Tail behavior of shot noise processes

Gennady Samorodnitsky
Cornell University

Abstract

We discuss the different ways heavy tails can arise in shot noise models and possible applications of the latter to financial modeling.

1 Introduction

A shot noise process is a stochastic process of the form

\[ X(t) = \sum_{T_i \leq t} R_i(t - T_i), \quad t \in R, \]

where \( T_i \)'s are the points of a point process, \( \{N(A), \ A \text{ a Borel set}\} \), and \( \{R_i(t), \ t \geq 0\}, i = 1, 2, \ldots \) is a sequence of i.i.d. measurable stochastic processes, independent of the point process. Regarding \( X(t) \) as describing the state of a certain system at time \( t \), there is a very natural interpretation of a model given by (1.1). Think of \( T_i \)'s being the instances of time a “shock” arrives at the system. The effect of the shock at the system is a dynamic and random process. That is, \( h \) units of time later, at time \( T_i + h \) this effect is equal to \( R_i(h) \). The total state of the system at time \( t \) is then the sum of the effects of all shocks occurring up to that time. The stochastic processes \( \{R_i(t), \ t \geq 0\}, i = 1, 2, \ldots \) are called the response processes.

Shot noise stochastic processes are, in certain cases, attractive as stochastic models because many phenomena seem to actually develop as the result of discrete “shocks”. This natural structure of shot noise processes is missing for certain “black box” type of stochastic models.

It is an important feature of shot noise processes that that the process (1.1) is stationary if the point process \( \{N(A), \ A \text{ a Borel set}\} \) is stationary, quite irrespectively of the structure of the response processes \( \{R_i(t), \ t \geq 0\}, i = 1, 2, \ldots \). Recall that a point process \( \{N(A), \ A \text{ a Borel set}\} \), which we always assume to be Radon (i.e. \( N(A) < \infty \) with probability 1 for every compact \( A \)) is stationary if \( (N(A_1 + s), \ldots, N(A_k + s)) \overset{d}{=} (N(A_1), \ldots, N(A_k)) \) for all \( k \geq 1 \), all compact sets \( A_1, \ldots, A_k \) and \( s \in R \). See Resnick [Res87] and Karr [Kar86] for more details on point processes.

---

*AMS 1991 subject classification. Primary 60E07, 60G55, 62P05
1Keywords and phrases: Shot noise processes, heavy tails, financial models.
2Research supported by the NSF grant DMS-94-00535 and the NSA grant MDA904-95-H-1036.
The rich literature on shot noise processes typically discusses various particular cases of the process (1.1). Those are obtained by specifying a particular class of response processes and/or a particular class of point processes \( \{T_i, i \geq 1\} \). A very common class of response processes is that of the kind

\[
R_i(t) = Z_i f(t), \quad t \geq 0,
\]

where \( \{Z_i\} \) is a sequence of i.i.d. random variables, and \( \{f(t), t \geq 0\} \) is a non-random measurable function. With this choice of response processes one explicitly separates the response amplitude \( Z_i \) from the dynamics of the response given by the function \( f \). The resulting shot noise process has then the form

\[
X(t) = \sum_{T_i \leq t} Z_i f(t - T_i), \quad t \in R.
\] (1.2)

The model (1.2) is simpler than the general shot noise model (1.1), but it still forms a class of models that is often adequate. A further simplification of the model is achieved by taking \( Z_i \equiv 1, \ i \geq 1 \), thus leaving only the shock arrival times \( T_i \)'s as sources of the randomness in the state of the system.

The simplest, and a very common, choice of the point process \( \{N(A), A \ a \ Borel \ set\} \) is that of a time homogeneous Poisson process. A more general setting is that of a renewal process, Here \( T_i = Y_1 + \ldots + Y_i \) and \( T_{-i} = -(Y_{i-1} + \ldots + Y_0) \) for \( i \geq 1 \). Here \( Y_1, Y_2, Y_1, Y_0, Y_1, \ldots \) is a positive i.i.d. sequence. Of course, unless the renewal process is, actually, Poisson, the resulting shot noise process will not be stationary. If the step distribution \( F_Y \) of of \( Y_i \)'s has a finite mean \( \mu_Y \) one can construct a stationary shot noise process by using a stationary renewal process. That is, we define \( (T_{-1}, T_1) = (US, (1 - U)S) \), with \( S \) and \( U \) being independent random variables, such that

\[
P(S \leq x) = (1/\mu_Y) \int_0^x t F_Y(dt), \quad x > 0, \quad \text{and} \quad U \text{ uniform in } (0,1), \text{ and then } T_i = T_1 + Y_2 + \ldots + Y_i \text{ and } T_{-i} = T_{-1} - (Y_{-1} + \ldots + Y_{-(i-1)}) \text{ for } i \geq 2.\]

Here \( Y_i \)'s are independent of \( (T_{-1}, T_1) \).

Taking Poisson or renewal processes as the basis, one often defines a shot noise process with \( \{N(A), A \ a \ Borel \ set\} \) being a \textit{cluster process}. A cluster point processes consists of “secondary” points in clusters centered around “primary” points of the Poisson or renewal process. Intuitively, the clusters of secondary points are thought of as bursts of activity resulting from the event corresponding to a primary point. In this paper we will look only at a cluster Poisson process, and it is defined as follows. Let \( \{p_k, k \geq 1\} \) be a probability law on positive integers, and \( F_D \) be a probability law on \( R \). Let \( (K_j, -\infty < j < \infty) \) be i.i.d. random variables with common law \( \{p_k, k \geq 1\} \), independent of the array \( (D_{ij}, i \geq 1, -\infty < j < \infty) \) of i.i.d. random variables with common law \( F_D \). Finally, let \( (C_j, -\infty < j < \infty) \) be the points of a time homogeneous Poisson process with intensity \( \lambda \) independent of \( (K_j, -\infty < j < \infty) \) and of \((D_{ij}, i \geq 1, -\infty < j < \infty)\). Then for every Borel set \( A \),

\[
N(A) = \sum_{j=-\infty}^{\infty} \sum_{i=1}^{K_j} 1(C_j + D_{ij} \in A).
\] (1.3)

That is, \( C_j \)'s are the cluster centers (“primary points”), \( K_j \)'s are the cluster sizes, \( C_j + D_{ij}, i = 1, \ldots, K_j \) are the points of the \( j \)th cluster, \( \{p_k, k \geq 1\} \) is the cluster size distribution and \( F_D \) is the displacement distribution within a cluster.

The intuitive appeal and relative simplicity of shot noise processes have attracted attention of researchers both with a view of further developing the theory of these processes and to discuss their
applications in particular situations. General expositions of shot noise processes trace back to Rice [Ric44]. See also Parzen [Par62], Daley [Dal71] and Vervaat [Ver79]. More references are listed in Hsing and Teugels [HT89]. Shot noise processes have been used as models for computer failure times (Lewis [Lew64]) and earthquake aftershocks (Vere-Jones [VJ70]), and applied in such diverse fields as acoustics (Kuno and Ikegaya [KI73]), risk theory (Klüppelberg and Mikosch [KM93a] and [KM93b]) and financial processes (Samorodnitsky [Sam95]). The latter work is quoted extensively in the present paper, for it is our purpose presently to discuss in what ways heavy tails can and cannot arise in shot noise models and the implications of this to financial modeling.

2 Heavy tails of shot noise processes

Let \( \{X(t), t \in R\} \) be a well defined shot noise processes given by (1.2). Throughout this paper such a process is understood to be defined as

\[
X(t) = \lim_{H \to \infty} \sum_{-H \leq T_i \leq t} Z_i f(t - T_i), \ t \in R,
\]

and the process is well defined if the limit exists in probability. Let us assume, for a time being, that \( T_i \)'s are the points of a time homogeneous Poisson process with intensity \( \lambda > 0 \). The shot noise processes is then stationary. If its marginal distribution is heavy tailed, where do the heavy tails come from?

It is obvious that the tails of \( X(0) \) are at least as heavy as the tails of the amplitude size \( Z_1 \). That is, if

\[
P(Z_1 > x) = x^{-p}L(x), \ x \to \infty
\]

where \( p > 0 \) and \( L \) is a slowly varying function, and if \( \text{Leb}\{t > 0 : f(t) > 0\} > 0 \), then

\[
\lim_{x \to \infty} \frac{P(X(0) > x)}{x^{-p}L(x)} > 0.
\]

Can the tails of \( X(0) \) be heavy if the tails of \( Z_1 \) are not?

Suppose that the function \( f \) is bounded. The following simple argument shows that, in this case, the tails of \( X(0) \) cannot be heavier than those of \( Z_1 \). Specifically, we claim that if for some \( p \geq 1 \)

\[
E|Z_1|^p < \infty,
\]

then \( E|X(0)|^p < \infty \) as well.

Let \( M \) be such that

\[
\text{Leb}\{x > 0 : |f(x)| > M\} = 0.
\]

The random variable \( X(0) \) is infinitely divisible , and its Lévy measure is given by

\[
\mu = (P \times \text{Leb}) \circ H^{-1},
\]

3
where \( H : \Omega \times (0, \infty) \to \mathbb{R} \) is given by

\[
H((\omega, x)) = Z_1(\omega)f(x).
\]

See also (3.2) below. Moreover,

\[
\int_R (|z| \wedge 1) \mu(dz) < \infty. \tag{2.4}
\]

Now, \( E|X(0)|^p < \infty \) will follow once we show that

\[
\int_R |z|^p \mu(dz) < \infty. \tag{2.5}
\]

If \( Z_1 \) is essentially bounded, then there is an \( 0 < a < \infty \) such that \( P(|Z_1| \leq a) = 1 \). Then

\[
\int_R |z|^p \mu(dz) = \int_R |z|^p 1(|z| \leq aM) \mu(dz) \leq (aM)^{p-1} \int_R |z| 1(|z| \leq aM) \mu(dz) < \infty
\]

by (2.4). If \( Z_1 \) is not essentially bounded, then there are \( 0 < a < b < \infty \) such that both \( P(a < |Z_1| < b) > 0 \) and \( P(|Z_1| > b) > 0 \). Then by (2.4)

\[
\infty > E \int_0^\infty (|Z_1 f(x)| \wedge 1) \, dx
\]

\[
\geq E\left(1(a < |Z_1| < b) \int_0^\infty (|Z_1 f(x)| \wedge 1) \, dx\right) + E\left(1(|Z_1| > b) \int_0^\infty (|Z_1 f(x)| \wedge 1) \, dx\right)
\]

\[
\geq aP(a < |Z_1| < b) \int_0^\infty |f(x)| 1(|f(x)| \leq \frac{1}{b}) \, dx + P(|Z_1| > b) \int_0^\infty 1(|f(x)| > \frac{1}{b}) \, dx.
\]

Therefore,

\[
\int_0^\infty |f(x)| 1(|f(x)| \leq \frac{1}{b}) \, dx < \infty \tag{2.6}
\]

and

\[
\int_0^\infty 1(|f(x)| > \frac{1}{b}) \, dx < \infty. \tag{2.7}
\]

We therefore have

\[
\int_R |z|^p \mu(dz) = E \int_0^\infty |Z|^p |f(x)|^p \, dx = E|Z|^p \int_0^\infty |f(x)|^p \, dx
\]

\[
\leq E|Z|^p \left(\frac{1}{b} \right)^{p-1} \int_0^\infty |f(x)| 1(|f(x)| \leq \frac{1}{b}) \, dx + M^p \int_0^\infty 1(|f(x)| > \frac{1}{b}) \, dx \right) < \infty,
\]

and so \( E|Z_1|^p < \infty \) implies \( E|X(0)|^p < \infty \) when \( f \) is a bounded function. See Doney and O’Brien [DO91] for a delicate related result in the absence of infinite divisibility.

Alternatively, heavy tails of \( X(0) \) can be caused by a special structure of the response function \( f \), even when \( Z_1 \) is light tailed. Indeed, the model \( X_{\gamma,d,m} \) below uses a (symmetrized) exponentially distributed \( Z_1 \). Models in which heavy tails are caused by a special structure of the response function \( f \) tend to be more flexible than those in which heavy tails are caused by a heavy tailed amplitude \( Z_1 \), in the following sense. Suppose, we would like to have a stationary model \( \{X(t), t \in \mathbb{R}\} \) in which, for some \( n_0 > 1 \) the probability tail of \( X(0) + X(1) + \ldots + X(n_0 - 1) \) is strictly lighter than that of
$X(0)$. This is desirable, for instance, in the case of financial modeling, where daily returns appear to have heavier tails than, say, weakly returns. This “tail cancelation” property is impossible to model by a shot noise process when the heavy tails are caused by a correspondingly heavy tailed amplitude $Z_1$, for it remains unaffected by taking linear combinations of the values of the process. On the other hand, forming such linear combinations can affect significantly the structure of the response function $f$ and, through it, the probability tails of the shot noise, if it is the response function that causes heavy tails. We exploit this idea in the sequel.

3 The structure of a cluster Poisson shot noise process

Let $\{X(t), t \in R\}$ be a shot noise process (1.2) with the point process $\{N(A), A$ a Borel set} being a cluster Poisson process (1.3). We refer to this process as a cluster Poisson shot noise process. The following proposition describes some of the basic properties of this process.

**Proposition 3.1** The cluster Poisson shot noise process is well defined by (2.1) if and only if

$$
\int_0^\infty \sum_{k=1}^\infty p_k E \left( \left( \sum_{i=1}^k Z_i f(x + D_{i1}) \right)^v \right) dx < \infty.
$$

(3.1)

In this case, it is a stationary infinitely divisible process. The marginal Lévy measure $\mu$ of $X(0)$ is given by

$$
\mu = \lambda \sum_{k=1}^\infty p_k \mu_k,
$$

(3.2)

where each $\mu_k$ is a Lévy measure defined by

$$
\mu_k = \left( P \times \text{Leb} \right) \circ H_k^{-1},
$$

with $H_k : \Omega \times (0, \infty) \to R$ given by

$$
H_k(\omega, x) = \sum_{i=1}^k Z_i(\omega) f(x + D_{i1}(\omega)).
$$

Moreover, if $EK_1^2 < \infty$, $EZ_1^2 < \infty$ and

$$
\int_0^\infty f(y)^2 dy < \infty,
$$

then the process $\{X(t), t \in R\}$ has a finite second moment, and its covariance function $(R(t), t \geq 0)$ is given by

$$
R(t) = \lambda E K_1 EZ_1^2 \int_0^\infty f(y) f(y + t) dy + \lambda(EK_1^2 - EK_1)(EZ_1)^2 \delta(t),
$$

(3.3)

with

$$
\delta(t) = \int_0^\infty f(y + t) \left( \int_R \int_R f(y + x_1 - x_2) F_D(dx_1) F_D(dx_2) \right) dy
$$

(3.4)

(\text{where} $f(y) = 0$ for $y < 0$).
See Section 2 of Samorodnitsky [Sam95] for the argument. Observe that the effect of the displacement distributions $F_D$ on the covariance function disappears if $EZ_1 = 0$. In the latter case

$$R(t) = \lambda E K_1 EZ_1^2 \int_0^\infty f(y)f(y + t)dy,$$

which makes it easy to see how long memory, another often observed property of financial processes, can be modeled with shot noise processes.

4 A shot noise model for financial processes

Before describing the particulars of the model, we list some of the more important properties of financial processes, like stock or exchange rate returns.

1. The data has observable periods of high activity and low activity, often measured through volatility.

2. The observations are heavy tailed.

3. As sampling interval increases, the tails tend to become less heavy. That is, adding up a certain number of consecutive observations may lead to lighter tails.

4. The data has long memory property (long range dependence).

See e.g. Akgiray and Booth [AB88] and Guillaume et al. [GDD+94]. We remark that the latter property means that observations far apart in time are highly dependent. The empirical autocorrelation function of the raw returns is often quite small, but the long memory is typically captured by the empirical autocorrelation function of the absolute returns.

Our purpose in this section is to show how we can account for such properties with a shot noise model. We will use a cluster Poisson shot noise process described in the previous section. An extremely important property of shot noise models is its intuitive structure. See Mandelbrot [Man82] for an emphatic argument for intuitive models. For simplicity we construct a symmetric model (even though it is well known that positive excursions of financial processes often look differently from their negative excursions). Nonsymmetric models can be constructed in a similar way.

Fix a $\gamma > 0$, $d > 0$ and $m \geq 1$, and define

$$f_{\gamma,d,m}(s) = \begin{cases} j^\gamma \cos \frac{j \pi}{m} & \text{if } s \in \left( j - (j + 1)^{-d+1}, j + (j + 1)^{-d+1} \right), j \geq 1. \\ 0 & \text{otherwise} \end{cases}$$

(4.1)

The function $f_{\gamma,d,m}$ is our response function. It is plotted on Fig. 4.1 for $\gamma = .4$, $d = .85$ and $m = 3$. Observe that the “spikes” of the response function tend to grow with time, but they occur on more and more narrow intervals. The spikes grow faster when the parameter $\gamma$ is big, and they become more narrow when the parameter $d$ is big. We expect, therefore, to see both heavier tails and longer memory when $\gamma$ is big, and $d$ is small. This becomes transparent from Theorem 4.1 below.
Figure 4.1

The presence of a trigonometric factor in (4.1) will be useful for the purpose of “tail cancelation”. Let \( Z_1 \overset{d}{=} eY \), where \( e \) is a Rademacher random variable (\( P(\epsilon = 1) = P(\epsilon = -1) = 1/2 \)), independent of the standard exponential random variable \( Y \), and let \( \{X_{\gamma,d,m}(t), t \in R\} \) be defined by (1.2) with the response function given by (4.1). The choice of a Laplace distribution for \( Z_1 \) is designed both to keep the latter light tailed, and to make tail computations more transparent via Tauberian-type arguments. Observe that the “cluster” structure of the shot noise process has built-in periods of high and low activity. The degree of distinction between them can be articulated in the model by varying the cluster size distribution and the displacement within cluster distribution. The following theorem shows that our model has other features listed above as well.

Note that we are discussing here only the equally spaced observations \( X_{\gamma,d,m}(i), i = 0, 1, 2, \ldots \) of \( \{X_{\gamma,d,m}(t), t \in R\} \). However, the process \( X_{\gamma,d,m} \) is, by its nature, a continuous time process, and, as such, can be naturally used as a continuous-time model.

**Theorem 4.1** (i) Suppose that \( E K_1 < \infty \). Then \( X_{\gamma,d,m} \) is a well defined stationary infinitely divisible stochastic process.

(ii) Assume that

\[
E K_1^2 < \infty \tag{4.2}
\]

and that the displacement distribution \( F_D \) is absolutely continuous with respect to the Lebesgue measure. Then

\[
P(X_{\gamma,d,m}(t) > x) \sim c_{d,\gamma,m}^{(1)} \lambda E K_1 x^{-d/\gamma}, \ x \to \infty, \tag{4.3}
\]

where

\[
c_{d,\gamma,m}^{(1)} = \frac{1}{2\gamma m} \Gamma(d/\gamma) \sum_{i=0}^{2m-1} |\cos \frac{i}{m}|^{d/\gamma}.
\]
(iii) Under the assumptions of (ii), if $\gamma > 1$, then

$$P\left( \sum_{k=0}^{2m-1} X_{\gamma,d,m}(t + k) > x \right) \sim \begin{cases} c^{(2)}_{d,\gamma,m} \lambda E K_1 x^{-(d+1)/\gamma} & \text{if } \frac{d+1}{\gamma} < \frac{d}{\gamma-1}, \\ c^{(3)}_{d,\gamma,m} \lambda E K_1 x^{-d/(\gamma-1)} & \text{if } \frac{d+1}{\gamma} > \frac{d}{\gamma-1}, \\ (c^{(2)}_{d,\gamma,m} + c^{(3)}_{d,\gamma,m}) \lambda E K_1 x^{-(d+1)/\gamma} & \text{if } \frac{d+1}{\gamma} = \frac{d}{\gamma-1}. \end{cases}$$

(4.4)

and if $\gamma \leq 1$

$$P\left( \sum_{k=0}^{2m-1} X_{\gamma,d,m}(t + k) > x \right) \sim c^{(2)}_{d,\gamma,m} \lambda E K_1 x^{-(d+1)/\gamma}$$

(4.5)

as $x \to \infty$. Here

$$c^{(2)}_{d,\gamma,m} = \frac{1}{m} \Gamma\left( \frac{(d + \gamma + 1)}{\gamma} \right) \sum_{i=1}^{2m-1} \sum_{j=1}^{2m} \left| \sum_{k=0}^{i-1} \cos \pi (j + k)/2m \right|^{(d+1)/\gamma} > 0$$

and

$$c^{(3)}_{d,\gamma,m} = \frac{1}{\gamma - 1} \left( \gamma m \right)^{d/(\gamma-1)} \Gamma\left( \frac{d}{\gamma - 1} \right).$$

(iv) Assume that (4.2) holds, and that

$$d > 2\gamma.$$  

(4.6)

Then the process $X_{\gamma,d,m}$ has a finite variance, and its covariance function $R(n)$ satisfies, as $n \to \infty$,

$$R(n) \sim \kappa_{d,\gamma,m} \lambda E K_1 n^{2\gamma - d} \sum_{i=1}^{2m} \cos \pi i/m \cos \pi (i + n)/m$$

(4.7)

in the sense that the product $n^{d-2\gamma} R(n)$ has at most $2m$ subsequential limits given by (4.7), all of which are finite, and some of which are positive. Here

$$\kappa_{d,\gamma,m} = 2m^{-1} \int_0^\infty x^{\gamma} (1 + x)^{\gamma - d - 1} \, ds.$$  

For the proof see Samorodnitsky [Sam95].

Observe that parts (ii) and (iii) of Theorem 4.1 show that our shot noise process has heavy tails in spite of light tailed exponential $Z$'s, and that the tail of the sum of $2m$ consecutive observations is lighter than that of a single observation. If one had a reliable estimator of the tail index, it would have been possible to estimate $d$ and $\gamma$ from (4.3) and (4.4). The trigonometric behavior of the correlation function is due to the choice of the response function $f_{\gamma,d,m}$. However, long memory is obvious. In particular, $\sum_{n=1}^{\infty} |R(n)| = \infty$ if $2\gamma < d \leq 2\gamma + 1$.

The plot on Figure 4.2 shows the result of a simulation of the cluster Poisson shot noise model, using a finite number of clusters. We have taken 600 clusters, with cluster distribution being geometric with mean 10.
The Poisson arrival rate has been taken to be equal to 1, and the displacement distribution $F_D$ to be uniform in $(0, 3)$. For the purpose of the above simulation we have chosen $m = 2$, $d = 1.2$ and $\gamma = .5$. Note that even though our process is symmetric (and ergodic), the displayed simulated path is mostly negative, which demonstrates the length of dependence in our model. This is not very characteristic of financial returns processes. However, our goal now is only to demonstrate certain features of cluster Poisson shot noise processes (one gets a very different type of behavior with a different choice of parameters.) One can easily see on this plot the periods of high and low volatility, and presence of heavy tails is also clear.
It follows from Proposition 3.1 that $X_{\gamma,d,m}(0)$ has a finite Lévy measure. Unavoidably, therefore, the process will have “flat” spots, observed on the plot. One can regard such a model as accounting only for the “most significant” events on the market. If we model the rest of the activity on the market by an additive Brownian motion, the result is the path shown on Figure 4.3.

![Figure 4.4](image)

Finally, to demonstrate the effect on the length of the memory of the process and the heaviness of the tails, we have reduced the parameter $d$ from 1.2 to 0.8.

![Figure 4.5](image)
The discussion above shows that we should expect to see longer memory and heavier tails. This is, of course, seen from the plot on Figure 4.4. We have used the same random numbers for the two simulations, hence both are skewed towards negative values. Finally, Figure 4.5 presents the latter path with an additive Brownian motion.

References


Genady Samorodnitsky
School of Operations Research and Industrial Engineering
Cornell University
Ithaca, NY 14853