*, pages 262–270, 2015.

Astrid Genet, Pavel Grabarnik, Olga Sekretenko, and David Pothier. Incorporating the mechanisms underlying inter-tree competition into a random point process model to improve spatial tree pattern analysis in forestry. *Ecological Modelling*, 288:143154, 09 2014.

Martin Haenggi, Jeffrey G. Andrews, François Baccelli, Olivier Dousse, and Massimo Franceschetti. Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE Journal on Selected Areas in Communications*, 27(7):1029–1046, 2009.

Matthew Hirn, Stéphane Mallat, and Nicolas Poilvert. Wavelet scattering regression of quantum chemical energies. *Multiscale Modeling and Simulation*, 15(2):827–863, 2017. arXiv:1605.04654.

Stéphane Mallat. Group invariant scattering. *Communications on Pure and Applied Mathematics*, 65(10):1331–1398, October 2012.

Edouard Oyallon and Stéphane Mallat. Deep roto-translation scattering for object classification. In *Proceedings in IEEE CVPR 2015 conference*, 2015. arXiv:1412.8659.

Frederic Paik Schoenberg. A note on the consistent estimation of spatial-temporal point process parameters. *Statistica Sinica*, 2016.

G. Shevchenko. Fractional Brownian motion in a nutshell. In *International Journal of Modern Physics Conference Series*, volume 36 of *International Journal of Modern Physics Conference Series*, page 1560002, January 2015.

Laurent Sifre and Stéphane Mallat. Combined scattering for rotation invariant texture analysis. In *Proceedings of the ESANN 2012 conference*, 2012.

Laurent Sifre and Stéphane Mallat. Rotation, scaling and deformation invariant scattering for texture discrimination. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2013.

Laurent Sifre and Stéphane Mallat. Rigid-motion scattering for texture classification. arXiv:1403.1687, 2014.

G. Wolf, S. Mallat, and S.A. Shamma. Audio source separation with time-frequency velocities. In *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Reims, France, 2014.

Guy Wolf, Stephane Mallat, and Shihab A. Shamma. Rigid motion model for audio source separation. *IEEE Transactions on Signal Processing*, 64(7):1822–1831, 2015.

L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, 67: 251–282, 1936.

Appendix A. Proof of Theorem 1

To prove Theorem 1 we will need the following lemma.

Lemma 7 *Let Z be a Poisson random variable with parameter λ. Then for all $\alpha \in \mathbb{R}$, $m \in \mathbb{N}$,*

$$\mathbb{E}\left[Z^\alpha \mathbb{1}_{\{Z>m\}}\right] = \sum_{k=m+1}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}k^\alpha \leq C_{m,\alpha}\lambda^{m+1}\,, \quad \forall\, 0<\lambda<1\,.$$

Proof For $0 < \lambda < 1$ and $k \in \mathbb{N}$, $e^{-\lambda}\lambda^k \leq 1$. Therefore,

$$\begin{aligned}\mathbb{E}\left[Z^\alpha \mathbb{1}_{\{Z>m\}}\right] &= \sum_{k=m+1}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}k^\alpha \\ &= \lambda^{m+1}\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{(k+m+1)!}(k+m+1)^\alpha \\ &\leq \lambda^{m+1}\sum_{k=0}^{\infty}\frac{(k+m+1)^\alpha}{(k+m+1)!} \\ &= C_{\alpha,m}\lambda^{m+1}\,.\end{aligned}$$

■

Proof [Theorem 1] Recalling the definitions of $Y(dt)$ and $S_{\gamma,p}Y(t)$, and setting $N_s(t) = N\left([t-s,t]^d\right)$, we see

$$\begin{aligned}S_{\gamma,p}Y(t) &= \mathbb{E}\left[\left|\int_{[s-t,t]^d} w\left(\frac{t-u}{s}\right)e^{i\xi\cdot(t-u)}\,Y(du)\right|^p\right] \\ &= \mathbb{E}\left[\left|\sum_{j=1}^{N_s(t)} A_j w\left(\frac{t-t_j}{s}\right)e^{i\xi\cdot(t-t_j)}\right|^p\right]\,,\end{aligned}$$

where $t_1, t_2, \dots t_{N_s(t)}$ are the points $N(t)$ in $[t-s,t]^d$. Conditioned on the event that $N_s(t) = k$, the locations of the k points on $[t-s,t]^d$ are distributed as i.i.d. random variables $Z_1, \dots, Z_k$ taking values in $[t-s,t]^d$ with density

$$p_Z(z) = \frac{\lambda(z)}{\Lambda_s(t)}, \quad z \in [t-s,t]^d .$$

Therefore, the random variables

$$V_i := \frac{t - Z_i}{s}$$

take values in the unit cube $Q_1 = [0,1]^d$ and have density

$$p_V(v) = \frac{s^d}{\Lambda_s(t)} \lambda(t - vs), \quad v \in Q_1 .$$

Note that in the special case that N is homogeneous, i.e. $\lambda(t) \equiv \lambda_0$ is constant, the V_i are uniform random variables on Q_1.

Our proof will be based off of conditioning on $N_s(t)$. For $N_s(t) = k \geq 1$,

$$\mathbb{E}\left[\left|\sum_{j=1}^{N_s(t)} A_j w\left(\frac{t-t_j}{s}\right) e^{i\xi\cdot(t-t_j)}\right|^p : N_s(t) = k\right] = \mathbb{E}\left[\left|\sum_{j=1}^{k} A_j w(V_j) e^{is\xi\cdot V_j}\right|^p\right]$$
$$\leq \frac{\|\lambda\|_\infty}{\lambda_{\min}} k^p \mathbb{E}[|A_1|^p] \|w\|_p^p , \tag{16}$$

where (16) follows from (i) the independence of the random variables A_j and V_j; (ii) the fact that for any sequence of i.i.d. random variables $Z_1, Z_2, \dots$,

$$\mathbb{E}\left[\left|\sum_{n=1}^{k} Z_n\right|^p\right] \leq k^{p-1} \mathbb{E}\left[\sum_{n=1}^{k} |Z_n|^p\right] = k^p \mathbb{E}[|Z_1|^p] ;$$

and (iii) the fact that

$$\mathbb{E}[|w(V_i)|^p] = \int_{Q_1} |w(v)|^p p_V(v)\, dv \leq \frac{\|\lambda\|_\infty}{\lambda_{\min}} \|w\|_p^p .$$

Therefore, since $\mathbb{P}[N_s(t) = k] = e^{-\Lambda_s(t)} \cdot {}^{(\Lambda_s(t))^k}/{}_{k!}$,

$$\begin{aligned}
&\mathbb{E}\left[\left|\sum_{j=1}^{N_s(t)} A_j w\left(\frac{t-t_j}{s}\right) e^{i\xi\cdot(t-t_j)}\right|^p\right] = \\
&= \sum_{k=0}^{\infty} e^{-\Lambda_s(t)} \frac{(\Lambda_s(t))^k}{k!} \mathbb{E}\left[\left|\sum_{j=1}^{N_s(t)} A_j w\left(\frac{t-t_j}{s}\right) e^{i\xi\cdot(t-t_j)}\right|^p : N_s(t) = k\right] \\
&= \sum_{k=1}^{\infty} e^{-\Lambda_s(t)} \frac{(\Lambda_s(t))^k}{k!} \mathbb{E}\left[\left|\sum_{j=1}^{k} A_j w(V_j) e^{is\xi\cdot V_j}\right|^p\right] \\
&= \sum_{k=1}^{m} e^{-\Lambda_s(t)} \frac{(\Lambda_s(t))^k}{k!} \mathbb{E}\left[\left|\sum_{j=1}^{k} A_j w(V_j) e^{is\xi\cdot V_j}\right|^p\right] + \epsilon(m,s,\xi,t) ,
\end{aligned}$$

where

$$\epsilon(m,s,t,\xi) := \sum_{k=m+1}^{\infty} e^{-\Lambda_s(t)} \frac{(\Lambda_s(t))^k}{k!} \mathbb{E}\left[\left|\sum_{j=1}^{k} A_j w(V_j) e^{is\xi\cdot V_j}\right|^p\right].$$

By (16) and Lemma 7, if s is small enough so that $\Lambda_s(t) \leq s^d \|\lambda\|_\infty < 1$, then:

$$\begin{aligned}
\epsilon(m,s,\xi,t) &= \sum_{k=m+1}^{\infty} e^{-\Lambda_s(t)} \frac{(\Lambda_s(t))^k}{k!} \mathbb{E}\left[\left|\sum_{j=1}^{k} A_j w(V_j) e^{is\xi\cdot V_j}\right|^p\right] \\
&\leq \frac{\|\lambda\|_\infty}{\lambda_{\min}} \mathbb{E}[|A_1|^p] \|w\|_p^p \sum_{k=m+1}^{\infty} e^{-\Lambda_s(t)} \frac{(\Lambda_s(t))^k}{k!} k^p \\
&\leq C_{m,p} \frac{\|\lambda\|_\infty}{\lambda_{\min}} \mathbb{E}[|A_1|^p] \|w\|_p^p (\Lambda_s(t))^{m+1} \\
&\leq C_{m,p} \frac{\|\lambda\|_\infty}{\lambda_{\min}} \mathbb{E}[|A_1|^p] \|w\|_p^p \|\lambda\|_\infty^{m+1} s^{d(m+1)}.
\end{aligned}$$

■

Appendix B. Proof of Theorem 2

Proof [Theorem 2] Let (s_k, ξ_k) be a sequence of scale and frequency pairs such that $\lim_{k\to\infty} s_k = 0$. Applying Theorem 1 with $m = 1$, we obtain:

$$\begin{aligned}
\frac{S_{\gamma_k,p} Y(t)}{s_k^d} &= e^{-\Lambda_{s_k}(t)} \frac{\Lambda_{s_k}(t)}{s_k^d} \mathbb{E}\left[\left|A_1 w(V_{1,k}) e^{is\xi\cdot V_{1,k}}\right|^p\right] + \frac{\epsilon(1,s_k,\xi_k,t)}{s_k^d} \\
&= e^{-\Lambda_{s_k}(t)} \frac{\Lambda_{s_k}(t)}{s_k^d} \mathbb{E}[|A_1|^p] \mathbb{E}[|w(V_{1,k})|^p] + \frac{\epsilon(1,s_k,\xi_k,t)}{s_k^d},
\end{aligned}$$

where we write $V_{1,k} = V_1$ to emphasize the fact that the density of $V_{1,k}$ is:

$$p_{V_k}(v) = \frac{s_k^d}{\Lambda_{s_k}(t)} \lambda(t - v s_k).$$

Using the error bound (9), we see that:

$$\lim_{k\to\infty} \frac{\epsilon(1,s_k,\xi_k,t)}{s_k^d} = 0.$$

Furthermore, since $0 \leq \Lambda_{s_k}(t) \leq s_k^d \|\lambda\|_\infty$, we observe that:

$$\lim_{k\to\infty} e^{-\Lambda_{s_k}(t)} = 1,$$

and by the continuity of $\lambda(t)$,

$$\lim_{k\to\infty} \frac{\Lambda_{s_k}(t)}{s_k^d} = \lim_{k\to\infty} \frac{1}{s_k^d} \int_{[s_k - t, t]^d} \lambda(u)\, du = \lambda(t). \quad (17)$$

Finally, by the continuity of $\lambda(t)$, we see that

$$p_{V_k}(v) \leq \frac{\|\lambda\|_\infty}{\lambda_{\min}} \quad \text{and} \quad \lim_{k\to\infty} p_{V_k}(v) = 1\,, \quad \forall\, v \in Q_1\,. \tag{18}$$

Therefore, by the bounded convergence theorem,

$$\lim_{k\to\infty} \mathbb{E}[|w(V_1)|^p] = \lim_{k\to\infty} \int_{Q_1} |w(v)|^p p_{V_k}(v)\,dv = \int_{Q_1} |w(v)|^p \lim_{k\to\infty} p_{V_k}(v)\,dv = \|w\|_p^p\,.$$

That completes the proof of (10).

To prove (11), we assume that $\lambda(t)$ is periodic with period T along each coordinate and again use Theorem 1 with $m = 1$ to observe,

$$\frac{SY(s_k, \xi_k, p)}{s_k^d} = \mathbb{E}[|A_1|^p]\frac{1}{T^d}\int_{Q_T} e^{-\Lambda_{s_k}(t)}\frac{\Lambda_{s_k}(t)}{s_k^d}\int_{Q_1}|w(v)|^p p_{V_k}(v)\,dv\,dt + \frac{1}{T^d}\int_{Q_1}\frac{\epsilon(1, s_k, \xi_k, t)}{s_k^d}\,dt\,.$$

By (9), the second integral converges to zero as $k \to \infty$. Therefore,

$$\lim_{k\to\infty}\frac{SY(s_k, \xi_k, p)}{s_k^d} = \mathbb{E}[|A_1|^p]\|w\|_p^p\frac{1}{T^d}\int_{Q_T}\lambda(t)\,dt\,,$$

by the continuity of $\lambda(t)$ and the bounded convergence theorem. ■

Appendix C. Proof of Theorem 3

Proof [Theorem 3] We apply Theorem 1 with $m = 2$ and obtain:

$$\begin{aligned} S_{\gamma_k,p}Y(t) &= e^{-\Lambda_{s_k}(t)}\Lambda_{s_k}(t)\mathbb{E}[|A_1|^p]\mathbb{E}[|w(V_{1,k})|^p] \\ &\quad + e^{-\Lambda_{s_k}(t)}\frac{(\Lambda_{s_k}(t))^2}{2}\mathbb{E}\left[\left|A_1 w(V_{1,k})e^{is_k\xi_k\cdot V_{1,k}} + A_2 w(V_{2,k})e^{is_k\xi_k\cdot V_{2,k}}\right|^p\right] + \epsilon(2, s_k, \xi_k, t)\,, \end{aligned} \tag{19}$$

where $V_{i,k}$, $i = 1, 2$, are random variables taking values on the unit cube $Q_1 = [0, 1]^d$ with densities,

$$p_{V_k}(v) = \frac{s_k^d}{\Lambda_{s_k}(t)}\lambda(t - vs_k)\,.$$

Dividing both sides in (19) by $s_k^{2d}\|w\|_p^p\mathbb{E}[|A_1|^p]$ and subtracting $\frac{\Lambda_{s_k}(t)}{s_k^{2d}}\frac{\mathbb{E}[|w(V_{1,k})|^p]}{\|w\|_p^p}$ yields:

$$\begin{aligned} \frac{S_{\gamma_k,p}Y(t)}{s_k^{2d}\|w\|_p^p\mathbb{E}[|A_1|^p]} - \frac{\Lambda_{s_k}(t)}{s_k^{2d}}\frac{\mathbb{E}[|w(V_{1,k})|^p]}{\|w\|_p^p} &= \frac{e^{-\Lambda_{s_k}(t)}\Lambda_{s_k}(t) - \Lambda_{s_k}(t)}{s_k^{2d}}\frac{\mathbb{E}[|w(V_{1,k})|^p]}{\|w\|_p^p} \\ &\quad + e^{-\Lambda_{s_k}(t)}\frac{(\Lambda_{s_k}(t))^2}{s_k^{2d}}\frac{\mathbb{E}\left[\left|A_1 w(V_{1,k})e^{is_k\xi_k\cdot V_{1,k}} + A_2 w(V_{2,k})e^{is_k\xi_k\cdot V_{2,k}}\right|^p\right]}{2\|w\|_p^p\mathbb{E}[|A_1|^p]} \\ &\quad + \frac{\epsilon(2, s_k, \xi_k, t)}{s_k^{2d}\|w\|_p^p\mathbb{E}[|A_1|^p]}\,. \end{aligned} \tag{20}$$

Using the error bound (9),

$$\lim_{k\to\infty} \frac{\epsilon(2, s_k, \xi_k, t)}{s_k^{2d}\|w\|_p^p \mathbb{E}[|A_1|^p]} = 0\,, \tag{21}$$

at a rate independent of t. Recalling (18) from the proof of Theorem 2, we use the fact that $\lim_{k\to\infty} p_{V_k} \equiv 1$ and the bounded convergence theorem to conclude,

$$\lim_{k\to\infty} \mathbb{E}\left[\left|A_1 w(V_{1,k})e^{is_k\xi_k\cdot V_{1,k}} + A_2 w(V_{2,k})e^{is_k\xi_k\cdot V_{2,k}}\right|^p\right] = \mathbb{E}\left[\left|A_1 w(U_1)e^{iL\cdot U_1} + A_2 w(U_2)e^{iL\cdot U_2}\right|^p\right]\,, \tag{22}$$

where U_i, $i = 1, 2$, are uniform random variables on the unit cube and $L = \lim_{k\to\infty} s_k\xi_k$. Similarly,

$$\lim_{k\to\infty} \frac{\mathbb{E}[|w(V_{1,k})|^p]}{\|w\|_p^p} = 1\,. \tag{23}$$

Lastly, recalling that $s_k \to 0$ as $k \to \infty$ and using (17) from the proof of Theorem 2, we see

$$\begin{aligned}
\lim_{k\to\infty} \frac{e^{-\Lambda_{s_k}(t)}\Lambda_{s_k}(t) - \Lambda_{s_k}(t)}{s_k^{2d}} &= \lim_{k\to\infty}\left(\frac{\Lambda_{s_k}(t)}{s_k^d}\right)\lim_{k\to\infty}\left(\frac{e^{-\Lambda_{s_k}(t)} - 1}{s_k^d}\right) \\
&= \lambda(t)\lim_{k\to\infty}\left(\frac{e^{-\Lambda_{s_k}(t)} - 1}{s_k^d}\right) \\
&= -\lambda(t)^2\,.
\end{aligned} \tag{24}$$

Now we integrate both sides of (20) over Q_T and divide by T^d. Taking the limit as $k \to \infty$, on the left hand side we get:

$$\begin{aligned}
&\lim_{k\to\infty} \frac{1}{T^d}\int_{Q_T}\left(\frac{S_{\gamma_k,p}Y(t)}{s_k^{2d}\|w\|_p^p\mathbb{E}[|A_1|^p]} - \frac{\Lambda_{s_k}(t)}{s_k^{2d}}\frac{\mathbb{E}[|w(V_{1,k})|^p]}{\|w\|_p^p}\right)dt \\
&\quad= \lim_{k\to\infty}\left(\frac{SY(s_k,\xi_k,p)}{s_k^{2d}\|w\|_p^p\mathbb{E}[|A_1|^p]} - \frac{\mathbb{E}[|w(V_{1,k})|^p]}{\|w\|_p^p}\frac{1}{T^d}\int_{Q_T}\frac{\Lambda_{s_k}(t)}{s_k^{2d}}\,dt\right) \\
&\quad= \lim_{k\to\infty}\left(\frac{SY(s_k,\xi_k,p)}{s_k^{2d}\mathbb{E}[|w(V_{1,k})|^p]\mathbb{E}[|A_1|^p]} - \frac{1}{T^d}\int_{Q_T}\frac{\Lambda_{s_k}(t)}{s_k^{2d}}\,dt\right)\,,
\end{aligned}$$

where we used the definition of the invariant scattering moments and (23). On the right hand side of (20), we use (23), (24) and the dominated convergence theorem to see that the first term is:

$$\begin{aligned}
\lim_{k\to\infty}\frac{1}{T^d}\int_{Q_T}\frac{e^{-\Lambda_{s_k}(t)}\Lambda_{s_k}(t) - \Lambda_{s_k}(t)}{s_k^{2d}}\frac{\mathbb{E}[|w(V_{1,k})|^p]}{\|w\|_p^p}\,dt &= \lim_{k\to\infty}\frac{1}{T^d}\int_{Q_T}\frac{e^{-\Lambda_{s_k}(t)}\Lambda_{s_k}(t) - \Lambda_{s_k}(t)}{s_k^{2d}}\,dt \\
&= -\frac{1}{T^d}\int_{Q_T}\lambda(t)^2\,dt\,.
\end{aligned}$$

Using (17), (22), and the bounded convergence theorem, the second term of (20) is:

$$\begin{aligned}
&\lim_{k\to\infty}\frac{1}{T^d}\int_{Q_T} e^{-\Lambda_{s_k}(t)}\frac{(\Lambda_{s_k}(t))^2}{s_k^{2d}}\frac{\mathbb{E}\left[\left|A_1 w(V_{1,k})e^{is_k\xi_k\cdot V_{1,k}} + A_2 w(V_{2,k})e^{is_k\xi_k\cdot V_{2,k}}\right|^p\right]}{2\|w\|_p^p\mathbb{E}[|A_1|^p]}\,dt \\
&\quad= \frac{\mathbb{E}[|A_1 w(U_1)e^{iL\cdot U_1} + A_2 w(U_2)e^{iL\cdot U_2}|^p]}{2\|w\|_p^p\mathbb{E}[|A_1|^p]}\left(\frac{1}{T^d}\int_{Q_T}\lambda(t)^2\,dt\right)\,.
\end{aligned}$$

Finally, the third term of (20) goes to zero using the bounded convergence theorem and (21). Putting together the left and right hand sides of (20) with these calculations finishes the proof. ∎

Appendix D. Proof of Theorem 4

Proof [Theorem 4] As in the proof of Theorem 1, let $N_s(t) = N\left([t-s,t]^d\right)$ denote the number of points in the cube $[t-s,t]^d$. Then since the support of w is contained in $[0,1]^d$,

$$(g_{\gamma_k} * Y)(t) = \int_{[t-s_k,t]^d} w\left(\frac{t-u}{s_k}\right) e^{i\xi_k\cdot(t-u)}\, Y(du) = \sum_{j=1}^{N_{s_k}(t)} A_j w\left(\frac{t-t_j}{s_k}\right) e^{i\xi_k\cdot(t-t_j)}\,,$$

where $t_1, t_2, \ldots, t_{N_{s_k}(t)}$ are the points of N in $[t-s_k,t]^d$. Therefore, in the event that $N_{s_k}(t) = 1$,

$$|\,(g_{\gamma_k} * Y)(t)|^p = (|g_{\gamma_k}|^p * |Y|^p)(t)\,,$$

and so, partitioning the space of possible outcomes based on $N_{s_k}(t)$, we obtain:

$$\begin{aligned}
|\,(g_{\gamma_k} * Y)(t)|^p &= |\,(g_{\gamma_k} * Y)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)=1\}} + (g_{\gamma_k} * Y)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)>1\}}|^p \\
&= |\,(g_{\gamma_k} * Y)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)=1\}}|^p + |\,(g_{\gamma_k} * Y)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)>1\}}|^p \\
&= (|g_{\gamma_k}|^p * |Y|^p)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)=1\}} + |\,(g_{\gamma_k} * Y)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)>1\}}|^p \\
&= (|g_{\gamma_k}|^p * |Y|^p)(t) + |\,(g_{\gamma_k} * Y)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)>1\}}|^p - (|g_{\gamma_k}|^p * |Y|^p)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)>1\}} \\
&= (|g_{\gamma_k}|^p * |Y|^p)(t) + e_k(t)\,,
\end{aligned}$$

where

$$e_k(t) := |\,(g_{\gamma_k} * Y)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)>1\}}|^p - (|g_{\gamma_k}|^p * |Y|^p)(t)\cdot \mathbb{1}_{\{N_{s_k}(t)>1\}}$$

Using the above, we can write the second order convolution term as:

$$\left(g_{\gamma'_k} * |g_{\gamma_k} * Y|\right)(t) = \left(g_{\gamma'_k} * |g_{\gamma_k}|^p * |Y|^p\right)(t) + \left(g_{\gamma'_k} * e_k\right)(t)\,.$$

The following lemma implies that $\left(g_{\gamma'_k} * e_k\right)(t)$ decays rapidly in $\mathbf{L}^{p'}$ at a rate independent of t.

Lemma 8 *There exists $\delta > 0$, independent of t, such that if $s_k < \delta$,*

$$\mathbb{E}\left[\left|\left(g_{\gamma'_k} * e_k\right)(t)\right|^p\right] \le C(p,p',w,c,L)\frac{\|\lambda\|_\infty}{\lambda_{\min}}\|\lambda\|_\infty^2 s_k^{d(p'+2)}\,.$$

Once we have proved Lemma 8, equation (13) will follow once we show,

$$\lim_{k\to\infty} \frac{\mathbb{E}\left[\left|\left(g_{\gamma'_k} * |g_{\gamma_k}|^p * |Y|^p\right)(t)\right|^{p'}\right]}{s_k^{d(p'+1)}} = K(p,p',w,c,L)\lambda(t)\mathbb{E}[|A_1|^q]\,. \tag{25}$$

Let us prove (25) first and postpone the proof of Lemma 8. We will use the fact that the support of $g_{\gamma'_k} * |g_{\gamma_k}|^p$ is contained in $[0, s_k + s'_k]^d$. Let $\tilde{s}_k := s_k + s'_k$, $N_k(t) := N_{\tilde{s}_k}(t)$, $\Lambda_k(t) := \Lambda_{\tilde{s}_k}(t)$, and let $t_1, t_2, \ldots, t_{N_k(t)}$ be the points of N in the cube $[t-\tilde{s}_k, t]^d$. We have that $\mathbb{P}[N_k(t) = n] = e^{-\Lambda_k(t)}\frac{(\Lambda_k(t))^n}{n!}$, and conditioned on the event that $N_k(t) = n$, the locations of the points $t_1, \ldots, t_n$ are distributed as i.i.d. random variables $Z_1(t), \ldots, Z_n(t)$ taking values in $[t-\tilde{s}_k, t]^d$

with density $p_{Z(t)}(z) = \frac{\lambda(z)}{\Lambda_k(t)}$. Therefore the i.i.d. random variables $\widetilde{V}_1(t), \dots, \widetilde{V}_n(t)$ defined by $\widetilde{V}_i(t) := t - Z_i(t)$ take values in $[0, \tilde{s}_k]^d$ and have density

$$p_{\widetilde{V}(t)}(v) = \frac{\lambda(t-v)}{\Lambda_k(t)}\,, \quad v \in [0, \tilde{s}_k]^d\,.$$

Now, we condition on $N_k(t)$ to see that

$$\begin{aligned}
\mathbb{E}\left[\left|\left(g_{\gamma_k'} * |g_{\gamma_k}|^p * |Y|^p\right)(t)\right|^{p'}\right] &= \mathbb{E}\left[\left|\sum_{j=1}^{N_k(t)} |A_j|^p \left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(t-t_j)\right|^{p'}\right] \\
&= \sum_{n=1}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} \mathbb{E}\left[\left|\sum_{j=1}^{N_k(t)} |A_j|^p \left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(t-t_j)\right|^{p'} : N_k(t) = n\right] \\
&= \sum_{n=1}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} \mathbb{E}\left[\left|\sum_{j=1}^{n} |A_j|^p \left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_j(t))\right|^{p'}\right] \\
&= e^{-\Lambda_k(t)} \Lambda_k(t) \mathbb{E}[|A_1|^q] \mathbb{E}\left[\left|\left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_1(t))\right|^{p'}\right] \qquad (26) \\
&\quad + \sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} \mathbb{E}\left[\left|\sum_{j=1}^{n} |A_j|^p \left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_j(t))\right|^{p'}\right]. \qquad (27)
\end{aligned}$$

The following lemma will be used to estimate the scaling of the term in (26).

Lemma 9 *For all $t \in \mathbb{R}^d$,*

$$\lim_{k\to\infty} \frac{\tilde{s}_k^d}{s_k^{d(p'+1)}} \mathbb{E}\left[\left|\left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_1(t))\right|^{p'}\right] = \|g_{c,L/c} * |g_{1,0}|^p\|_{p'}^{p'}\,. \qquad (28)$$

Furthermore, there exists $\delta > 0$, independent of t, such that if $s_k < \delta$ then

$$\frac{\tilde{s}_k^d}{s_k^{d(p'+1)}} \mathbb{E}\left[\left|\left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_1(t))\right|^{p'}\right] \le 2\frac{\|\lambda\|_\infty}{\lambda_{\min}} C(p, p', w, c, L)\,. \qquad (29)$$

Proof Making a change of variables in both u and v, and recalling the assumption that $s'_k = cs_k$, we observe that

$$\begin{aligned}
&\frac{\tilde{s}_k^d}{s_k^{d(p'+1)}}\mathbb{E}\left[\left|\left(g_{\gamma'_k} * |g_{\gamma_k}|^p\right)(\widetilde{V}_1(t))\right|^{p'}\right] \\
&\quad= \frac{\tilde{s}_k^d}{s_k^{d(p'+1)}}\int_{\mathbb{R}^d} p_{\widetilde{V}(t)}(v)\left|\int_{\mathbb{R}^d} w\left(\frac{v-u}{s'_k}\right)e^{i\xi'_k\cdot(v-u)}\left|w\left(\frac{u}{s_k}\right)\right|^p du\right|^{p'} dv \\
&\quad= \tilde{s}_k^d\int_{\mathbb{R}^d} p_{\widetilde{V}(t)}(s_k v)\left|\int_{\mathbb{R}^d} w\left(\frac{s_k(v-u)}{s'_k}\right)e^{is_k\xi'_k\cdot(v-u)}|w(u)|^p\, du\right|^{p'} dv \\
&\quad= \int_{\mathbb{R}^d}\frac{\tilde{s}_k^d\lambda(t-s_kv)}{\Lambda_k(t)}\left|\int_{\mathbb{R}^d} w\left(\frac{u-v}{c}\right)e^{is'_k\xi'_k\cdot(u-v)/c}|w(u)|^p\, du\right|^{p'} dv\,. \qquad (30)
\end{aligned}$$

The continuity of $\lambda(t)$ implies that

$$\lim_{k\to\infty}\frac{\tilde{s}_k^d\lambda(t-s_kv)}{\Lambda_k(t)} = 1\,, \quad \forall\, v\in[0,1+c]^d\,.$$

Furthermore, the assumption $0<\lambda_{\min}\leq\|\lambda\|_\infty<\infty$ implies

$$\frac{\tilde{s}_k^d\lambda(t-s_kv)}{\Lambda_k(t)} \leq \frac{\|\lambda\|_\infty}{\lambda_{\min}}\,, \quad \forall\, k\geq 1\,. \qquad (31)$$

Therefore, (28) follows from the dominated convergence theorem and by the observation that the inner integral of (30) is zero unless $v\in[0,1+c]^d$. Equation (29) follows from inserting (31) into (30) and sending k to infinity. ■

Since

$$\lim_{k\to\infty}\frac{\Lambda_k(t)}{\tilde{s}_k^d} = \lambda(t)\,,$$

the independence of $\widetilde{V}_1(t)$ and A_1, the continuity of $\lambda(t)$, and Lemma 9 imply that taking $k\to\infty$ in (26) yields:

$$\begin{aligned}
&\lim_{k\to\infty}\left(\frac{e^{-\Lambda_k(t)}\Lambda_k(t)\mathbb{E}[|A_1|^q]\mathbb{E}\left[|g_{\gamma'_k} * |g_{\gamma_k}|^p(\widetilde{V}_1(t))|^{p'}\right]}{s_k^{d(p'+1)}}\right) \\
&\quad= \lim_{k\to\infty}\left(e^{-\Lambda_k(t)}\frac{\Lambda_k(t)}{\tilde{s}_k^d}\mathbb{E}[|A_1|^q]\frac{\tilde{s}_k^d}{s_k^{d(p'+1)}}\mathbb{E}\left[|g_{\gamma'_k} * |g_{\gamma_k}|^p(\widetilde{V}_1(t))|^{p'}\right]\right) \\
&\quad= K(p,p',c,w,L)\lambda(t)\mathbb{E}[|A_1|^q]\,.
\end{aligned}$$

The following lemma shows that (27) is $O\left(s_k^{d(p'+2)}\right)$ (and converges at a rate independent of t), and therefore completes the proof of (13) subject to proving Lemma 8.

Lemma 10 *For all $\alpha \in \mathbb{R}$ there exists $\delta > 0$, independent of t, such that if $s_k < \delta$, then*

$$\sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} n^\alpha \mathbb{E}\left[\left|\sum_{j=1}^{n} |A_j|^p \left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_j(t))\right|^{p'}\right] \leq C(p,p',w,c,\alpha,L) \frac{\|\lambda\|_\infty}{\lambda_{\min}} \|\lambda\|_\infty^2 \mathbb{E}[|A_1|^q] s_k^{d(p'+2)} .$$

Proof For any sequence of i.i.d. random variables, $Z_1, Z_2, \ldots$, it holds that

$$\mathbb{E}\left[\left|\sum_{n=1}^{k} Z_n\right|^p\right] \leq k^{p-1} \mathbb{E}\left[\sum_{n=1}^{k} |Z_n|^p\right] = k^p \mathbb{E}\left[|Z_1|^p\right].$$

Therefore, by Lemma 7, Lemma 9, and the fact that the $\widetilde{V}_j(t)$ and A_i are i.i.d. and independent of each other, we see that if $s_k < \delta$, where δ is as in (29),

$$\begin{aligned}
&\sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} n^\alpha \mathbb{E}\left[\left|\sum_{j=1}^{n} |A_i|^p \left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_j(t))\right|^{p'}\right] \\
\leq &\sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} n^\alpha n^{p'} \mathbb{E}\left[|A_1|^q \left|\left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_1(t))\right|^{p'}\right] \\
= &\sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} n^{p'+\alpha} \mathbb{E}[|A_1|^q] \mathbb{E}\left[\left|\left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_1(t))\right|^{p'}\right] \\
= &\mathbb{E}[|A_1|^q] \mathbb{E}\left[\left|\left(g_{\gamma_k'} * |g_{\gamma_k}|^p\right)(\widetilde{V}_1(t))\right|^{p'}\right] \sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} n^{p'+\alpha} \\
\leq &C(p,p',w,c,L) \frac{\|\lambda\|_\infty}{\lambda_{\min}} \mathbb{E}[|A_1|^q] \frac{s_k^{d(p'+1)}}{\tilde{s}_k^d} \sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} n^{p'+\alpha} \\
\leq &C(p,p',w,c,L,\alpha) \frac{\|\lambda\|_\infty}{\lambda_{\min}} \mathbb{E}[|A_1|^q] \frac{s_k^{d(p'+1)}}{\tilde{s}_k^d} (\Lambda_k(t))^2 \\
\leq &C(p,p',w,c,L,\alpha) \frac{\|\lambda\|_\infty}{\lambda_{\min}} \|\lambda\|_\infty^2 \mathbb{E}[|A_1|^q] s_k^{d(p'+2)},
\end{aligned}$$

where the last inequality uses the fact that $\Lambda_k(t) \leq \tilde{s}_k^d \|\lambda\|_\infty = (1+c)^d s_k^d \|\lambda\|_\infty$. ■

We will now complete the proof of the theorem by proving Lemma 8.

Proof [Lemma 8] Since

$$e_k(t) = |\left(g_{\gamma_k} * Y\right)(t) \mathbb{1}_{\{N_{s_k}(t) > 1\}}|^p - \left(|g_{\gamma_k}|^p * |Y|^p\right)(t) \mathbb{1}_{\{N_{s_k}(t) > 1\}},$$

we see that

$$\left|g_{\gamma_k'} * e_k(t)\right| \leq \left|g_{\gamma_k'} * \left(\left|\left(g_{\gamma_k} * Y\right) \mathbb{1}_{\{N_{s_k}(\cdot) > 1\}}\right|^p\right)(t)\right| + \left|g_{\gamma_k'} * \left(\left(|g_{\gamma_k}|^p * |Y|^p\right) \mathbb{1}_{\{N_{s_k}(\cdot) > 1\}}\right)(t)\right|.$$

First turning our attention to the second term, we note that

$$\begin{aligned}
&\left|g_{\gamma_k'} * \left(\left(|g_{\gamma_k}|^p * |Y|^p\right) \mathbb{1}_{\{N_{s_k}(\cdot)>1\}}\right)(t)\right| \\
&\quad = \left|\int_{[t-s_k',t]^d} w\left(\frac{t-u}{s_k'}\right) e^{i\xi_k'\cdot(t-u)} \left(|g_{\gamma_k}|^p * |Y|^p\right)(u) \mathbb{1}_{\{N_{s_k}(u)>1\}}\, du\right| \\
&\quad \leq \mathbb{1}_{\{N_k(t)>1\}} \int_{[t-s_k',t]^d} w\left(\frac{t-u}{s_k'}\right) \left(|g_{\gamma_k}|^p * |Y|^p\right)(u)\, du \\
&\quad = \mathbb{1}_{\{N_k(t)>1\}} \left(g_{s_k',0} * |g_{\gamma_k}|^p * |Y|^p\right)(t). \qquad (32)
\end{aligned}$$

since $N_{s_k}(u) \leq N_{s_k+s_k'}(t) = N_{\tilde{s}_k}(t) = N_k(t)$ for all $u \in [t-s_k', t]^d$. Therefore, conditioning on $N_k(t)$, if $s_k < \delta$,

$$\begin{aligned}
&\mathbb{E}\left[\left|g_{\gamma_k'} * \left(\left(|g_{\gamma_k}|^p * |Y|^p\right) \mathbb{1}_{\{N_{s_k}(\cdot)>1\}}\right)(t)\right|^{p'}\right] \\
&\quad \leq \mathbb{E}\left[\left|\mathbb{1}_{\{N_k(t)>1\}} \left(g_{s_k',0} * |g_{\gamma_k}|^p * |Y|^p\right)(t)\right|^{p'}\right] \\
&\quad = \sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^n}{n!} \mathbb{E}\left[\left|\sum_{j=1}^{n} |A_j|^p \left(g_{s_k',0} * |g_{\gamma_k}|^p\right)(\widetilde{V}_j(t))\right|^{p'}\right] \\
&\quad \leq C(p,p',w,c,L) \frac{\|\lambda\|_\infty}{\lambda_{\min}} \|\lambda\|_\infty^2 \mathbb{E}[|A_1|^q] s_k^{d(p'+2)}
\end{aligned}$$

by Lemma 10. Now, turning our attention to the first term, note that

$$|(g_{\gamma_k} * Y)(t)|^p \, \mathbb{1}_{\{N_{s_k}(t)>1\}} \leq N_{s_k}(t)^{p-1} \left(|g_{\gamma_k}|^p * |Y|^p\right)(t) \mathbb{1}_{\{N_{s_k}(t)>1\}}.$$

Therefore, by the same logic as in (32)

$$\begin{aligned}
&\left|g_{\gamma_k'} * \left(\left|(g_{\gamma_k} * Y)\, \mathbb{1}_{\{N_{s_k}(\cdot)>1\}}\right|^p\right)(t)\right| \\
&\quad \leq \int_{[t-s_k',t]^d} w\left(\frac{t-u}{s_k'}\right) N_{s_k}(u)^{p-1} \left(|g_{\gamma_k}|^p * |Y|^p\right)(u) \mathbb{1}_{\{N_{s_k}(u)>1\}}\, du \\
&\quad \leq \mathbb{1}_{\{N_k(t)>1\}} N_k(t)^{p-1} \int_{[t-s_k',t]^d} w\left(\frac{t-u}{s_k'}\right) \left(|g_{\gamma_k}|^p * |Y|^p\right)(u)\, du \\
&\quad \leq \mathbb{1}_{\{N_k(t)>1\}} N_k(t)^{p-1} \left(g_{s_k',0} * \left(|g_{\gamma_k}|^p * |Y|^p\right)\right)(t).
\end{aligned}$$

So again conditioning on $N_k(t)$, and applying Lemma 10, we see that if $s_k < \delta$

$$\begin{aligned}
&\mathbb{E}\left[\left|g_{\gamma_k'} * \left(\left|(g_{\gamma_k} * Y)\, \mathbb{1}_{\{N_{s_k}(\cdot)>1\}}\right|^p\right)(t)\right|^{p'}\right] \\
&\quad \leq \sum_{n=2}^{\infty} e^{-\Lambda_k(t)} \frac{(\Lambda_k(t))^k}{n!} n^{p-1} \mathbb{E}\left[\left|\sum_{j=1}^{n} |A_j|^p \left|\left(g_{s_k',0} * |g_{\gamma_k}|\right)(\widetilde{V}_j(t))\right|^p\right|^{p'}\right] \\
&\quad \leq C(p,p',w,c,L) \frac{\|\lambda\|_\infty}{\lambda_{\min}} \|\lambda\|_\infty^2 \mathbb{E}[|A_1|^q] s_k^{d(p'+2)}.
\end{aligned}$$

■

This completes the proof of (13). Line (14) follows from integrating with respect to t, observing that the error bounds in Lemmas 8 and 9 are independent of t, and applying the bounded convergence theorem.

■

Appendix E. Proofs of Results from Section 5

In order to prove Theorems 5 and 6, we will need the following lemma which shows that the scaling relationship of a self-similar process $X(t)$induces a similar relationship on stochastic integrals against $dX(t)$.

Lemma 11 *If X is a stochastic process that satisfies the scaling relation*

$$X(st) \stackrel{d}{=} s^{\beta} X(t) \tag{33}$$

for some $\beta > 0$, then for any measurable function $f : \mathbb{R} \to \mathbb{R}$,

$$\int_0^s f(u)\, dX(u) \stackrel{d}{=} s^{\beta} \int_0^1 f(su)\, dX(u)\,.$$

Proof Let $X = (X(t))_{t\in\mathbb{R}}$ be a stochastic process satisfying (33), and let $\mathcal{P}_n = \{0 = t_0^n < t_1^n < \ldots < t_{K_n}^n = 1\}$ be a sequence of partitions of $[0, 1]$ such that

$$\lim_{n\to\infty} \max_k \left\{ |t_k^n - t_{k-1}^n| \right\} = 0.$$

Then, by the scaling relation (33),

$$\begin{aligned}\int_0^s f(u)\, dX(u) &= \lim_{n\to\infty} \sum_{k=0}^{K_n-1} f(st_k^n) \left(X(st_{k+1}^n) - X(st_k^n) \right) \\ &\stackrel{d}{=} s^{\beta} \lim_{n\to\infty} \sum_{k=0}^{K_n-1} f(st_k^n) \left(X(t_{k+1}^n) - X(t_k^n) \right) = s^{\beta} \int_0^1 f(su)\, dX(u)\,.\end{aligned}$$

■

We will now use Lemma 11 to prove Theorems 5 and 6.

Proof [Theorem 5] Let $X = (X(t))_{t\in\mathbb{R}}$ be the α-stable process, $p < \alpha \le 2$. Since X has stationary increments, its scattering coefficients do not depend on t and it suffices to analyze

$$\mathbb{E}\left[|(g_{\gamma_k} * dX)(0)|^p\right] = \mathbb{E}\left[\left|\int_{-s_k}^0 g_{\gamma_k}(u)\, dX(u)\right|^p\right] = \mathbb{E}\left[\left|\int_0^{s_k} g_{\gamma_k}(u)\, dX(u)\right|^p\right],$$

where the second equality uses the fact the distribution of X does not change if it is run in reverse, i.e.

$$(X(t))_{t\in\mathbb{R}} \stackrel{d}{=} (X(-t))_{t\in\mathbb{R}}$$

It is well known that $X(t)$ satisfies (33) for $\beta = 1/\alpha$. Therefore, by Lemma 11

$$\mathbb{E}\left[|(g_{\gamma_k} * dX)(0)|^p\right] = \mathbb{E}\left[\left|\int_0^{s_k} w\left(\frac{u}{s_k}\right) e^{i\xi_k u}\, dX(u)\right|^p\right] = s_k^{p/\alpha}\mathbb{E}\left[\left|\int_0^1 w(u) e^{i\xi_k s_k u}\, dX(u)\right|^p\right].$$

So,

$$\frac{\mathbb{E}\left[|(g_{\gamma_k} * dX)(0)|^p\right]}{s_k^{p/\alpha}} = \mathbb{E}\left[\left|\int_0^1 w(u) e^{i\xi_k s_k u}\, dX(u)\right|^p\right].$$

The proof will be complete as soon as we show that

$$\lim_{k\to\infty}\left(\mathbb{E}\left[\left|\int_0^1 w(u) e^{i\xi_k s_k u}\, dX(u)\right|^p\right]\right)^{1/p} = \left(\mathbb{E}\left[\left|\int_0^1 w(u) e^{iLu}\, dX(u)\right|^p\right]\right)^{1/p}.$$

By the triangle inequality,

$$\begin{aligned}&\left|\left(\mathbb{E}\left[\left|\int_0^1 w(u) e^{i\xi_k s_k u}\, dX(u)\right|^p\right]\right)^{1/p} - \left(\mathbb{E}\left[\left|\int_0^1 w(u) e^{iLu}\, dX(u)\right|^p\right]\right)^{1/p}\right| \\ &\leq \left(\mathbb{E}\left[\left|\int_0^1 w(u)\left(e^{i\xi_k s_k u} - e^{iLu}\right)\, dX(u)\right|^p\right]\right)^{1/p}.\end{aligned}$$

Since $1 \leq p < \alpha$, we may choose p' strictly greater than 1 such that $p \leq p' < \alpha$, and note that by Jensen's inequality

$$\left(\mathbb{E}\left[\left|\int_0^1 w(u)\left(e^{i\xi_k s_k u} - e^{iLu}\right)\, dX(u)\right|^p\right]\right)^{1/p} \leq \left(\mathbb{E}\left[\left|\int_0^1 w(u)\left(e^{i\xi_k s_k u} - e^{iLu}\right)\, dX(u)\right|^{p'}\right]\right)^{1/p'},$$

and since $X(t)$ is a p'-integrable martingale, the boundedness of martingale transforms (Burkholder, 1988) (see also Bañuelos and Wang (1995)) implies

$$\begin{aligned}&\left(\mathbb{E}\left[\left|\int_0^1 w(u)\left(e^{i\xi_k s_k u} - e^{iLu}\right)\, dX(u)\right|^{p'}\right]\right)^{1/p'} \\ &\leq C_{p'} \sup_{0\leq u\leq 1}\left|w(u)\left(e^{i\xi_k s_k u} - e^{iLu}\right)\right| \mathbb{E}\left[|X_1|^{p'}\right] \leq C_{p'}|s_k\xi_k - L|\|w\|_\infty \mathbb{E}\left[|X_1|^{p'}\right],\end{aligned}$$

which converges to zero by the continuity of w on $[0, 1]$ and the assumption that $s_k\xi_k$ converges to L. ■

Proof [Theorem 6] Similarly to the proof of Theorem 5, it suffices to show that if a $(X(t))_{t\in\mathbb{R}}$ is fractional Brownian motion with Hurst parameter H, then

$$\lim_{k\to\infty}\left(\mathbb{E}\left[\left|\int_0^1 w(u)\left(e^{i\xi_k s_k u} - e^{iLu}\right)\, dX(u)\right|^p\right]\right)^{1/p} = 0.$$

However, fractional Brownian motion is not a semi-martingale so we cannot apply Burkholder's theorem as we did in the proof of Theorem 5. Instead, we use the following result first established in

Young (1936) (see also Friz and Hairer (2014), p. 48) which states that if $x(u)$ is any (deterministic) function with bounded variation, and $y(u)$ is any function which is α-Hölder continuous, $0 < \alpha < 1$, then

$$\int_0^1 x(u)\, dy(u)$$

is well-defined as the limit of Riemann sums and

$$\left| \int_0^1 x(u)\, dy(u) - x(0)\,(y(1) - y(0)) \right| \le C_\alpha \|x\|_{BV} \|y\|_\alpha \,,$$

where $\|\cdot\|_{BV}$ and $\|\cdot\|_\alpha$ are the bounded variation and α-Hölder seminorms respectively. For all k, the function $h_k(u) \coloneqq w(u)\left(e^{i\xi_k s_k u} - e^{iLu}\right) \coloneqq w(u) f_k(u)$ satisfies, $h_k(0) = 0$ and

$$\|h_k\|_{BV} \le \|w\|_\infty \|f_k\|_{BV} + \|w\|_{BV} \|f_k\|_\infty \,.$$

One can check that the fact that $s_k \xi_k$ converges to L implies that f_k converges to zero in both $\mathbf{L}^\infty$ and in the bounded variation seminorm, and that therefore that $\|h_k\|_{BV}$ converges to zero.

It is well-known that fractional Brownian motion with Hurst parameter H admits a continuous modification which is α-Hölder continuous for any $\alpha < H$. Therefore,

$$\mathbb{E}\left[\left| \int_0^1 w(u) \left(e^{i\xi_k s_k u} - e^{iLu} \right) \, dX(u) \right|^p \right] \le C_\alpha^p \|h_k\|_{BV}^p \mathbb{E}\left[\|X\|_\alpha^p \right] \,.$$

Lastly, one can use the Garsia-Rodemich-Rumsey inequality Garsia et al. (1970), to show that

$$\mathbb{E}[\|X\|_\alpha^p] < \infty \,.$$

for all $1 < p < \infty$. For details we refer the reader to the survey article Shevchenko (2015). Therefore,

$$\lim_{k \to 0} \mathbb{E}\left[\left| \int_0^1 w(u) \left(e^{i\xi_k s_k u} - e^{iLu} \right) \, dX(u) \right|^p \right] = 0$$

as desired.

■

Remark 12 *The assumption that w has bounded-variation was used to justify that the stochastic integral against fractional Brownian motion was well defined as the limit of Riemann sums because of its Hölder continuity and the above mentioned result of Young (1936). This allowed us to avoid the technical complexities of defining such an integral using either the Malliavin calculus or the Wick product.*

Appendix F. Details of Numerical Experiments

F.1. Definition of Filters

For all the numerical experiments, we take the window function w to be the smooth bump function

$$w(t) = \begin{cases} \exp\left(-\frac{1}{4t - 4t^2}\right), & t \in (0,1) \\ 0, & \text{otherwise.} \end{cases}$$

Therefore for $\gamma = (s, \xi)$, our filters are given by

$$g_\gamma(t) = e^{i\xi t} w(t) = \begin{cases} e^{i\xi t} e^{-s^2/(4ts-4t^2)}, & t \in (0, s) \\ 0, & \text{otherwise} \end{cases} .$$

F.2. Frequencies

In all of our experiments, we hold the frequency, ξ, which we sample uniformly at random from $(0, 2\pi)$, constant while allowing the scale to decrease to zero.

F.3. Simulation of Poisson point process

We use the standard method to generate a realization of a Poisson point process. For Poisson point process with intensity λ, the time interval between two neighbor jumps follows exponential distribution:

$$\Delta_j \coloneqq t_j - t_{j-1} \sim \text{Exp}(\lambda) .$$

Therefore, taking the inverse cumulative distribution function, we sample the time interval between two neighbor jumps through:

$$\Delta_j = -\frac{\log U_j}{\lambda},$$

where U_j are i.i.d. uniform random variables on $[0, 1]$, and assign the charge A_j to the jump at location t_j.

For inhomogeneous Poisson process with intensity funciton $\lambda(t)$, we simulate the time interval based on the proposition from Cinlar (1975). First define the cumulated intensity:

$$\Lambda(t) = \int_0^t \lambda(s) ds ,$$

then generate the location of jumps t_j by the following algorithm:

initialize $V = 0, t = 0$
while $t < N$ **do**
    generate $U \sim \mathcal{U}([0, 1])$
    $V \leftarrow V - \log U$
    $t = \inf\{v : \Lambda(v) < V\}$
    deliver t
end

Algorithm 1: Algorithm for simulating inhomogeneous Poisson point process