

# Integer-valued Lévy processes and low latency financial econometrics

OLE E. BARNDORFF-NIELSEN

*The T.N. Thiele Centre for Mathematics in Natural Science,  
Department of Mathematical Sciences,  
University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark  
& CREATES, University of Aarhus  
oebn@imf.au.dk*

DAVID G. POLLARD

*Pollards Et Filles Ltd,  
Tanglewood, Emmer Green, Reading RG4 8UE, UK  
dpollard@dircon.co.uk*

NEIL SHEPHARD

*Nuffield College, New Road, Oxford OX1 1NF, UK,  
Department of Economics, University of Oxford  
& Oxford-Man Institute, University of Oxford,  
neil.shephard@economics.ox.ac.uk*

January 11, 2012

## Abstract

Motivated by features of low latency data in financial econometrics we study in detail integer-valued Lévy processes as the basis of price processes for high frequency econometrics. We propose using models built out of the difference of two subordinators. We apply these models in practice to low latency data for a variety of different types of futures contracts.

Keywords: futures markets; high frequency econometrics; low latency data; negative binomial; Skellam; tempered stable.

JEL Classifications: C01, C14, C32

## 1 Introduction

In this paper we provide an exploratory analysis of low latency financial data. Our focus is on the unconditional distributional features of returns at times of trades only, establishing the framework of integer-valued Lévy processes as a fundamental starting point for models of low latency data. This can be thought of as a first step towards more realistic stochastic process modelling, which in particular would involve time-change to allow for serial dependence in returns and diurnal features.

Recently low latency data have become available for research. These data from specialist data providers are recorded very close to the data exchange itself and are therefore of the highest available

quality. Typically low latency data are added to the data providers database less than 1 millisecond after they leave the exchange.

There has been considerable interest in using high frequency financial data to aid decision making. Recent reviews are given by Russell and Engle (2010) and Bauwens and Hautsch (2009). Leading applied reasons include:

(i) Building models to design efficient trading methods with low transaction costs. These methods are typically implemented electronically and are called “automated trading”. An interesting recent example being Avellaneda and Stoikov (2008). Such methods often study the relative utility of market and limit orders, see for example Lehmann (2008) for a theoretical discussion. An aspect of this is to provide better short term forecasting of market levels.

(ii) Harnessing the data to better estimate medium term financial volatility or dependence e.g. by Andersen et al. (2001), Barndorff-Nielsen and Shephard (2002), Barndorff-Nielsen et al. (2008) and Mykland and Zhang (2011). There is a small literature on realised volatility type statistics applied to Lévy related processes. References include Barndorff-Nielsen and Shephard (2006) and Bacry et al. (2011). The latter, which was carried out independently and concurrently with this paper, shares our interest in the discreteness of prices.

(iii) Studies of the relationships between the many quantities of economic interest. For example relationships between trade volumes and price changes have been studied by Potters and Bouchaud (2003) and Lo and Wang (2010) amongst many and between order flow and tick price changes by Weber and Rosenow (2005) and others.

In this paper we systematically develop a continuous time integer-valued Lévy process which has features which are attractive for low latency data. In particular this process delivers prices which obey the tick structure we observe empirically in low latency data. Its most basic form is based on the Skellam distribution and can be thought of as modeling price changes as the difference between two Poisson processes, but we generalise this to processes based on the difference of two integer-valued subordinators. We also show that any integer-valued Lévy process is the sum of a compound Poisson process with only zero innovations and the difference of two independent subordinators. We also generalise this decomposition to the multivariate case.

The structure of integer-valued Lévy processes means our models will evolve over the tick structure of high frequency data. Related integer-valued econometric models include those discussed by, for example, Hausman et al. (1992), Rydberg and Shephard (2003), Zeng (2003), Russell and Engle (2006), Hasbrouck (1999), Phillips and Yu (2008), Rosenblau (2009) and Hansen and Horel (2009).

The model we discuss in this paper is not fully fledged. However, it can be extended us-

ing time-changes to yield volatility clustering as well as allowing serial correlation due to market microstructure effects.

The structure of this paper is as follows. In Section 2 we first discuss some features of low latency data and their empirical statistics. We detail what we call pure mid-prices, which are a variant on the usual mid-prices that preserve the tick structure of the bid and ask prices recorded in low latency data. Section 3 looks at the mathematics of integer-valued continuous time processes. We introduce the basic continuous time Skellam process and discuss its properties. We generalise it in a number of ways, to allow for heavier tails. We apply compound Poisson processes to the financial data in Section 4 and fit the so-called  $\Delta$ NB Lévy and  $\Delta$ PTS processes, classes of processes we introduce here, in Section 5. We then draw some conclusions in Section 6. Derivations of important properties of integer-valued Lévy processes are in the Appendices.

## 2 Low latency data and integer-valued distributions

### 2.1 Low latency futures data

We will study tick price processes in low latency data from futures exchanges<sup>1</sup>. Futures exchanges trade many assets ranging from equity indices to interest rate products and commodities. Liquidity on the electronic marketplace in many of these futures contracts is good and the exchanges well established. They are able to provide low latency data feeds recording every price and new order update seen on the matching engine’s order book.

We study, in particular, futures data for the S&P500 (mini) contract, the US Treasury 10 year note, the NYMEX benchmark Crude Oil contract and the IMM Eurodollar futures contract. These markets are sufficiently different to demonstrate a range of tick price behaviour. These data were provided to us by QuantHouse ([www.quanthouse.com](http://www.quanthouse.com)) from data feeds at the Chicago Mercantile Exchange (CME) which is one of the largest Futures exchanges.

### 2.2 Mid-price changes

By ‘tick price process’ we refer to the continuous time evolution of the ‘best’ price in the market as it changes over time from data update to update. The mid-price is the arithmetic average of the best ask  $P_{ask,t}$  and best bid price  $P_{bid,t}$

$$P_{mid,t} = 0.5(P_{ask,t} + P_{bid,t}), \quad t \in \mathbb{R}_{\geq 0}.$$

---

<sup>1</sup>By low latency data we mean high frequency data that are not conflated and that contain very few data errors such as zero or negative bid-ask spreads. “Conflation” is a process by which a number of consecutive data updates are aggregated with only the aggregate result being reported by the data provider. This process produces an incorrect record of market activity. Apart from conflation some data feeds that are not fast enough may also simply drop or miss data updates. Such data losses can result in inconsistent bid and ask prices that may present as zero or negative bid-ask spreads. With good low latency data we go as far as current technology allows to getting an accurate record of market activity.

This price is computed in continuous time and its value changes when either the ask or bid is altered.

The minimum price change allowed by the exchange on any market, the ‘tick size’, means that exchange prices map to the positive integers and mid-prices to the positive integers and half-integers.

We will do some rudimentary filtering by restricting attention to the period of the day when the market trades actively<sup>2</sup> and then selecting only those data updates at times when trades occur. The times at which there are trades will be written as

$$\tau_i, \quad i = 1, 2, \dots, N.$$

The justification being that when trades occur there is agreement by at least two market participants about the market price and so we have more confidence in its accuracy. Figure 1 shows this for the Euro-Dollar IMM FX futures contract during 10th November 2008, which had 33,074 trades on that day. The upper left hand graph plots the bid and ask at the times of the first 80 trades. It also shows the pure mid-prices, which we will define in a moment. The corresponding pure mid-price returns<sup>3</sup> are given in the upper right hand graph in the figure. It shows integer returns, with most being  $-1$ ,  $0$  and  $1$ . However, there is also a return of 2 ticks.

Consider changes in  $p_{mid,\tau_i}$ ,

$$c_i^* = \frac{P_{mid,\tau_{i+1}} - P_{mid,\tau_i}}{tickSize}, \quad c_i^* \in \left\{ \dots, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \dots \right\}, \quad (1)$$

between consecutive trades at times  $\tau_{i+1}, \tau_i$ . For the above contract the tick size is 0.0001 of a unit, i.e. prices move from, for example, 1.2768 to 1.2767 U.S. Dollar to the Euro. Then these changes  $c_i^*$  mostly live on the integers but have some mass on the half integers mostly caused by one sided moves in the spread, i.e. the ask moving up one tick but no move in the bid, causing the spread to widen and the mid-price to move up by a half a tick. It turns out these spread induced half tick changes are difficult to model for various reasons (including they make the distribution of price changes non-monotonic as we go away from zero).

### 2.3 Pure mid-prices

We can reduce these spread induced changes by using what we call “pure mid-prices”. Pure mid-prices move the price as little as possible subject to keeping the pure mid-price between the bid and the ask at times of trade. This can be formalised in the following way.

Pure mid-prices are defined recursively by the following criteria

$$P_{pure,t} = \arg \min_x |x - P_{pure,\tau_i}|, \quad t \in (\tau_i, \tau_{i+1}]$$

<sup>2</sup>The ‘active day’ is restricted to the period 12.30-18.00 GMT, for each of the datasets we study.

<sup>3</sup>By ‘returns’ here and elsewhere we mean the tick-size normalised raw price changes as opposed to the log-returns or arithmetic-returns more conventionally used in lower frequency financial analysis.

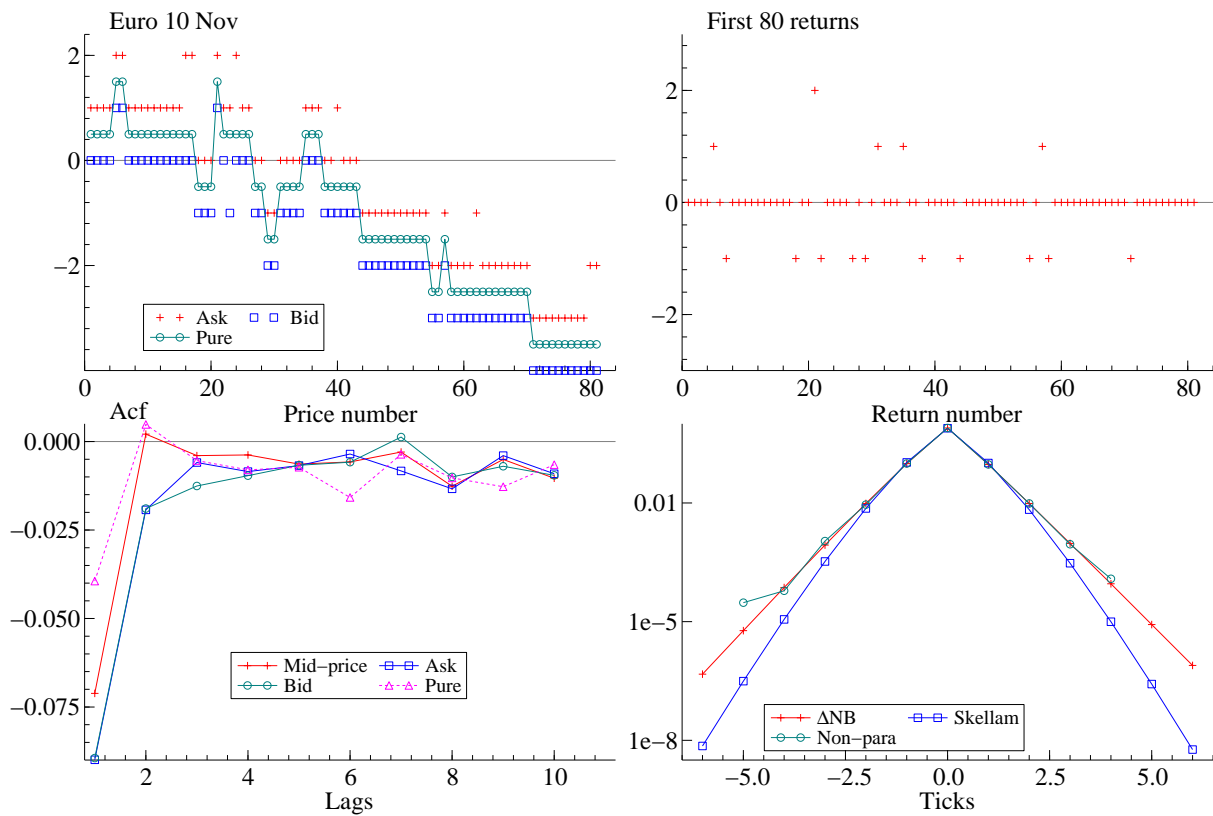


Figure 1: *Euro-Dollar IMM FX futures contract on 10th November 2008. Top left: ask, bid and pure mid-price for the first 80 trades of the day minus the price at the start of the day. Top right: returns from pure mid-price. They are all integers. Bottom left: correlogram for mid-price, ask, bid and pure mid-price for entire day. Bottom right: log-histogram of pure mid-price returns: non-parametric, Skellam and  $\Delta NB$  distributions.*

subject to the discrete time constraining knots

$$P_{bid,\tau_i} < P_{pure,\tau_i} < P_{ask,\tau_i}, \quad i = 1, 2, \dots, N.$$

This means that pure mid-prices are not effected by a widening of the spread. This definition requires an initial value to be given to the recursion, which we take as  $P_{pure,\tau_1} = P_{mid,\tau_1}$ .

In tick space the assets we discuss in this paper will have a spread which will be one or more ticks. As a result it makes sense from an econometric modelling viewpoint to add a second criteria to scaled pure mid-prices — that they are half-integers. That is we only allow

$$\frac{P_{pure,t}}{tickSize} \in \left\{ \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \dots \right\} = \mathbb{Z}_{\geq 0} + \frac{1}{2}.$$

Hence if, for example the tick size was one,  $P_{ask,\tau_i} = 101$  and  $P_{bid,\tau_i} = 100$  then  $P_{pure,\tau_i} = 100.5$ , while if this is followed by  $P_{ask,\tau_{i+1}} = 103$  and  $P_{bid,\tau_{i+1}} = 101$  then  $P_{pure,t}$  keeps at 100.5 until time  $\tau_{i+1}$  when it instantly jumps up to  $P_{pure,\tau_{i+1}} = 101.5$ . Likewise if  $P_{ask,\tau_{i+1}} = 102$  and  $P_{bid,\tau_{i+1}} = 100$  then  $P_{pure,\tau_{i+1}} = 100.5$ . This then delivers an integer return sequence from the half-integer scaled pure mid-prices. This will turn out to be relatively easy to model<sup>4</sup>.

We should note that if the futures contract is traded on a so-called one-tick market (see, for example, Field and Large (2008)), where depths are so large that the spread is always one tick wide, then the pure mid-price and the usual mid-price will always be identical.

These remarks are illustrated in Figure 1 which plots (in the upper panel) returns on pure mid-prices at the times of the trades that occurred on the Euro-Dollar FX contract during 10th November 2008. Pure mid-prices returns are integers.

The bottom graphs hold some summaries of returns. On the left are the correlograms and they show the usual small amount of negative autocorrelation due to market microstructure effects (e.g. Hansen and Lunde (2006) and the references contained within it). Strong autocorrelation basically lasts one lag and is more modest for the pure mid-price return series (it is well known the trades themselves will live on the lattice structure but will have a great deal of autocorrelation). The latter point seems a robust feature across a lot of assets we have studied. Interestingly the pure mid-price returns have less autocorrelation than the returns from mid-prices, asks or bids.

The bottom right hand side plot shows the unconditional histogram for the pure mid-price returns for the whole day of data. The non-parametrically estimated log-probabilities seem to be declining roughly linearly in the tails for this dataset.

The pictures change over time, but many features are constant. Figure 2 shows the analysis on 7th November 2008, a US Non-farm payroll day. Now tick changes of order  $\pm 40$  occur during

---

<sup>4</sup>Note both  $\frac{P_{ask,t} - P_{pure,t}}{tickSize} \in \mathbb{Z}_{\geq 0} + \frac{1}{2}$  and  $\frac{P_{pure,t} - P_{bid,t}}{tickSize} \in \mathbb{Z}_{\geq 0} + \frac{1}{2}$ . Hence a very basic factor model for the bid and ask in continuous time is to model a discrete-valued martingale  $\frac{P_{pure,t}}{tickSize} - \frac{1}{2}$  and two stationary non-negative discrete-valued processes  $\frac{P_{ask,t} - P_{pure,t}}{tickSize} - \frac{1}{2}$  and  $\frac{P_{pure,t} - P_{bid,t}}{tickSize} - \frac{1}{2}$ .

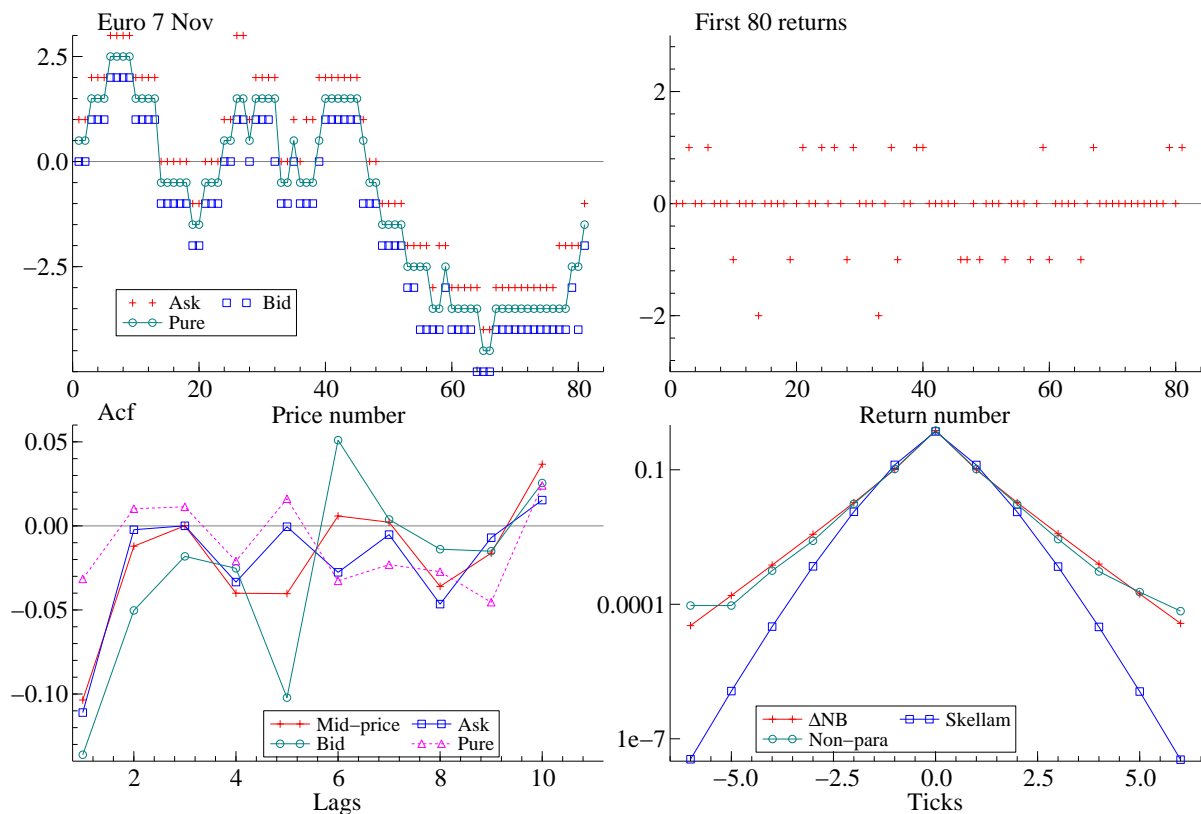


Figure 2: *Euro-Dollar IMM FX futures contract on 7th November 2008. Top left: ask, bid and pure mid-price for the first 80 trades of the day minus the price at the start of the day. Top right: returns from pure mid-price. They are all integers. Bottom left: correlogram for mid-price, ask, bid and pure mid-price for entire day. Bottom right: log-histogram of pure mid-price returns: non-parametric, Skellam and  $\Delta NB$  distributions.*

the day, just after the announcement, and the log probability plot shows more extended tails as a consequence. Again the correlogram is closer to being white noise for the pure mid-price changes than for the alternatives we considered.

For many other markets a similar picture holds for pure mid-price changes. Later we will illustrate this using data from the Ten Year US treasury note (TNC), Nymex/CME benchmark crude oil contract (CLN) and the mini S&P500 contract (ESPC).

For each of these series Table 1 provides summary statistics, indicating the number of low latency returns available. The Table also shows the standard deviations when scaled prices are computed using mid-prices, asks, bids and pure mid-prices. As expected the standard deviations are lower for mid-prices than for asks and bids, which reflects their smaller amount of autocorrelation. An interesting feature of the Table is that the standard deviation of the pure mid-price returns are typically smaller than that for the mid-price returns.

	# of trades	Standard deviation			
		Mid	Ask	Bid	Pure
Euro 07/11/08	42,592	0.834	0.849	0.999	0.723
Euro 10/11/08	33,074	0.545	0.584	0.584	0.538
ESPC 10/11/08	163,970	0.260	0.267	0.268	0.260
CLN 10/11/08	90,762	0.822	0.937	0.990	0.760
TNC 10/11/08	26,764	0.319	0.326	0.324	0.318

Table 1: *Summary statistics for the five low latency data sets used in this paper. Shows the sample size on each day and the standard deviations of the returns, having scaled the returns so they are in ticks. The returns are computed using mid-prices, asks, bids or using pure mid-prices.*

### 3 Mathematics of integer-valued Lévy processes

#### 3.1 Introduction

In order to build models of integer-valued price changes it is important to have an understanding of continuous time processes which deliver independent and stationary increments which are integer-valued. These processes can be time-changed to deliver empirically plausible models with both diurnal features and time-varying volatility, in the same way Brownian motion is often time-changed to deliver stochastic volatility.

The basis of our analysis will be integer-valued Lévy processes. Recall a càdlàg stochastic process  $L = \{L_t\}_{t \geq 0}$  with  $L_0 = 0$  is a Lévy process if and only if it has independent and (strictly) stationary increments. See the reviews of Lévy processes by, for example, Sato (1999) and Cont and Tankov (2004). An integer-valued Lévy process has its law concentrated on  $\mathbb{Z} = \{i : i = 0, \pm 1, \pm 2, \dots\}$ . The simplest example of this class is the Poisson process, but clearly this is not satisfactory for our tick process.

The following proposition indicates the way we can build these kinds of models.

**Proposition 1** *Suppose  $L$  is a integer-valued Lévy process. Then the Lévy measure  $\nu$  of  $L$  is concentrated on  $\mathbb{Z} \setminus \{0\}$  and has finite mass.*

**Proof.** Given in the Appendix.

The finiteness of  $\nu$  implies that  $L$  is of finite-jump activity, i.e. it has at most finitely many jumps in any finite time interval. Consequently, without loss of generality,  $L$  can be written in the form

$$L_t = L_t^+ - L_t^-,$$

where the paths of  $L_t^+$  and  $L_t^-$  can be deduced from the single path of  $L_t$  simply by summing the positive and negative jumps of  $L$  separately. Thus  $L^+$  and  $L^-$  are both discrete subordinators (Lévy



processes with non-negative increments<sup>5</sup>), whose Lévy measures  $\nu^+$  and  $\nu^-$  are the restrictions of  $\nu$  to the positive and negative half axes, respectively. Since  $\nu^+$  and  $\nu^-$  are both finite measures,  $L^+$  and  $L^-$  are compound Poisson (CP) processes, and, in obvious notation,  $L$  may be written as

$$L_t = \sum_{j=1}^{N_t^+} C_j^+ - \sum_{j=1}^{N_t^-} C_j^-.$$

where  $\{N_t^+, N_t^-\}$  are independent (homogeneous) Poisson processes with intensities  $\psi^+ = \nu((0, \infty))$  and  $\psi^- = \nu((-\infty, 0))$ , while the  $C_j^\pm$  are strictly positive integer innovations. The fact that they are greater than or equal to one is important. Notice that with probability one the paths of  $N_t^+$  and  $N_t^-$  jump at different times. The class of Lévy processes that are piecewise constant are known to coincide with the class of compound Poisson processes, see (Sato, 1999, Theorem 21.2).

**Remark 1** *The tick process itself can be written as a compound Poisson process*

$$L_t = \sum_{j=1}^{N_t} C_j,$$

where  $N_t$  is the number of trades up to time  $t$  and  $C_j$  are the potential moves when there is a trade. In this case  $C_j$  has an atom at 0 as many trades will not move the price. Without observing the counting process  $N_t$  the process would not be identified due to the  $C_j$  having an atom at zero. Compound Poisson models with general, not necessarily an integer, returns have a long history, examples include Press (1967) and Madan and Seneta (1984).

**Remark 2** *The proof of the validity of the splitting  $L = L^+ - L^-$ , discussed above, can be extended to multivariate integer-valued Lévy processes. We say that a  $d$ -dimensional Lévy process  $L$  is integer-valued if it takes values in  $\mathbb{Z}^d \setminus \{\mathbf{0}\}$  (where  $\mathbf{0}$  is the origin in  $\mathbb{Z}^d$ ). Thus if, for example,  $L$  is a bivariate integer-valued Lévy process then it can be represented as a sum of 8 independent compound Poisson processes, one for each of the four quadrants and one for each of the four half axes.*

### 3.2 Cumulants

A characterising feature of Lévy processes is that, so long as they exist,  $\kappa_{j,t} = t\kappa_j$ ,  $j = 1, 2, \dots$ , where  $\kappa_{j,t}$  and  $\kappa_j$  are the  $j$ -th cumulant of  $L_t$  and  $L_1$ , respectively.

The cumulant function of any Lévy process  $Y_t$  has the form

$$C\{\theta \ddagger Y_t\} = \log[\mathbb{E} \exp\{i\theta Y_t\}] = tC\{\theta \ddagger Y_1\}.$$

---

<sup>5</sup>Discrete infinite divisibility for distributions on  $\mathbb{N}_0 = \{i : i = 0, 1, 2, \dots\}$  is discussed briefly in Bondesson (1992) and more extensively in Steutel and Van Harn (2004).

This implies for the integer-valued process  $L$  that  $C\{\theta \dagger L_t\} = tC\{\theta \dagger L_1^+\} + tC\{-\theta \dagger L_1^-\}$  and consequently

$$\kappa_{j,t} = t\kappa_j^+ + t(-1)^j \kappa_j^-$$

where  $\kappa_j^+$  and  $\kappa_j^-$  denote the cumulants of  $L_1^+$  and  $L_1^-$ , respectively. Further, since  $L^\pm$  are compound Poisson with rates  $\psi^\pm$  we have  $C\{\theta \dagger L_1^\pm\} = -\psi^\pm [1 - \exp\{C(\theta \dagger C_1^\pm)\}]$ , where  $C(\theta \dagger C_1^\pm)$  is the cumulant function of  $C_1^\pm$ . Hence

$$\kappa_j^\pm = -\psi^\pm (1 - \mu_j'^\pm) = \psi^\pm (\mu_j'^\pm - 1),$$

where the uncentred moments  $\mu_j'^\pm = E\{(C_1^\pm)^j\}$ . Note that  $\mu_j'^\pm \geq 1$  by construction.

### 3.3 Skellam Lévy process

In the simplest case where all the jumps are unit, then with probability one,  $C_n^\pm = 1$  so prices move a single tick at a time. Then  $L_t^\pm = N_t^\pm$  so

$$L_t = N_t^+ - N_t^-.$$

We call this a Skellam Lévy process, for the process is the Lévy process generated from the Skellam distribution, introduced by Irwin (1937) and studied more recently by Abdulhamid and Omair (2010). That distribution is the law of the difference of two independent Poisson distributions, with parameters  $\psi^+$  and  $\psi^-$ , say, and we will denote it by  $Sk(\psi^+, \psi^-)$ . Then we have that

$$L_t - L_s \sim Sk((t-s)\psi^+, (t-s)\psi^-), \quad t > s. \quad (2)$$

For  $k \in \mathbb{N}_0$  the point probabilities of the Skellam distribution,  $Sk(\psi^+, \psi^-)$ , are

$$\begin{aligned} p_k &= \sum_{n=0}^{\infty} \Pr(N_t^+ = n+k) \Pr(N_t^- = n) = e^{-\psi^+ - \psi^-} (\psi^+)^k \sum_{n=0}^{\infty} \frac{(\psi^+ \psi^-)^n}{n!(k+n)!} \\ &= e^{-\psi^+ - \psi^-} (\psi^+)^k \left(\sqrt{\psi^+ \psi^-}\right)^{-k} I_{|k|}(2\sqrt{\psi^+ \psi^-}), \end{aligned}$$

where  $I_k(x)$  is a modified Bessel function of the first kind (Abramowitz and Stegun, 1970, p. 375, (9.6.10))

$$I_k(x) = \left(\frac{1}{2}x\right)^k \sum_{n=0}^{\infty} \frac{\left(\frac{1}{4}x^2\right)^n}{n!\Gamma(k+n+1)}.$$

By symmetry, the probability mass function for an arbitrary  $k \in \mathbb{Z}$  can be expressed as

$$p_k = e^{-\psi^+ - \psi^-} \left(\frac{\psi^+}{\psi^-}\right)^{k/2} I_{|k|}(2\sqrt{\psi^+ \psi^-}). \quad (3)$$

We note that  $E(L_t) = (\psi^+ - \psi^-)t$  and  $\text{Var}(L_t) = (\psi^+ + \psi^-)t$ . Hence if  $\psi^+ = \psi^-$  the process is a martingale.

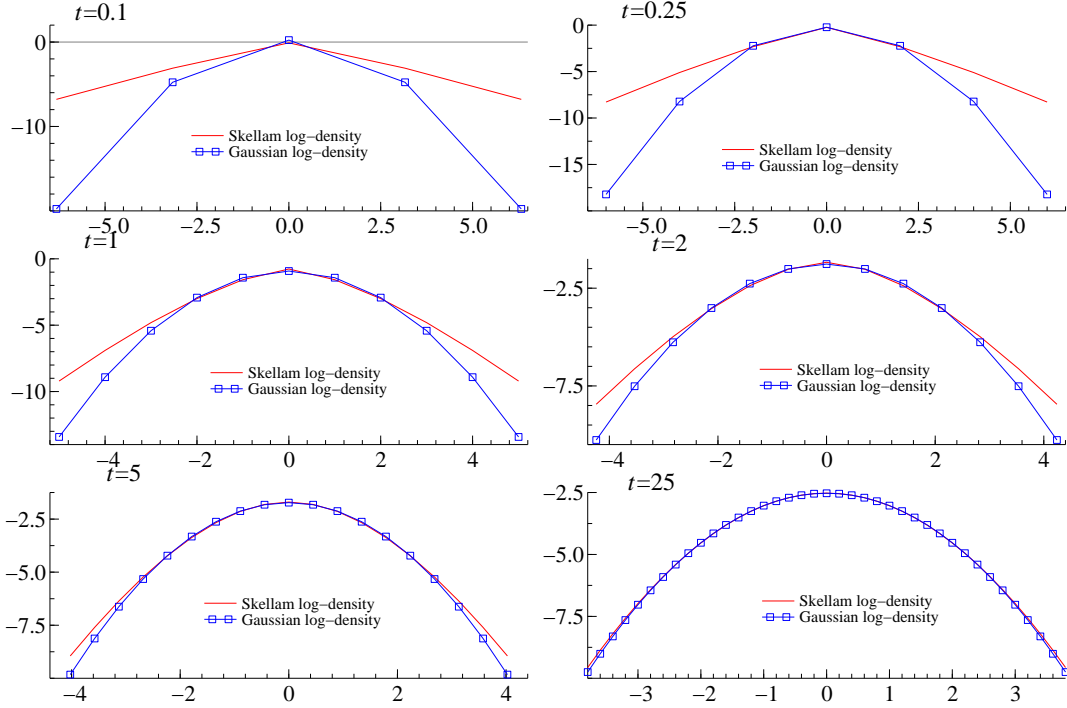


Figure 3: Log-density of the normal and Skellam distributions for  $S_t/\sqrt{t}$ . Code: `skellam.ox`.

**Remark 3** The most important special case is the **standard Skellam process** when  $\psi^+ = \psi^- = 1/2$  and then

$$C\{\theta \dagger L_1\} = \frac{1}{2} \left( -2 + e^{i\theta} + e^{-i\theta} \right) = -(1 - \cos \theta).$$

We will use the notation  $S_t$ ,  $t \in \mathbb{R}_{\geq 0}$ ,  $S_0 = 0$ , to denote the standard Skellam process. Clearly this is a martingale with unit variance per unit of time and

$$S_t \sim Sk\left(\frac{1}{2}t, \frac{1}{2}t\right), \quad \kappa_1 = 0, \quad \kappa_2 = 1, \quad \kappa_3 = 0, \quad \kappa_4 = 1.$$

Hence this is an integer-valued analogue of Brownian motion. This process has a unit expected number of price changes per unit of time. Let us study the distribution of  $S_t/\sqrt{t}$ . Expanding in small  $\theta$ ,

$$C\left\{\theta \dagger S_t/\sqrt{t}\right\} = \frac{\theta^2}{2} + \frac{\theta^4}{24t} + \dots$$

and hence, as  $t \rightarrow \infty$ , so  $S_t/\sqrt{t} \xrightarrow{d} N(0,1)$ . Figure 3 shows the log-density of  $S_t/\sqrt{t}$ . It is slightly sub-linear in the tails for small  $t$  and it becomes quadratic as  $t$  increases.

### 3.4 $\Delta$ NB Lévy process

#### 3.4.1 Negative binomial precursor

We now study a more general model, based upon the negative binomial distribution. The negative binomial distribution comes from mixing a Poisson distribution

$$\Pr(L_1^+ = k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \dots,$$

with a random intensity parameter  $\lambda$

$$\lambda \sim Ga\left(\delta, \frac{p}{1-p}\right), \quad p \in (0, 1), \quad \delta > 0, \quad \mathbb{E}(\lambda) = \delta \frac{p}{1-p}, \quad \text{Var}(\lambda) = \delta \left(\frac{p}{1-p}\right)^2,$$

which is gamma distributed. Then, the following is well known

$$\begin{aligned} \Pr(L_1^+ = k) &= \int_0^\infty \frac{\lambda^k e^{-\lambda}}{k!} \lambda^{\delta-1} \frac{\exp\{-\lambda(1-p)/p\}}{\left(\frac{p}{1-p}\right)^\delta \Gamma(\delta)} d\lambda = \frac{1}{k!} \frac{1}{\left(\frac{p}{1-p}\right)^\delta \Gamma(\delta)} \int_0^\infty \lambda^{k+\delta-1} \exp(-\lambda/p) d\lambda \\ &= \frac{1}{k!} \frac{1}{\left(\frac{p}{1-p}\right)^\delta \Gamma(\delta)} p^{\delta+k} \Gamma(\delta+k) \\ &= \frac{1}{k!} \frac{\Gamma(\delta+k)}{\Gamma(\delta)} (1-p)^\delta p^k, \end{aligned}$$

that is the negative binomial distribution, which we will write as  $NB(\delta, p)$ . This distribution goes back to Greenwood and Yule (1920). The first four cumulants of  $NB(\delta, p)$  are

$$\kappa_1 = \delta \frac{p}{1-p}, \quad \kappa_2 = \delta \frac{p}{(1-p)^2}, \quad \kappa_3 = \delta \frac{p+p^2}{(1-p)^3}, \quad \kappa_4 = \delta \frac{p+4p^2+p^3}{(1-p)^4},$$

and the cumulant function is

$$C\{\theta \dagger L_1^+\} = \delta \left\{ \log(1-p) - \log(1-pe^{i\theta}) \right\}.$$

It follows immediately, as is well known, that this distribution is infinitely divisible and it supports a negative binomial Lévy process with

$$L_t^+ \sim NB(t\delta, p).$$

This process is overdispersed as  $\kappa_{1,t}/\kappa_{2,t} = \kappa_1/\kappa_2 = 1-p \in (0, 1]$ . It is well known that the negative binomial process can also be generated as  $L_t^+ = N \circ T_t = N_{T_t}$ , where the subordinator  $T$  is a gamma process stochastically independent from  $N$  a standard Poisson process.

We recall that the negative binomial nests the Poisson distribution. In fact, reparameterising from  $(\delta, p)$  to  $(\psi, p)$  by letting

$$\psi = \delta \frac{p}{1-p},$$

then

$$\Pr(L_1^+ = k) = \frac{\psi^k \Gamma(\delta+k)}{k! \Gamma(\delta)} (1-p)^{\delta+k}$$

with

$$\mathbb{E}(\lambda) = \delta \frac{p}{1-p} = \psi, \quad \text{Var}(\lambda) = \delta \left( \frac{p}{1-p} \right)^2 = \psi \frac{p}{1-p}.$$

For fixed  $\psi$ , as  $p \downarrow 0$  so  $\lambda \xrightarrow{p} \psi$ . Hence, using this parameterisation, Poisson is the extreme case of  $p = 0$ . We will use this  $\psi$  parameterisation in our empirical work.

**Remark 4** *The negative binomial process can be thought of as a compound Poisson process*

$$L_t^* = \sum_{j=1}^{N_t} X_j, \quad X_j \sim i.i.d., \quad X \perp\!\!\!\perp N,$$

where the innovations are logarithmic variables

$$\Pr(X_j = k) = \frac{p^k}{|\log(1-p)| k}, \quad k = 1, 2, \dots,$$

while  $N_t$  is a Poisson process with intensity

$$\delta |\log(1-p)| = \psi \frac{(1-p)}{p} |\log(1-p)|.$$

A derivation of this known result will be given in Example 2 below. This implies

$$\log \Pr(X_j = k) = k \log p - \log k - \log \{-\log(1-p)\},$$

so the log-histogram of the innovations will appear approximately linear in the tails. Statistically this is a convenient form, it means the  $p$  parameter entirely controls the size of the moves and the  $\delta$  parameter can be freely set to control the intensity of the moves. Note as  $p \downarrow 0$  so  $|\log(1-p)| \sim p$  so  $\Pr(X_j = k) \simeq p^{k-1}/k$ , which will have nearly all of its mass at one; furthermore, the intensity tends to  $\psi$ .

### 3.4.2 $\Delta$ NB Lévy process in detail

We work with  $L_t$  where

$$L_t = L_t^+ - L_t^-, \quad L_t^+ \perp\!\!\!\perp L_t^-, \quad L_t^+ \sim NB(t\delta^+, p^+), \quad L_t^- \sim NB(t\delta^-, p^-),$$

with  $p^\pm \in (0, 1)$ ,  $\delta^\pm > 0$ .

Introducing the rising factorial, or Pochhammer symbol,

$$(a)_n = a(a+1) \cdots (a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)},$$

we have

$$\Pr(L_1^+ = k) = \frac{(\delta)_k}{k!} p^k (1-p)^\delta$$

and, for  $k \in \mathbb{N}_0$ , the point probabilities of  $L_t$  are

$$\begin{aligned}
p_k &= \sum_{n=0}^{\infty} \Pr(L_1^+ = n+k) \Pr(L_1^- = n) \\
&= \sum_{n=0}^{\infty} \frac{(\delta^+)_{n+k}}{(n+k)!} (p^+)^{n+k} (1-p^+)^{\delta^+} \frac{(\delta^-)_n}{n!} (p^-)^n (1-p^-)^{\delta^-} \\
&= (1-p^-)^{\delta^-} (1-p^+)^{\delta^+} (p^+)^k \sum_{n=0}^{\infty} \frac{(\delta^+)_{n+k}}{(n+k)!} \frac{(\delta^-)_n}{n!} (p^+ p^-)^n.
\end{aligned}$$

Now,  $(\delta)_{n+k} = (\delta)_k (\delta+k)_n$  and

$$\frac{\Gamma(n+k+1)}{\Gamma(k+1)} = (k+1)_n,$$

so

$$\begin{aligned}
p_k &= (1-p^-)^{\delta^-} (1-p^+)^{\delta^+} \frac{(p^+)^k (\delta^+)_k}{\Gamma(k+1)} \sum_{n=0}^{\infty} \frac{(\delta^+ + k)_n (\delta^-)_n}{(k+1)_n n!} (p^+ p^-)^n \\
&= (1-p^-)^{\delta^-} (1-p^+)^{\delta^+} \frac{(p^+)^k (\delta^+)_k}{k!} F(\delta^+ + k, \delta^-; k+1; p^+ p^-),
\end{aligned}$$

where

$$F(\alpha, \beta; \gamma; z) = \sum_{n=0}^{\infty} \frac{(\alpha)_n (\beta)_n}{(\gamma)_n} \frac{z^n}{n!}, \quad z \in [0, 1), \quad \alpha, \beta, \gamma > 0, \quad (4)$$

is the classical hypergeometric function which has many properties and applications (see, for example, (Abramowitz and Stegun, 1970, Ch. 15)). To be explicit, here  $\alpha = \delta^+ + m > 0$ ,  $\beta = \delta^- > 0$ ,  $\gamma = m + 1 \geq 1$ ,  $z = p^+ p^- \in (0, 1)$ . By symmetry, for any  $k \in \mathbb{Z}$ ,

$$p_k = (1-p^-)^{\delta^-} (1-p^+)^{\delta^+} \frac{(p^+)^{k^+} (p^-)^{k^-} (\delta^+)_{k^+} (\delta^-)_{k^-}}{k^+! k^-!} F(\delta^+ + k^+, \delta^- + k^-; k^+ + k^- + 1; p^+ p^-)$$

where  $k^+ = \max(0, k)$  and  $k^- = \max(0, -k)$ . This seems to be a new type of four parameter distribution. We write it as a  $\Delta NB(\delta^+, p^+, \delta^-, p^-)$  distribution or, using the parametrisation discussed above,

$$\Delta NB(\psi^+, p^+, \psi^-, p^-).$$

The latter is convenient for us as it allows a simple comparison with the Skellam distribution and it will be used throughout our empirical work.

Clearly

$$\begin{aligned}
\mathbb{C}\{\theta \dagger L_1\} &= \mathbb{C}\{\theta \dagger L_1^+\} + \mathbb{C}\{-\theta \dagger L_1^-\} = \delta^+ \left\{ \log(1-p^+) - \log(1-p^+ e^{i\theta}) \right\} \\
&\quad + \delta^- \left\{ \log(1-p^-) - \log(1-p^- e^{-i\theta}) \right\},
\end{aligned}$$

which directly demonstrates it is infinitely divisible. We call the resulting Lévy process a  $\Delta NB$  process and it has the property that

$$L_t \sim \Delta NB(t\delta^+, p^+, t\delta^-, p^-),$$

or in the alternative parameterisation  $L_t \sim \Delta NB(t\psi^+, p^+, t\psi^-, p^-)$ . Then

$$\begin{aligned}\kappa_1 &= \delta^+ \frac{p^+}{1-p^+} - \delta^- \frac{p^-}{1-p^-}, & \kappa_2 &= \delta^+ \frac{p^+}{(1-p^+)^2} + \delta^- \frac{p^-}{(1-p^-)^2}, \\ \kappa_3 &= \delta^+ \frac{p^+ + p^{+2}}{(1-p^+)^3} - \delta^- \frac{p^- + p^{-2}}{(1-p^-)^3}, & \kappa_4 &= \delta^+ \frac{p^+ + 4p^{+2} + p^{+3}}{(1-p^+)^4} + \delta^- \frac{p^- + 4p^{-2} + p^{-3}}{(1-p^-)^4}.\end{aligned}$$

**Remark 5** *The symmetric version of this distribution is distinct from a symmetric Skellam distribution  $Sk(\psi, \psi)$  with gamma distributed intensity  $\psi$ . Hence this process is not a Skellam process time changed by a gamma process.*

**Remark 6** *The most important special case is the standard symmetric process when*

$$\delta^+ = \delta^- = \delta, \quad p^+ = p^- = p, \quad \delta = \frac{1(1-p)^2}{2p},$$

then at time one

$$\kappa_1 = 0, \quad \kappa_2 = 2\delta \frac{p}{(1-p)^2} = 1, \quad \kappa_3 = 0, \quad \kappa_4 = 2\delta \frac{p + 4p^2 + p^3}{(1-p)^4} = \frac{1 + 4p + p^2}{(1-p)^2}.$$

Thus unlike the standard Skellam distribution it has the flexibility to deliver any value of  $\kappa_4 \geq 1$ . Of course as  $p^+ \downarrow 0$  so  $\kappa_4 \downarrow 1$  (the standard Skellam case), while as  $p^+ \uparrow 1$  then  $\kappa_4 \uparrow \infty$ . Figure 4 compares a sample path from the standard  $\Delta NB$  process, given in part (a) of the Figure, to that of a standard Skellam process, given in part (b). Clearly the  $\Delta NB$  process has a smaller number of jumps, but some of the jumps are more than a single tick. Part (c) shows a log-histogram of the returns from the  $\Delta NB$  process computing each return over one thousandth of a unit of time. It shows a very large probability of a zero, with some probability at  $\pm 1$ . What is important is there is a small positive probability of moves to  $\pm 2$  and even some observed  $\pm 3$ . For the Skellam process the corresponding log-histogram has no mass outside  $\pm 1$ . This is important empirically.

**Remark 7** *If  $\delta^+ = \delta^- = 1$  then  $L_1^\pm$  become geometric random variables with parameters  $p^\pm$ . As a result  $L_1$  is a “integer-valued skewed Laplace” random variable. It has the feature that*

$$p_k = \Pr(L_1 = k) = \frac{(1-p^-)(1-p^+)}{1-p^-p^+} (p^+)^{k^+} (p^-)^{k^-}, \quad k \in \mathbb{Z},$$

where  $k^+ = \max(0, k)$  and  $k^- = \max(0, -k)$ . This follows as  $F(k+1, 1; k+1; z) = \sum_{n=0}^{\infty} (1)_n \frac{z^n}{n!} = \frac{1}{1-z}$ ,  $(1)_n = \frac{\Gamma(n+1)}{\Gamma(1)} = n!$ . Some properties of this kind of distribution appear in Kozubowski and Inusah (2006) and Inusaha and Kozubowski (2006). The attractions of this distribution are two

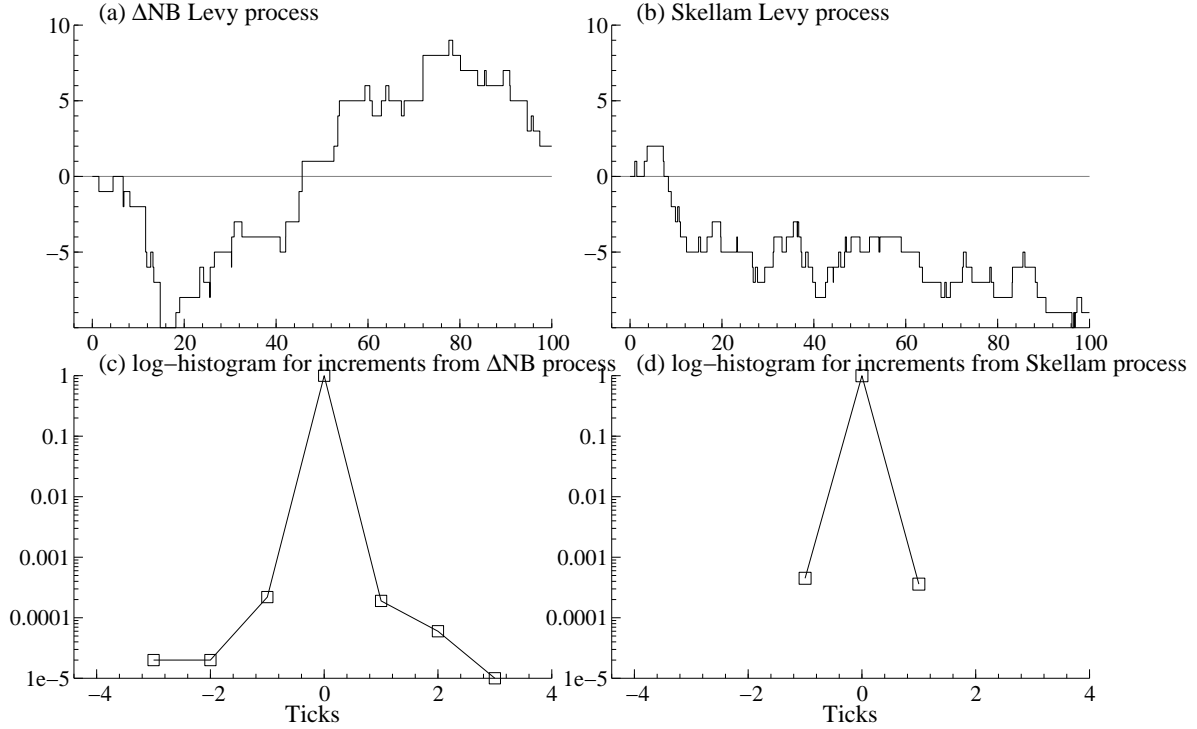


Figure 4: *Top figures: sample path from the standard  $\Delta NB$  (with  $p^+ = 0.32$ ) and Skellam Lévy processes. Bottom figures: log-histogram from the increments of  $\Delta NB$  and Skellam Lévy process. Code: hyper.ox.*

*fold: (i) it is very simple to estimate using a likelihood function, (ii) the log-probabilities decay linearly and this is in line with some empirical features found above. However,  $L_t$  is not skewed Laplace, instead its distribution is  $\Delta NB(t, p^+, t, p^-)$ .*

### 3.5 A generic precursor

The discussion of the negative binomial distribution, as a gamma time-changed Poisson process, is nested within the following setup which may be useful for the development of more general models. Let  $N$  be a Poisson process with unit intensity and  $T$  be a subordinator (i.e. a non-negative Lévy processes) such that  $N \perp\!\!\!\perp T$  and let

$$L^+ = N \circ T$$

be the subordination of  $N$  by  $T$  i.e.  $L_t^+ = N_{T_t}$ .

To analyse this class it is helpful to take a step back and introduce some well known mathematics through the kumulant function

$$\bar{K}(\theta \dagger X) = \log E \{ \exp(-\theta X) \}$$

for a random variable  $X \geq 0$ . Then it is well known that the Lévy-Khintchine representation for



all non-negative Lévy processes can be written as

$$\overline{\mathbb{K}}\{\theta \ddagger L_t^+\} = -at\theta - t \int_0^\infty (1 - e^{-\theta u}) \nu(du \ddagger L_1^+), \quad (5)$$

where the drift  $a \geq 0$  and  $\nu$  is a measure on  $\mathbb{R}_{>0}$  such that  $\int_0^\infty \min\{1, y\} \nu(du \ddagger L_1^+) < \infty$ . Hence all non-negative processes can be classified by their Lévy measure  $\nu(du \ddagger L_1^+)$  and the drift.

**Theorem 1**  $L^+ = N \circ T$  can be written as a compound Poisson process

$$L_t^+ = \sum_{j=1}^{N_t^*} U_j, \quad U_j \sim i.i.d., \quad N^* \perp\!\!\!\perp U,$$

where  $N_t^*$  is a Poisson process with constant rate

$$\lambda^{N^*} = -\overline{\mathbb{K}}(1 \ddagger T_1) < \infty$$

and  $U_j$  is concentrated on the positive integers, with

$$\Pr(U_j = m) = \left(\lambda^{N^*}\right)^{-1} q_m, \quad q_m = \int_0^\infty \frac{u^m}{m!} e^{-u} \nu(du \ddagger T_1).$$

**Proof.** Given in the Appendix.

This result gives a complete characterisation of this class of time-changed processes, showing it is always representable as a compound Poisson process. Further, the rate of the intensity is known, as is the probability function of the innovations  $U_j$ .

We call models of the form  $L_t = L_t^+ - L_t^-$ , where  $L_t^+ = N_{T_t^+}^+$  and  $L_t^- = N_{T_t^-}^-$ , “type  $P$  processes”, mimicing the nomenclature “type  $G$  processes” used where Brownian motion is time changed by a subordinator (see Rosinski (1991)).

**Example 1** Suppose  $T_t \sim IG(t\delta, \gamma)$ , which means it is inverse Gaussian. Then

$$\nu(du \ddagger T_1) = \frac{\delta}{\sqrt{2\pi}} u^{-\frac{3}{2}} e^{-\frac{1}{2}\gamma^2 u} du, \quad \overline{\mathbb{K}}(\theta \ddagger T_1) = \delta \left\{ \gamma - (\gamma^2 + 2\theta)^{1/2} \right\},$$

so that

$$\begin{aligned} q_m &= \frac{\delta}{\sqrt{2\pi}} \frac{1}{m!} \int_0^\infty u^m e^{-u} u^{-\frac{3}{2}} e^{-\frac{1}{2}\gamma^2 u} du = \frac{\delta}{\sqrt{2\pi}} \frac{1}{m!} \frac{\Gamma(m - \frac{1}{2})}{\left(1 + \frac{\gamma^2}{2}\right)^{m - \frac{1}{2}}} \\ &= \frac{\delta}{\pi} (2)^{\frac{1}{2}} B\left(\frac{3}{2}, m - \frac{1}{2}\right) \frac{1}{\left(1 + \frac{1}{2}\gamma^2\right)^{m - \frac{1}{2}}}, \end{aligned} \quad (6)$$

where  $B(\cdot)$  is the beta function and the intensity is  $\lambda^{N^*} = \delta \left\{ (\gamma^2 + 2)^{1/2} - \gamma \right\}$ . We call the resulting  $N_{T_t}$  a Poisson inverse Gaussian Lévy process or PIG process, noting that the PIG( $\delta, \gamma$ ) distribution of  $N_{T_1}$  was introduced by Holla (1967), see the discussion in, for example, Karlis and

Xekalaki (2005). If  $T_t^\pm$  are independent inverse Gaussian Lévy processes with parameters  $(\delta^\pm, \gamma^\pm)$  then we call  $L_t = N_{T_t^+}^+ - N_{T_t^-}^-$  a delta Poisson inverse Gaussian Lévy process and write it as

$$L_t \sim \Delta\text{PIG}(t\delta^+, \gamma^+, t\delta^-, \gamma^-).$$

**Example 2** This example reproduces results from the previous subsection but uses a different route.

Suppose  $T_t \sim \text{Ga}(t\delta, 1/\alpha)$ , so relating to previously  $\alpha = (1-p)/p$ . For this process

$$\nu(\text{du} \ddagger T_1) = \delta u^{-1} e^{-\alpha u} \text{du}, \quad \bar{K}(\theta \ddagger T_1) = -\delta \log\left(1 + \frac{\theta}{\alpha}\right).$$

So noting  $(1+\alpha)^{-1} = p$ , then

$$q_m = \frac{\delta}{m!} \int_0^\infty e^{-u} u^{m-1} e^{-\alpha u} \text{du} = \delta \frac{p^m}{m}, \quad \lambda^{N^*} = \delta \log\left(1 + \frac{1}{\alpha}\right) = \delta |\log(1-p)|. \quad (7)$$

The latter term is the intensity of the Poisson process  $N_t^*$ . Hence the law of the innovations for this compound Poisson process is

$$\Pr(U_j = m) = \frac{1}{|\log(1-p)|} \frac{p^m}{m}.$$

That is, the innovations follow the logarithmic distribution, which is well known to be infinitely divisible. Furthermore, the law of  $L_1^+$  is the negative binomial with point probabilities

$$\Pr(L_1^+ = k) = \binom{k+\delta-1}{k} (1-p)^\delta p^k.$$

It is well known that a Poisson number of i.i.d. logarithmic variables follows a negative binomial distribution and that the negative binomial is infinitely divisible. In fact, the negative binomial has the stronger property of being discrete selfdecomposable, cf. Steutel and Van Harn (2004).

Finally we establish a converse to Theorem 1.

**Theorem 2** Let  $Y$  be a compound Poisson process of the form  $Y_t = \sum_{j=1}^{N_t} U_j$  where the innovations  $U_j$  take only positive integer values, with point probabilities  $p_m$ ,  $m = 1, 2, \dots$ . Let  $g$  be the moment generating function of  $U_j$ , i.e.  $g(\phi) = \sum_{m=1}^\infty p_m \phi^m$  ( $\phi \in [0, 1]$ ) and let  $h(\phi) = 1 - g(1 - \phi)$ . Then  $Y$  is representable as a time-changed Poisson process, where the time change is a subordinator, if and only if  $h$  is of the form  $h(\phi) = \bar{K}\{\phi \ddagger S\}$  where  $\bar{K}\{\phi \ddagger S\}$  is the kumulant function of a positive infinitely divisible random variable  $S$ .

**Proof.** Given in the Appendix.

### 3.5.1 Delta Poisson tempered stable processes

Let  $p(x; \kappa, \delta)$  denote the probability density function of the positive  $\kappa$ -stable law  $S(\kappa, \delta)$  with cumulant transform  $-\delta(2\theta)^\kappa$ ,  $0 < \kappa < 1$ , and let  $p(x; \kappa, \delta, \gamma)$  denote the tempered (exponentially tilted) version of  $p(x; \kappa, \delta)$  defined by

$$p(x; \kappa, \delta, \gamma) = e^{\delta\gamma} p(x; \kappa, \delta) e^{-\frac{1}{2}\gamma^{1/\kappa}x}. \quad (8)$$

The distribution with density (8) ( $\kappa \in (0, 1)$ ,  $\delta > 0$ ,  $\gamma \geq 0$ ) will be referred to as a *tempered stable law* and we denote it by  $TS(\kappa, \delta, \gamma)$ . This process is infinitely divisible and so supports a Lévy process. Simulation of tempered stable variables is studied by Brix (1999) and extensively by Devroye (2009), the latter provides a very reliable algorithm across all parameters.

Suppose  $T_t \sim TS(\kappa, t\delta, \gamma)$ , a tempered stable process. Here  $0 < \kappa < 1$ ,  $\delta, \gamma > 0$ . For this process

$$\nu(du \dagger T_1) = cu^{-1-\kappa} \exp\left(-\frac{1}{2}\gamma^{1/\kappa}u\right) du, \quad \bar{K}(\theta \dagger T_1) = \delta\gamma - \delta\left(\gamma^{1/\kappa} + 2\theta\right)^\kappa,$$

so  $T_t \sim TS(\kappa, t\delta, \gamma)$  and

$$E(T_1) = 2\kappa\delta\gamma^{(\kappa-1)/\kappa}, \quad \text{Var}(T_1) = 4\kappa(1-\kappa)\delta\gamma^{(\kappa-2)/\kappa}.$$

Here  $c = \delta 2^\kappa \kappa / \Gamma(1-\kappa)$ . This process nests the gamma ( $\kappa \downarrow 0$ ), the positive stable ( $\gamma \downarrow 0$ ) and inverse Gaussian ( $\kappa = 1/2$ ). Then the intensity of  $N_t^\pm$  is  $\lambda^{N^*} = \delta \left\{ (\gamma^{1/\kappa} + 2)^\kappa - \gamma \right\}$  and

$$q_m = \frac{c}{m!} \int_0^\infty u^{m-\kappa-1} e^{-u(1+\frac{1}{2}\gamma^{1/\kappa})} du = \frac{c}{m!} \frac{\Gamma(m-\kappa)}{\left(1 + \frac{\gamma^{1/\kappa}}{2}\right)^{m-\kappa}}.$$

Note that Shepard's formula states that as  $m \rightarrow \infty$  so  $\Gamma(m+a)/\Gamma(m+b) \sim m^{a-b}$ , which means that

$$q_m \sim cg^\kappa m^{\kappa-1} \exp(-gm), \quad g = \log\left(1 + \frac{\gamma^{1/\kappa}}{2}\right).$$

Hence this distribution has a semi-heavy tail and will be heavy tailed in the limit as  $\gamma \downarrow 0$ . We call this the Poisson tempered stable distribution.

This distribution was derived using a different approach by Zhu and Joe (2009), who call it a “generalised Poisson-inverse Gaussian family”. The class of tempered stable distributions was introduced by Tweedie (1984). See also Hougaard (1986), Jørgensen (1987), Brix (1999) and Rosinski (2007). The normal variance-mean mixtures with tempered stable mixing were introduced by Barndorff-Nielsen and Shephard (2001b), while Carr et al. (2002) have used a related type of stable tempered process in finance under the nomenclature “CGMY processes”. If  $T_t^\pm$  are

independent tempered stable processes with parameters  $(\kappa^\pm, \delta^\pm, \gamma^\pm)$  then we call  $L_t = N_{T_t^+}^+ - N_{T_t^-}^-$  a delta Poisson tempered stable process and write it as

$$L_t \sim \Delta PTS(\kappa^+, t\delta^+, \gamma^+, \kappa^-, t\delta^-, \gamma^-).$$

Finally it makes sense at some points to reparameterise this process so using

$$\psi^\pm = 2\kappa^\pm \delta^\pm (\gamma^\pm)^{(\kappa^\pm - 1)/\kappa^\pm} = \mathbb{E}(T_1^\pm),$$

instead of  $\delta^\pm$ .

### 3.6 Integer-valued Lévy processes, semimartingales and arbitrage

Any Lévy process with integer-valued and finite measure  $\nu$  can be written as

$$L_t = L_t^+ - L_t^- = A_t + M_t,$$

where

$$A_t = t(\kappa_1^+ - \kappa_1^-), \quad M_t = (L_t^+ - t\kappa_1^+) + (L_t^- - t\kappa_1^-).$$

Hence  $A$  is of locally bounded variation and  $M$  is a martingale, so of course  $L$  is a semimartingale (indeed all Lévy processes are semimartingales). As a result (Delbaen and Schachermeyer, 2006, Theorem 1) applies, which states that the concept of “no free lunch without vanishing risk” is equivalent to the existence of an equivalent martingale measure (EMM)  $Q$ .

Suppose we have a semimartingale that is locally absolutely continuous to  $L$  then necessarily the semimartingale is a Lévy process  $L^\#$  whose integer-valued Levy measure  $\nu^\#$  is equivalent to  $\nu$  (e.g. (Barndorff-Nielsen and Shiryaev, 2010, Theorem 6.1 and Example 6.2)). There are of course many such processes  $L^\#$ , and one way to obtain an EMM would be to use an Esscher transformation (e.g. (Barndorff-Nielsen and Shiryaev, 2010, Section 7.1.3)). However, then  $L^\#$  would not be a integer-valued Lévy process, and there is some appeal in choosing  $L^\#$  to be also a integer-valued Lévy process. All one has to do is to change one or more of the point masses of  $\nu$  so as to ensure that  $\mathbb{E}(L_t^\#) = 0$  while keeping the total mass of  $\nu^\#$  finite. Note that having  $\mathbb{E}(L_t^\#) = 0$  is then equivalent to requiring that the integral of  $\int x\nu^\#(dx) = 0$ .

**Example 3** Suppose the Lévy process  $L_t \sim Sk(t\psi^+, t\psi^-)$ , then  $L_t^\# \sim Sk(t\psi, t\psi)$  where  $\psi \in \mathbb{R}_{>0}$ . Hence there are an infinite number of EMMs in this case.

**Example 4** For a call option written at time 0 with an exercise price  $K$  at time  $t$ , then  $K$  must be an integer for it to be tradeable. We then have the integer-valued Lévy process option price

$$C(L_0, t, K) = \exp(-rt)\mathbb{E}^\# \left[ \left\{ \max(L_t^\# - K, 0) \right\} \mid L_0^\# = L_0 \right]$$

$$\begin{aligned}
&= \exp(-rt) \mathbb{E}^\# \left[ \max \left\{ \left( L_t^\# - L_0 \right) - (K - L_0), 0 \right\} \mid L_0^\# = L_0 \right] \\
&= \exp(-rt) \mathbb{E}^\# \left\{ \max \left( L_t^\circ - K^\circ, 0 \right) \right\}, \quad L_t^\circ = L_t^\# - L_0, \quad K^\circ = K - L_0, \\
&= \exp(-rt) \sum_{j=1}^{\infty} j \Pr(L_t^\circ = j + K^\circ),
\end{aligned}$$

where  $L_t^\circ$  is the increment in the integer-valued Lévy process  $L^\#$  from time 0 to time  $t$ . Here  $\mathbb{E}^\#$  denotes the expectation under  $L^\#$ . Note that  $C(L_0, t, K)$  only exists when  $L_0$  is an integer so the usual Black-Scholes risk management tools “delta”  $\partial C(L_0, t, K) / \partial L_0$  and “gamma”  $\partial^2 C(L_0, t, K) / \partial L_0^2$  do not exist. This integer driven result is distinct from the usual features of option pricing under Lévy processes where the Lévy measure has support on the real line (e.g. (Cont and Tankov, 2004, Chapter 12)). It is related to, but distinct from, Pelsser and Vorst (1994) and Chung and Shackleton (2002) who studied greeks for binomial tree approximations to options for geometric Brownian motion. In the concrete Skellam case  $L_t^\# \sim Sk(t\psi, t\psi)$  then

$$C(L_0, t, K) = \exp \{ - (r + 2\psi) t \} \sum_{j=1}^{\infty} j I_{|j+K-L_0|} (2t\psi),$$

recalling  $I_k(x)$  is the modified Bessel function of the third kind. Such option pricing formula may be helpful in valuing the economic costs of limit orders more accurately due to the inherent tick structure of the limit order book, e.g. Copeland and Galai (1983).

This approach to constructing  $L^\#$  has the advantage that if the general time change,  $T$ , which is not necessarily a subordinator, is constructed to be probabilistically independent of the martingale integer-valued Lévy process  $L^\#$  then  $L_{T_t}^\#$  is also an integer-valued martingale. This allows for the introduction of diurnal and time-varying volatility features.

**Example 5** If we extend Example 4 to the time-changed case then one approach is to compute

$$C(L_0, t, K) = \exp(-rt) \sum_{j=1}^{\infty} j \mathbb{E}^{T_t} \left\{ \Pr(L_{T_t}^\circ = j + K^\circ \mid T_t) \right\}.$$

Here  $\mathbb{E}^{T_t}$  denotes the expectation over the distribution of the random  $T_t$ . This approach echoes the approach to option pricing under Gaussian stochastic volatility by Renault and Touzi (1996).

## 4 Fitting CP processes to futures tick data

### 4.1 General case

As we discussed in Section 3.1 one approach is to model

$$L_t = \sum_{j=1}^{N_t} C_j,$$

where  $\{C_j\}$  are i.i.d. integer-valued innovations independent from the Poisson process  $N$  which in turn generates the times of trades. If  $C_j$  has a distribution called  $G$ , then we will call  $L_t$  a CP- $G$  process.

An approach to inference is to estimate the intensity of  $N$  by counting the number of data points during a day and separately estimating the probability function of  $C$ . We will focus on this approach in this section.

Write a sample of innovations as  $C_1, \dots, C_{N_1}$  and then a simple non-parametric estimate of the discrete probabilities is

$$p_k = \frac{1}{N_1} \sum_{j=1}^{N_1} 1_{C_j=k}, \quad k \in \mathbb{Z},$$

which we will compare to various parametric fits written as  $g_k = \Pr(C_j = k)$ .

Throughout we will use the log-likelihood as a measure of fit for  $G$ . It is defined as

$$\log L(G) = \sum_{j=1}^{N_1} \log g_{C_j},$$

evaluating the probability function only at points where there have been observations. Notice that  $\log L(p)$  maximises the potential log-likelihood, for  $p_k$  is the non-parametric maximum likelihood estimator of  $\Pr(C_j = k)$ .

## 4.2 Likelihood

For all the models considered in this paper we can directly compute the likelihood. This is built out of three straightforward observations. First, for a homogeneous Poisson process with intensity  $\psi > 0$  the joint density of the arrival times  $\tau = (\tau_1, \dots, \tau_k)'$  between 0 and  $T$  is well known to be (e.g. (Cox and Lewis, 1966, p. 27))

$$p(\tau|\psi) = \exp(-T\psi) \psi^k.$$

Second, for a compound Poisson process the jumps or “marks”  $C = (C_1, \dots, C_k)'$  are independent of the times, so the joint density is

$$p(C, \tau|\theta) = \exp(-T\psi) \psi^k \prod_{j=1}^k \Pr(C_j|\theta),$$

where  $\theta$  denotes a general parameter. Finally, we put this together and it delivers

$$\begin{aligned} p(L^+|\theta)p(L^-|\theta) &= \exp(-T\psi^+) (\psi^+)^{N_T^+} \prod_{j=1}^{N_T^+} \Pr(C_j^+|\theta^+) \\ &\quad \times \exp(-T\psi^-) (\psi^-)^{N_T^-} \prod_{j=1}^{N_T^-} \Pr(C_j^-|\theta^-). \end{aligned}$$

Here  $\psi^\pm$  is the intensity of  $L^\pm$  and  $C_j^\pm$  are the corresponding jumps, while  $\theta = (\psi^+, \psi^-, \theta^+, \theta^-)'$ . Of course for a Skellam Lévy process  $\Pr(C_j^- = 1) = \Pr(C_j^+ = 1) = 1$ . If  $\psi^+, \psi^-, \theta^+, \theta^-$  are variation free (e.g. Engle et al. (1983)) then the maximum likelihood estimators of  $\psi^+$  and  $\psi^-$  are  $\hat{\psi}^+ = N_T^+/T$ ,  $\hat{\psi}^- = N_T^-/T$ .

### 4.3 CP-Skellam and CP- $\Delta$ NB processes

Table 2 shows the ML estimates of the innovation distributions in the CP-Skellam and CP- $\Delta$ NB cases for the Euro-Dollar IMM FX futures contract on 7th and 10th of November, 2008. Figures 1 and 2 show the corresponding computed probability function, as well as superimposing the corresponding non-parametric fit, in the lower right graphs.

In the case of the relatively tranquil 10th November sample path, the Skellam distribution is not too poor, it is slightly thinner in the tails than the data and perhaps struggles at  $\pm 4$  ticks. The  $\Delta$ NB is statistically stronger, but there are small signs that even it is not sufficiently fat tailed. The difference between the CP-Skellam and CP- $\Delta$ NB is modest although statistically significant (recall the  $\Delta$ NB nests the Skellam model as a special case).

	CP-Skellam		CP- $\Delta$ NB				logL
	$\hat{\psi}^+$	$\hat{\psi}^-$	$\hat{\psi}^+$	$\hat{\psi}^-$	$\hat{p}^+$	$\hat{p}^-$	
Euro 7/11/08	0.1810	0.1817					-38,705
			0.1734	0.1742	0.2366	0.2286	-37,542
Euro 10/11/08	0.1328	0.1375					-24,739
			0.1314	0.1360	0.0868	0.0700	-24,655
ESPC 10/11/08	0.0329	0.0339					-47,002
			0.0329	0.0338	0.0072	0.0087	-46,993
CLN 10/11/08	0.1334	0.1378					-72,669
			0.1254	0.1298	0.4619	0.4525	-64,237
TNC 10/11/08	0.0539	0.0516					-10,662
			0.0539	0.0517	0	0	-10,662

Table 2: *ML estimation of CP-Skellam and CP- $\Delta$ NB models. Each fit is for data from the 7th or 10th of November, 2008.*

For the much more challenging 7th November case the differences are more stark. The Skellam log-probability function looks sub-linear and cannot really deal with data which are at  $\pm 6$  ticks. The  $\Delta$ NB log-probability function is linear in the tails (but more flexible at the centre of the distribution and so allows it to be better fitting in the tails), like a Laplace distribution.

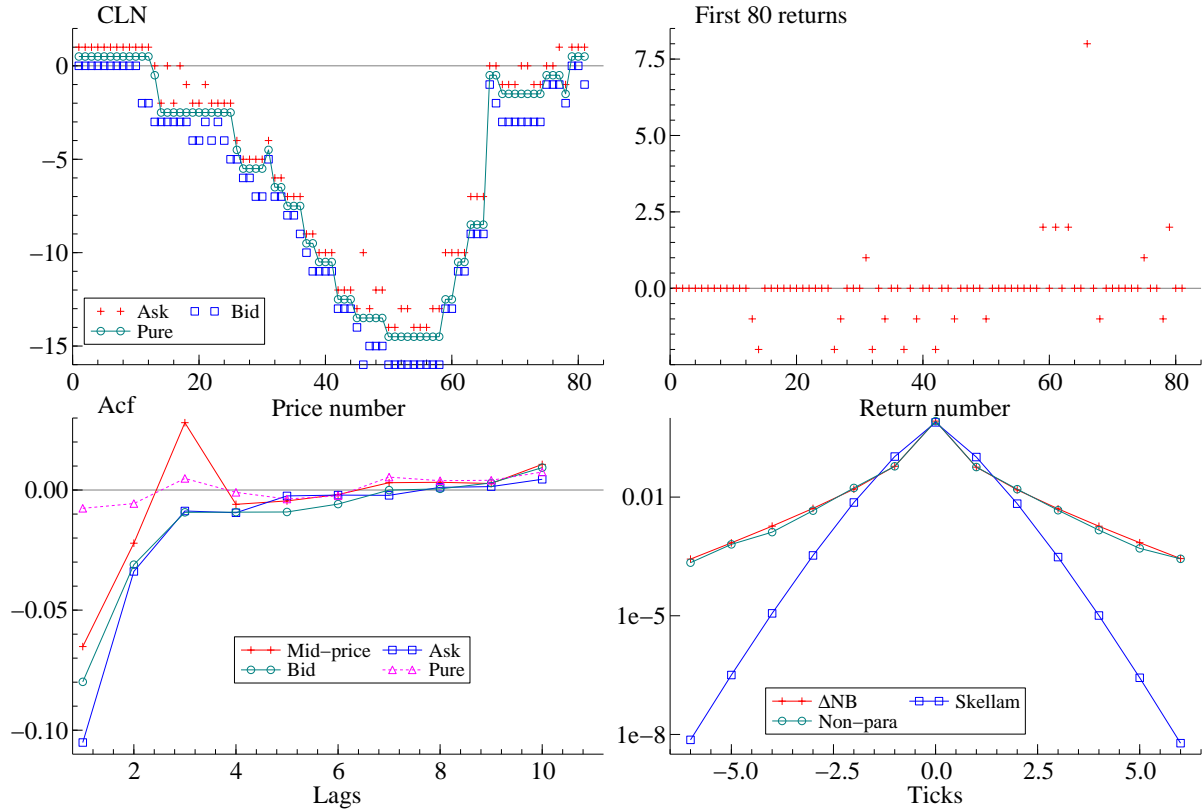


Figure 5: *Nymex/CME benchmark crude oil contract (CLN) on 10th November 2008. Top left: ask, bid and pure mid-price for the first 80 trades of the day. Top right: returns from pure mid-price. They are all integers. Bottom left: correlogram for mid-price, ask, bid and pure mid-price for entire day. Bottom right: log-histogram of pure mid-price returns: non-parametric, Skellam and  $\Delta$ NB distributions.*

## 4.4 Other examples

### 4.4.1 Oil futures

Next we will look at the Nymex/CME benchmark crude oil contract (CLN) series on 10th November 2008. The tick size is 0.01 of a unit, i.e. prices move from, for example, 64.41 to 64.42 dollars per barrel. On the 10th November there were 90,762 trades.

The results from the Skellam and  $\Delta$ NB distribution are given in Table 2 and Figure 5. It shows again the  $\Delta$ NB distribution doing much better in the tails of the distribution and having a substantially higher likelihood. Here  $p^+$  and  $p^-$  have roughly similar values, which means the estimated distribution is roughly symmetric in this case. Interestingly the  $\Delta$ NB tails decay less fast than linearly. Indeed this is a pretty heavy tailed integer-valued process.

### 4.4.2 Ten Year US treasury note

The Ten Year US treasury note (TNC) series on 10th November 2008 has a tick size of  $\frac{1}{64}$  of a dollar, so the movements are from, for example,  $115\frac{34}{64}$  to  $115\frac{35}{64}$  dollars. On the 10th November



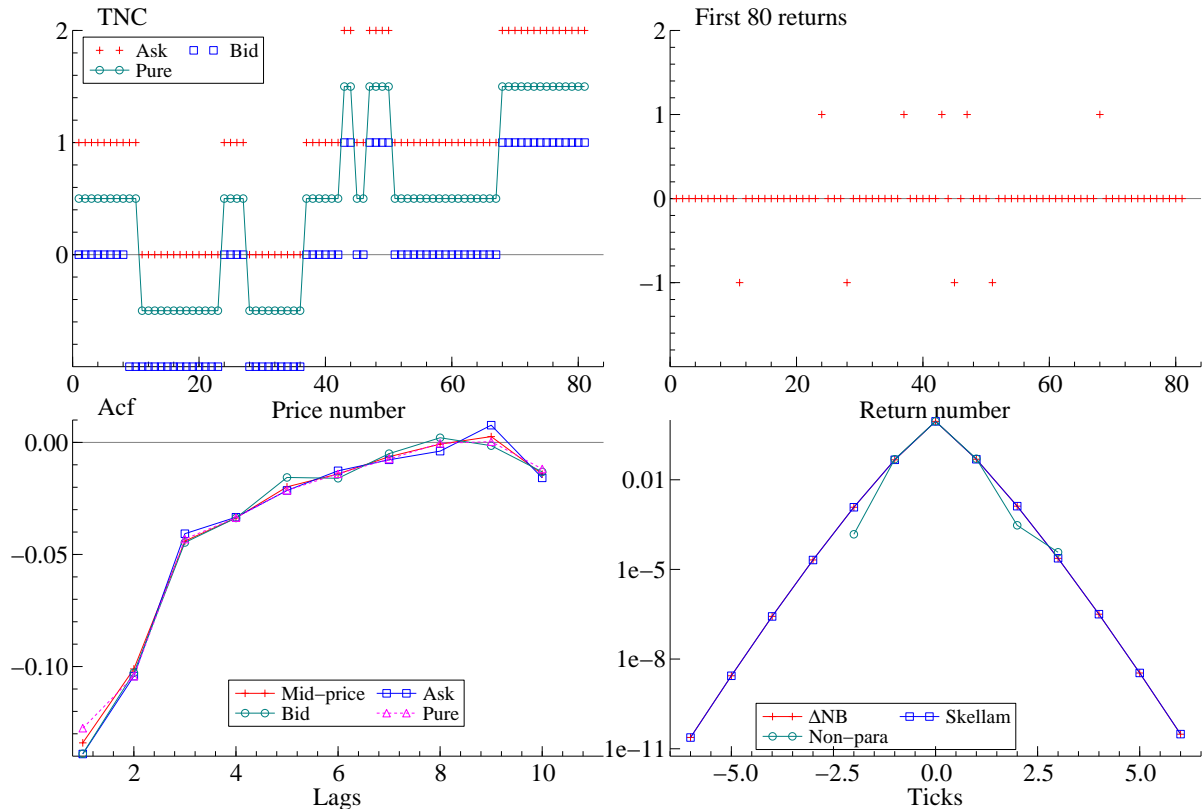


Figure 6: *Ten Year US treasury note (TNC) contract on 10th November 2008. Top left: ask, bid and pure mid-price for the first 80 trades of the day. Top right: returns from pure mid-price. They are all integers. Bottom left: correlogram for mid-price, ask, bid and pure mid-price for entire day. Bottom right: log-histogram of pure mid-price returns: non-parametric, Skellam and  $\Delta NB$  distributions.*

there were 26,754 trades. Of these observations only a tiny fraction of moves in the pure mid-price which are larger than  $\pm 1$  hence for this dataset the Skellam distribution will be nearly satisfactory. This is reflected in Table 2 which shows no improvement by using the more complicated  $\Delta NB$  distribution with the estimated  $p^+$  and  $p^-$  parameters being close to zero. The resulting graphs are in Figure 6.

#### 4.4.3 The mini S&P500 contract

Finally we look at the mini S&P500 contract (ESPC) series on 10th November 2008. The corresponding graphs are in Figure 7. The tick size is 0.25 of a unit, i.e. prices move from, for example, 952.00 to 951.75 dollars. On the 10th November there were 163,974 trades. Again for these data the Skellam distribution seems satisfactory as there is hardly any mass outside  $\pm 1$  ticks, but there is a tiny amount of unmodelled mass at quite large moves, such as  $\pm 5$  ticks. Table 2 shows no improvement by using the more complicated  $\Delta NB$  model. Note the correlogram for pure mid-prices changes is very slightly closer to that of white noise than the corresponding correlogram for

the mid-price changes.

## 5 Estimating generic integer-valued Lévy processes from futures tick data

### 5.1 Econometric framework

We now turn to the perhaps more interesting problem of the econometrics of integer-valued Lévy processes estimated directly from futures tick data.

We will write a continuous time pure mid-price process during a single day as

$$L_t = L_0 + \sum_{j=1}^{N_t} C_j = L_0 + \sum_{j=1}^{N_t^+} C_j^+ - \sum_{j=1}^{N_t^-} C_j^-, \quad t \in [0, 1],$$

where  $N_t$  is the number of trades up to time  $t$ ,  $N_t^+$  are the number of trades which deliver an uptick in the price and  $N_t^-$  are the number of trades which yield a downtick in the price. Clearly  $N_t \geq N_t^+ + N_t^-$  as many trades occur without the pure mid-price moving. Here the innovations are  $C_j^+, C_j^- \in \{1, 2, \dots\}$ . One of the attractive features of the high frequency data is that we are able to separately observe the five processes  $N_t, N_t^+, N_t^-, C_j^+, C_j^-$ . This is helpful econometrically. This component view of high frequency data echoes earlier work by, for example, Engle (2000), Rydberg and Shephard (2003), Barndorff-Nielsen et al. (2009) and Russell and Engle (2010).

**Remark 8** *We can go from a compound Poisson process  $\{N_t, C_j\}$  for trades, which includes innovations of zeros, into a Lévy model for  $\{N_t^+, N_t^-, C_j^+, C_j^-\}$  which exclude the zeros. In particular, writing the intensity of  $N_t$  as  $\lambda$ , then*

$$\begin{aligned} \lambda^+ &= \lambda \Pr(C_j \geq 1), & \lambda^- &= \lambda \Pr(C_j \leq -1), \\ \Pr(C_j^+ = k) &= \frac{\Pr(C_j = k)}{\Pr(C_j \geq 1)}, & \Pr(C_j^- = k) &= \frac{\Pr(C_j = -k)}{\Pr(C_j \leq -1)}, \quad k = 1, 2, \dots \end{aligned}$$

### 5.2 Skellam Lévy process and one-tick markets

For a Skellam Lévy process then

$$L_t - L_s \sim Sk((t-s)\psi^+, (t-s)\psi^-), \quad t > s.$$

But in continuous time the price process is constant until an innovation hits. These are  $\{C_j^+, C_j^-\}$  and are degenerate, being a sequence of ones with probability one. Hence in that case

$$L_t = L_0 + N_t^+ - N_t^-, \quad t \in [0, 1].$$

This process does not allow instantaneous moves in the price of more than one tick, which limits its direct application to so-called one-tick markets (see, for example, Field and Large (2008)). Hence

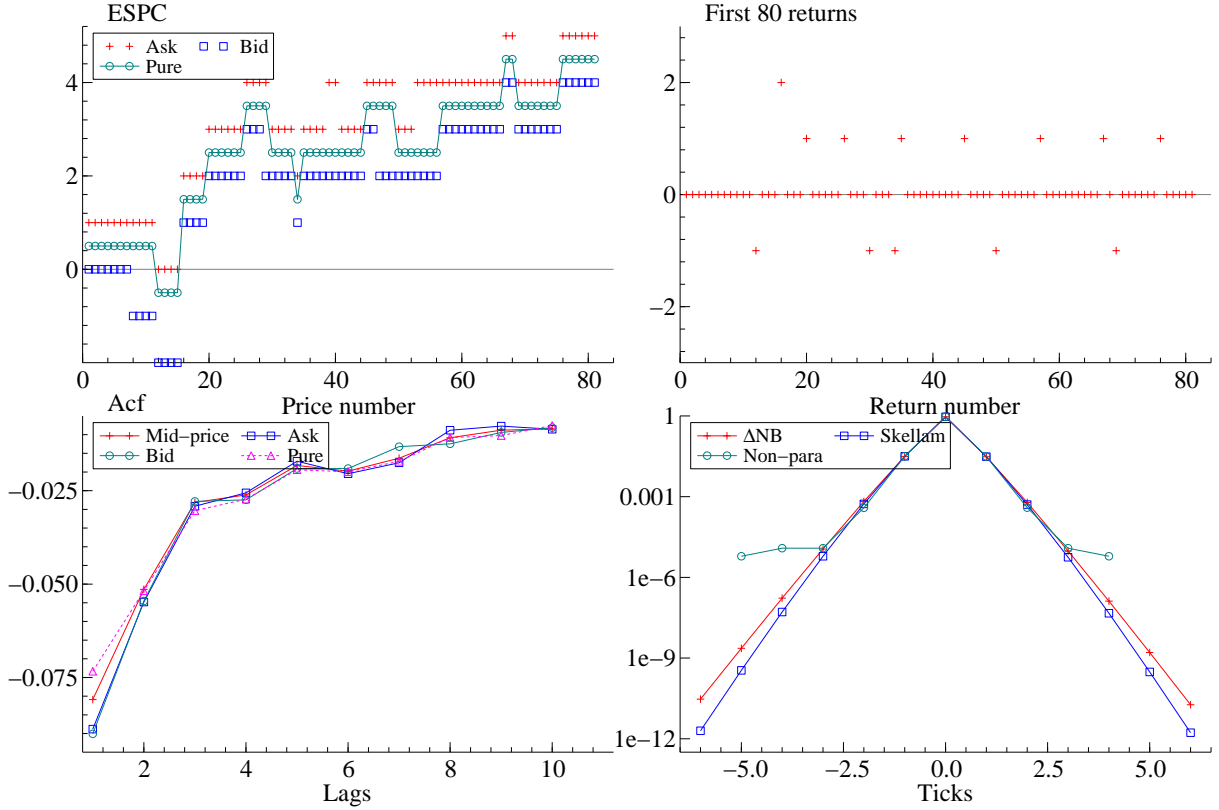


Figure 7: *Mini S&P500 contract (ESPC) on 10th November 2008. Top left: ask, bid and pure mid-price for the first 80 trades of the day. Top right: returns from pure mid-price. They are all integers. Bottom left: correlogram for mid-price, ask, bid and pure mid-price for entire day. Bottom right: log-histogram of pure mid-price returns: non-parametric, Skellam and  $\Delta NB$  distributions.*

the Skellam Lévy process is fundamentally different from the CP-Skellam process. In the latter cases the innovations can be larger than one. This single tick empirical limitation of the Skellam Lévy process means we will not continue with its application here.

### 5.3 Generic integer-valued Lévy processes

#### 5.3.1 Econometrics

If we return to the generic approach of Section 3.5 then the conditional likelihood for positive moves is

$$\log L(p^+; C_1^+, \dots, C_{N_1^+}^+ | N_1^+) = \sum_{j=1}^{N_1^+} \log \left\{ \Pr(C_j^+) \right\} = \sum_{j=1}^{N_1^+} \left\{ \log q_{C_j^+} - \log \left( \lambda^{N_1^+} \right) \right\}.$$

The same approach can be followed for negative moves.

**Example 6** *In the  $\Delta PTS$  process case then*

$$\lambda^{N^+} = \delta^+ \left\{ \left( (\gamma^+)^{1/\kappa^+} + 2 \right)^{\kappa^+} - \gamma^+ \right\}, \quad q_m = \frac{\delta^+ 2^{\kappa^+} \kappa^+ / \Gamma(1 - \kappa^+)}{m!} \frac{\Gamma(m - \kappa^+)}{\left( 1 + \frac{\gamma^{1/\kappa^+}}{2} \right)^{m - \kappa^+}}.$$

Notice that  $\log L(p^+; C_1^+, \dots, C_{N_1}^+ | N_1^+)$  does not depend upon  $\delta$ , solely  $\kappa^+$  and  $\gamma^+$ . Once  $\kappa^+$  and  $\gamma^+$  are estimated as  $\widehat{\kappa}^+$  and  $\widehat{\gamma}^+$ , respectively, then

$$\widehat{\delta}^+ = \frac{N_1^+}{\left(\widehat{\gamma}^+ 1/\widehat{\kappa}^+ + 2\right)^{\widehat{\kappa}^+} - \widehat{\gamma}^+}.$$

Throughout we then reparameterise  $\delta^\pm$  to

$$\psi^\pm = 2\kappa^\pm \delta^\pm (\gamma^\pm)^{(\kappa^\pm - 1)/\kappa^\pm},$$

so that  $E(L_t^+ | L_0^+) = L_0^+ + t\psi^+$ .

**Example 7** In the  $\Delta NB$  process special case

$$\lambda^{N^+*} = \delta^+ |\log(1 - p^+)|, \quad q_m = \delta^+ \frac{(p^+)^m}{m}. \quad (9)$$

Notice that  $\log L(p^+; C_1^+, \dots, C_{N_1}^+ | N_1^+)$  does not depend upon  $\delta^+$ , solely  $p^+$ , and has a very simple form in this case with  $\sum_{j=1}^{N_1^+} C_j^+$  being a sufficient statistic for  $p^+$ . Throughout we use the parameterisation  $\delta^+$  to  $\psi^+ = \delta^+ p^+ / (1 - p^+)$ .

### 5.3.2 Empirical results

All of the resulting fits look at the entire day and the results are given in Tables 3 and 4. Table 3 shows the counted up and down moves on each day which determines the intensity of the process, while the estimated  $\psi^+$  and  $\psi^-$  are the estimated expected total up and down ticks seen during the day. These are the same whatever model is being fitted. In the ESPC and TNC cases these

	Intensities			
	Up	Down	$\widehat{\psi}^+$	$\widehat{\psi}^-$
Euro 7/11/08	5,298	5,348	6,513	6,544
Euro 10/11/08	3,468	3,633	3,861	4,015
ESPC 10/11/08	5,148	5,292	5,232	5,382
CLN 10/11/08	7,320	7,649	10,612	11,009
TNC 10/11/08	1,363	1,308	1,373	1,312

Table 3: *ML estimation of intensities*  $\psi^+ = E(L_1^+)$  and  $\psi^- = E(L_1^-)$ . Also recorded are the number of updates which move the price up or down.

are slightly above the counts, for the other assets they are quite a lot above the counts due to those series having quite frequent multiple tick moves. The tail thickness of up moves is determined by  $p^+$  and show quite thick tails for the Euro and CLN futures prices.

The fitted probabilities for the  $\Pr(C_j^+ = k)$  from the  $\Delta NB$  Lévy process are shown in Figure 8 for the five series, together with a non-parametric fit and the result for the  $\Delta PTS$  process. The

	Estimated $\Delta\text{NB}$ Lévy process				Estimated $\Delta\text{PTS}$ Lévy process					
	$\hat{p}^+$	logL	$\hat{p}^-$	logL	$\hat{\gamma}^+$	$\hat{\kappa}^+$	logL	$\hat{\gamma}^-$	$\hat{\kappa}^-$	logL
Euro 7/11/08	.329	-3,089	.323	-3,070	1.37	0.386	-3,076	1.38	0.350	-3,060
Euro 10/11/08	.190	-1,272	.178	-1,261	1.00	0.001	-1,272	1.00	0.002	-1,261
ESPC 10/11/08	.031	-424	.033	-449	2.46	0.948	-401	1.79	0.959	-415
CLN 10/11/08	.504	-6,497	.497	-6,719	1.04	0.072	-6,496	1.00	0.001	-6,719
TNC 10/11/08	.014	-59	.006	-27	5.25	0.953	-57	1.02	0.003	-27

Table 4: *ML estimation of  $\Delta\text{NB}$  Lévy process. Each fit uses all the data on that day. Up moves records the number of upmovements during the day, downmoves looks at down moves.*

graphs are reasonably promising, although there is some evidence that the  $\Delta\text{NB}$  distribution is slightly too thin for the Euro series on the 7th November. The most significant issue is shown for ESPC, where the moves are nearly almost always one or two ticks, but there is a small chance of quite a large move. The  $\Delta\text{NB}$  process is quite unable to pick up those surprise shocks, but the  $\Delta\text{PTS}$  copes. These large moves drive the high estimated values for  $\kappa^\pm$ . The graphs also shows the probabilities for  $\Pr(C_j^- = k)$ , the results are broadly similar. Of course there is no probability at the atom at zero for these innovations.

## 6 Conclusion

In this paper we developed an exploratory analysis of integer-valued low latency financial data. Our focus is on the unconditional distributional features of returns at times of trades only, establishing the framework of integer-valued Lévy processes as a fundamental starting point for models of low latency data. This can be thought of as a first step towards more realistic stochastic process modelling, which in particular would involve time-change to allow for volatility clustering, statistical leverage and diurnal features. We also need a device to induce some serial dependence in short lags to allow for microstructure dependence. Finally, the models are invariant to the level of prices, which is unrealistic over long periods of time<sup>6</sup>.

In this work high quality, low latency tick price data from futures exchanges were used. With this we demonstrated that the CP-Skellam process (a compound Poisson process with Skellam innovations) provides a good fit to the unconditional distribution of mid-price changes on ‘normal’ times. Further we exhibit how unconditional price change distributions are affected by large economic events such as the release of US non-farm payroll numbers. On those days the quality of CP-Skellam fits tends to be poor. We also used the data to illustrate differences between the empirical, unconditional price change distributions of different futures markets showing, for example, log-linear tails for Crude Oil futures price changes. Such markets also pose fitting problems for the simple Skellam distribution.

---

<sup>6</sup>The latter point is easy to fix by allowing the intensity process of price changes to vary with the price level.

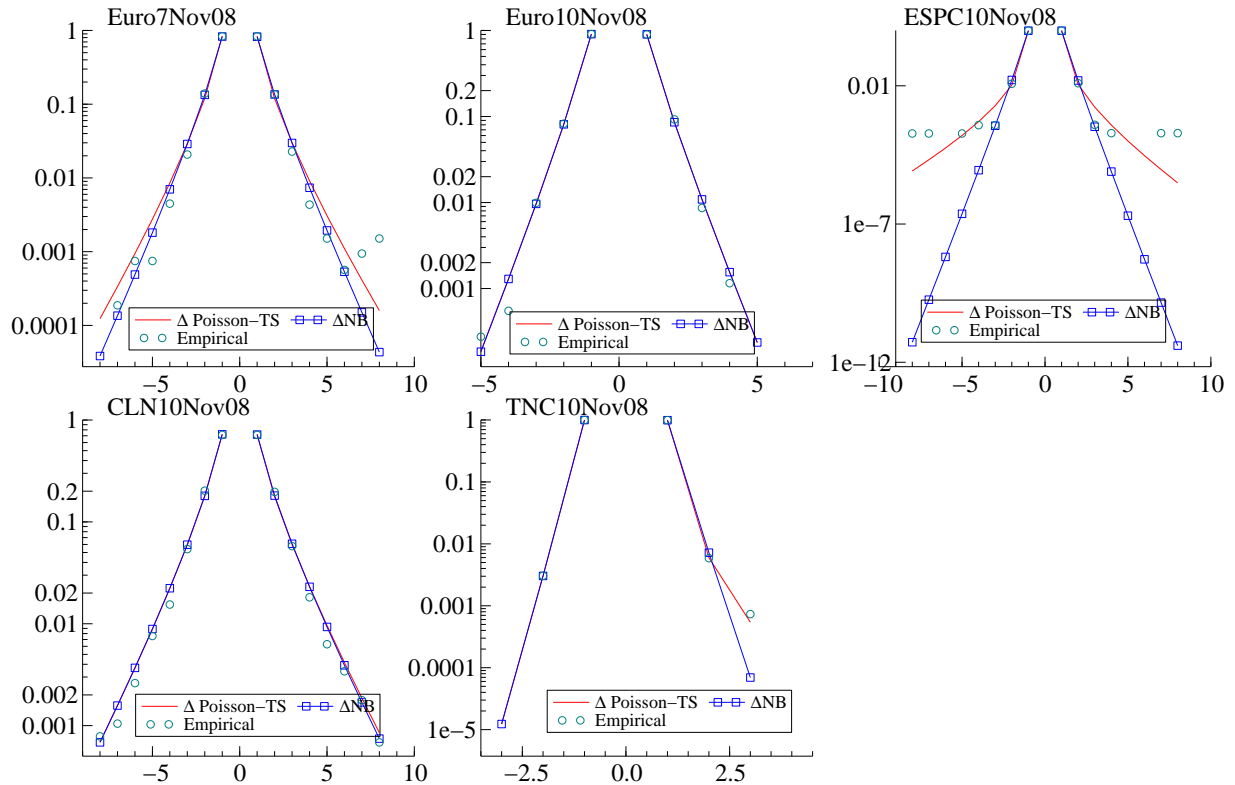


Figure 8: *Fits of the  $\Delta NB$  and  $\Delta Poisson-TS$  processes, together with a nonparametric fit. The results are given on the log scale.*

We have addressed the cases where simple Skellam fitting proves inadequate. Our mathematical theory has developed alternative distributions to the Skellam. Notable amongst these is the  $\Delta NB$  distribution for which we have derived a tractable distribution law. This distribution is more flexible and consequently more able to model the pure mid-price innovations for futures prices. The resulting process is nested inside the  $\Delta PTS$ .

We should mention the following. The simple binomial model of Cox et al. (1979) is related to the simple Skellam process in continuous time. Over a very small amount of time, in the Skellam process with probability one the price either stays the same, goes up one tick or goes down one tick. Hence the model is closest to a continuous time trinomial tree, discussed by for example Hull and White (1996) and Boyle (1986). A recent paper on this subject is, for example, Yuen and Yang (2010). Related mathematical finance work is carried out by Kirch and Runggaldier (2004) who look at modelling derivative prices based upon Poisson processes.

Finally, the basic building blocks developed here can be extended to allow for volatility clustering and statistical leverage using time-changes, while it would be also attractive to allow for limited amounts of autocorrelation to deal with the remaining microstructure noise in the pure mid-price.

The former problem is modelled in a follow up paper to this one in Holger Fink's 2011 Ph.D.

thesis, which adapts the Barndorff-Nielsen and Shephard (2001a) stochastic volatility model to the integer case. This new model has prices  $P_t = R_t^+ - R_t^-$ , where

$$R_t^+ = L_{T_t + \alpha \psi t E(Z_1)}^+, \quad R_t^- = L_{T_t + \psi Z_t}^-.$$

Here  $L_t^+$  and  $L_t^-$  are independent integer-valued subordinators, independent of  $Z$  which is a subordinator and  $T$  which is a general time-change. Note that it is not possible to write  $P$  as being a univariate Lévy process which is time-changed by a single process nor is it possible to write it as a stochastic volatility process  $\int_0^t \sigma_u dL_u$  (e.g. Renault and Werker (2009)). The class of time-changes explicitly studied puts

$$T_t = a_t + \int_0^t k(t, s) dZ_s.$$

Here the non-stochastic kernel  $k(t, s)$  is assumed to be bound and non-negative for any  $t, s \geq 0$ , while the non-stochastic  $a_t \geq 0$ .

Since submission of this paper in the autumn of 2010 we have been made aware of Bacry et al. (2011) which studies the modelling of low latency data using the difference of two Hawkes processes. This may be particularly attractive for one tick markets, but could equally be extended to allow larger moves by replacing the underlying Poisson process of the Hawkes process with the integer subordinators developed in this paper.

## 7 Acknowledgements

We are grateful to the low latency data providers QuantHouse ([www.quanthouse.com](http://www.quanthouse.com)) for allowing us to use their data in this paper. We also would like to acknowledge useful and informative discussions with Richard Adams, Giuliana Bordigoni, Holger Fink, Tom Kelly, Anthony Ledford, Andrew Patton and Kevin Sheppard. The comments of three referees and the editor Frederi Viens were also helpful.

## A Appendix

### A.1 Proof of Proposition 1

**Proof** Clearly  $L$  has no Gaussian component and so its Lévy-Ito representation has the form

$$L_t = at + \int_0^t \int_{|x| \geq \varepsilon} x N(dx ds) + \int_0^t \int_{|x| < \varepsilon} x (N(dx ds) - \nu(dx) ds) \quad (\text{A.1})$$

for any  $\varepsilon > 0$  and where  $N$  is a Poisson random measure with compensator  $E\{N(dx ds)\} = \nu(dx) ds$ . Since  $L$  is taking integer values only, by choosing  $\varepsilon < 1$  the last term in (A.1) disappears, and it

follows that  $a$  must be 0. Thus, in fact,

$$L_t = \int_0^t \int_{\mathbb{R}} x N(dx ds).$$

Furthermore, again since  $L$  is integer valued, for any  $i \in Z$  and  $t > 0$  we have that  $N((i-1, i) \times [0, t])$  is almost surely 0 and therefore

$$E\{N((i-1, i) \times [0, t])\} = \nu((i-1, i))t = 0$$

implying that  $\nu$  is concentrated on  $Z \setminus \{0\}$ . Recalling that any Lévy measure has the property

$$\int \min\{1, x^2\} \nu(dx) < \infty$$

we furthermore conclude that  $\nu$  is a finite measure.

The finiteness of  $\nu$  means that  $L$  is a finite activity process. Hence, splitting the jumps of  $L$  into positive and negative values, we can reexpress  $L$  as the difference  $L^+ - L^-$  between two discrete subordinators  $L^+$  and  $L^-$ .

## A.2 Proof of Theorem 1

**Proof.** The kumulant functions (log Laplace transforms) of  $L^+$  and  $T$  are related by

$$\bar{K}(\theta \dagger L_t^+) = \log E\{\exp(-\theta L_t^+)\} = t\bar{K}\left(\left(1 - e^{-\theta}\right) \dagger T_1\right).$$

Now

$$\bar{K}\left(\left(1 - e^{-\theta}\right) \dagger T_1\right) = \int_0^\infty \left\{e^{-(1-e^{-\theta})u} - 1\right\} \nu(du \dagger T_1).$$

Then

$$\begin{aligned} \bar{K}\left(\left(1 - e^{-\theta}\right) \dagger T_1\right) &= \int_0^\infty (e^{-u} - 1) \nu(du \dagger T_1) + \int_0^\infty (e^{e^{-\theta}u} - 1) e^{-u} \nu(du \dagger T_1) \\ &= \bar{K}\{1 \dagger T_1\} + \sum_{m=1}^\infty e^{-m\theta} \int_0^\infty \frac{u^m}{m!} e^{-u} \nu(du \dagger T_1) \\ &= \bar{K}\{1 \dagger T_1\} + \sum_{m=1}^\infty e^{-m\theta} q_m, \end{aligned}$$

where, crucially,

$$q_m = \int_0^\infty \frac{u^m}{m!} e^{-u} \nu(du \dagger T_1).$$

Using the fact that  $\bar{K}(0 \dagger L_t^+) = 0$  we obtain

$$\bar{K}(\theta \dagger L_t^+) = t \sum_{m=1}^\infty (e^{-m\theta} - 1) q_m.$$



Consequently, as is easily checked by direct calculation, the Lévy measure of  $L_t^+$  equals

$$\nu(dx \dagger L_t^+) = t \sum_1^\infty q_m \delta_m(dx)$$

where  $\delta_m$  denotes the delta measure at  $m$ . It follows in particular that

$$\nu((0, \infty) \dagger L_t^+) = t \int (1 - e^{-u}) \nu(du \dagger T_1) = -t\bar{K}(1 \dagger T_1) < \infty$$

and hence  $L^+$  is representable as a compound Poisson process of rate  $\kappa = \nu((0, \infty) \dagger L_1^+)$  and with innovation summands  $U_1, U_2, \dots$  having probability law

$$P(dx \dagger U) = \frac{\nu(dx \dagger L_1^+)}{\nu((0, \infty) \dagger L_1^+)}. \quad (\text{A.2})$$

In other words, the point probabilities of  $U$  are

$$\Pr(U = m) = \kappa^{-1} q_m.$$

This completes the proof.

### A.3 Proof of Theorem 2

We let  $U$  denote a random variable having the same law as the  $U_j$ .

Suppose first that  $Y$  is time change representable. Then, by Theorem 1, the point probabilities  $p_m$  must be of the form  $p_m = q_m$  with  $q_m$  of the form given in the statement of Theorem 1, and hence, for  $\phi \in [0, 1]$  and for some subordinator  $T$ ,

$$\begin{aligned} g(\phi) &= \kappa \int_0^\infty \sum_{j=1}^\infty \frac{(\phi u)^j}{j!} e^{-u} \nu(du \dagger T_1) \\ &= \kappa \int_0^\infty (e^{\phi u} - 1) e^{-u} \nu(du \dagger T_1) \\ &= \kappa \int_0^\infty (e^{-(1-\phi)u} - e^{-u}) \nu(du \dagger T_1) \\ &= \kappa \int_0^\infty (e^{-(1-\phi)u} - 1) \nu(du \dagger T_1) - \kappa \int_0^\infty (e^{-u} - 1) \nu(du \dagger T_1). \end{aligned}$$

Thus

$$g(1) = 1 = -\kappa \int_0^\infty (e^{-u} - 1) \nu(du \dagger T_1)$$

and

$$\begin{aligned} h(\phi) &= -\kappa \int_0^\infty (e^{-\phi u} - 1) \nu(du \dagger T_1) \\ &= -\int_0^\infty (e^{-\phi u} - 1) \nu(du \dagger S) \\ &= \bar{K}\{\phi \dagger S\} \end{aligned}$$

where  $S = T_\kappa$ .

On the other hand, if  $h(\phi) = \bar{K} \{\phi \dagger S\}$  on  $\phi \in [0, 1]$ , for a positive infinitely divisible random variable  $S$ , then, exactly retracing the steps in the above derivation, we find that  $U$  has moment generating function

$$g(\phi) = \int_0^\infty \sum_{j=1}^\infty \frac{(\phi u)^j}{j!} e^{-u} \nu(du \dagger S)$$

i.e.

$$p_m = \int_0^\infty \frac{u^m}{m!} e^{-u} \nu(du \dagger S).$$

But these point probabilities are precisely those in the compound Poisson representation of a time-changed Poisson process where the subordinator has law at time  $\kappa$  equal to that of  $S$ .

## References

- Abdulhamid, A. A. and M. Omair (2010). On the poisson difference distribution inference and applications. *Bulletin of the Malaysian Mathematical Sciences Society* 2, 17–45.
- Abramowitz, M. and I. A. Stegun (1970). *Handbook of Mathematical Functions*. New York: Dover Publications Inc.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association* 96, 42–55.
- Avellaneda, M. and S. Stoikov (2008). High frequency trading in a limit order book. *Quantitative Finance* 8, 217–224.
- Bacry, E., M. H. S Delattre, and J. F. Muzy (2011). Modelling microstructure noise with mutually exciting point processes. Unpublished paper.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 76, 1481–1536.
- Barndorff-Nielsen, O. E., S. Kinnebroeck, and N. Shephard (2009). Measuring downside risk: realised semivariance. In T. Bollerslev, J. Russell, and M. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, pp. 117–136. Oxford University Press.
- Barndorff-Nielsen, O. E. and N. Shephard (2001a). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, Series B* 63, 167–241.
- Barndorff-Nielsen, O. E. and N. Shephard (2001b). Normal modified stable processes. *Theory of Probability and Mathematical Statistics* 65, 1–19.
- Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* 64, 253–280.
- Barndorff-Nielsen, O. E. and N. Shephard (2006). Impact of jumps on returns and realised variances: econometric analysis of time-deformed Lévy processes. *Journal of Econometrics* 131, 217–252.
- Barndorff-Nielsen, O. E. and A. N. Shiryaev (2010). *Change of Time and Change of Measure*. Singapore: World Scientific.

- Bauwens, L. and N. Hautsch (2009). Modelling financial high frequency data using point processes. In T. G. Andersen, R. A. Davis, J. P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 953–979. Springer-Verlag.
- Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*. Heidelberg: Springer-Verlag.
- Boyle, P. P. (1986). Option valuation using a three jump process. *International Options Journal* 3, 7–12.
- Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Advances in Applied Probability* 31, 929–953.
- Carr, P., H. Geman, D. B. Madan, and M. Yor (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business* 75, 305–332.
- Chung, S.-L. and M. Shackleton (2002). The binomial Black-Scholes model and the greeks. *The Journal of Futures Markets* 22, 143–153.
- Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. London: Chapman and Hall.
- Copeland, T. E. and D. Galai (1983). Information effects in the bid-ask spread. *Journal of Finance* 38, 1457–1469.
- Cox, D. R. and P. A. W. Lewis (1966). *The Statistical Analysis of Series of Events*. London: Chapman & Hall.
- Cox, J. C., S. A. Ross, and M. Rubinstein (1979). Option pricing: a simplified approach. *Journal of Financial Economics* 7, 229–263.
- Delbaen, F. and W. Schachermeyer (2006). *The Mathematics of Arbitrage*. Springer.
- Devroye, L. (2009). Random variable generation for exponentially and polynomially tilted stable distributions. *ACM Transactions on Modeling and Computer Simulation* 19, 18:1–18:20.
- Engle, R. F. (2000). The econometrics of ultra-high frequency data. *Econometrica* 68, 1–22.
- Engle, R. F., D. F. Hendry, and J. F. Richard (1983). Exogeneity. *Econometrica* 51, 277–304.
- Field, J. and J. Large (2008). Pro-rata matching and one-tick markets. Unpublished paper: Oxford-Man Institute.
- Greenwood, M. and G. U. Yule (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society* 83, 255–279.
- Hansen, P. R. and G. Horel (2009). Quadratic variation by Markov chains. Unpublished paper: Department of Economics, Stanford University.
- Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise (with discussion). *Journal of Business and Economic Statistics* 24, 127–218.
- Hasbrouck, J. (1999). The dynamics of discrete bid and ask quotes. *Journal of Finance* 54, 2109–2142.
- Hausman, J., A. W. Lo, and A. C. MacKinlay (1992). An ordered probit analysis of transaction stock prices. *Journal of Financial Economics* 31, 319–30.
- Holla, M. (1967). On a Poisson-inverse Gaussian distribution. *Metrika* 11, 115–121.
- Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387–396.
- Hull, J. and A. White (1996). Using Hull-White interest rate trees. *Journal of Derivatives* 3(3), 26–36.

- Inusah, S. and T. J. Kozubowski (2006). A discrete analogue of the laplace distribution. *Journal of Statistical Planning and Inference* 136, 1090–1102.
- Irwin, J. O. (1937). The frequency distribution of the difference between two independent variates following the same Poisson distribution. *Journal of the Royal Statistical Society, Series A* 100, 415–416.
- Jørgensen, B. (1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society, Series B* 49, 127–162.
- Karlis, D. and E. Xekalaki (2005). Mixed poisson distributions. *International Statistical Review* 73, 35–58.
- Kirch, M. and W. Runggaldier (2004). Efficient hedging when asset prices follow a geometric Poisson process with unknown intensities. *SIAM Journal on Control and Optimization* 43, 1174–1195.
- Kozubowski, T. J. and S. Inusah (2006). A skew Laplace distribution on integers. *Annals of the Institute of Statistical Mathematics* 58, 555–571.
- Lehmann, B. (2008). Arbitrage-free limit order books and the pricing of order flow risk. NBER discussion paper.
- Lo, A. W. and J. Wang (2010). Stock market trading volume. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics: volume 2 — applications*, pp. 241–342.
- Madan, D. B. and E. Seneta (1984). Compound Poisson models for economic variable movements. *Sankhya: The Indian Journal of Statistics, Series B* 46, 174–187.
- Mykland, P. A. and L. Zhang (2011). The econometrics of high frequency data. In M. Kessler, A. Lindner, and M. Sørensen (Eds.), *Statistical Methods for Stochastic Differential Equations*. Chapman & Hall/CRC Press. Forthcoming.
- Pelsser, A. and T. Vorst (1994). The binomial model and the greeks. *Journal of Derivatives* 1, 45–49.
- Phillips, P. C. B. and J. Yu (2008). Information loss in volatility measurement with flat price trading. Unpublished paper: Cowles Foundation for Research in Economics, Yale University.
- Potters, M. and J.-P. Bouchaud (2003). More statistical properities of order books and price impact. *Physics A: Stat. Mech. Appl.* 324, 133–140.
- Press, S. J. (1967). A compound events model for security prices. *Journal of Business* 40, 317–335.
- Renault, E. and N. Touzi (1996). Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance* 6, 279–302.
- Renault, E. and B. Werker (2009). Stochastic volatility models with transaction time risk. Department of Economics, University of North Carolina.
- Rosenblau, M. (2009). Integrated volatility and round-off error. *Bernoulli* 15, 687–720.
- Rosinski, J. (1991). On a class of infinitely divisible processes represented as mixtures of Gaussian processes. In S. Cambanis, G. Samorodnitsky, and M. S. Taqqu (Eds.), *Stable Processes and Related Topics*, pp. 27–41. Basel: Birkhäuser.
- Rosinski, J. (2007). Tempered stable processes. *Stochastic Processes and Their Applications* 117, 677–707.
- Russell, J. R. and R. F. Engle (2006). A discrete-state continuous-time model of financial transaction prices and times. *Journal of Business and Economic Statistics* 23, 166–180.
- Russell, J. R. and R. F. Engle (2010). Analysis of high-frequency data. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics: volume 1 — tools and techniques*, pp. 383–426.
- Rydberg, T. H. and N. Shephard (2003). Dynamics of trade-by-trade price movements: decomposition and models. *Journal of Financial Econometrics* 1, 2–25.

- Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.
- Steutel, F. W. and K. Van Harn (2004). *Infinitely Divisibility of Probability Distributions on the Real Line*. Marcel Dekker.
- Tweedie, M. (1984). An index which distinguishes between some important exponential families. In J. Ghosh and J. Roy (Eds.), *Statistics: Applications and New Directions: Proc. Indian Statistical Institute Golden Jubilee International Conference*, pp. 579–604.
- Weber, P. and B. Rosenow (2005). Order book approach to price impact. *Quantitative Finance* 5, 357–364.
- Yuen, F. L. and H. Yang (2010). Option pricing with regime switching by trinomial tree method. *Journal of Computational and Applied Mathematics* 233, 1821–1833.
- Zeng, Y. (2003). A partially observed model for micromovement of asset prices with Bayes estimation via filtering. *Mathematical Finance* 13, 411–444.
- Zhu, R. and H. Joe (2009). Modelling heavy-tailed count data using a generalized Poisson-inverse Gaussian family. *Statistics and Probability Letters* 79, 1695–1703.