# An Improved Test for Continuous Local Martingales

**Article** *in* Communication in Statistics- Theory and Methods · May 2014

DOI: 10.1080/03610926.2013.788709



**2 authors:**

David A. Rolls
The Commonwealth Scientific and Industrial Research Organisation


Owen D Jones
Cardiff University

# An improved test for continuous local martingales

David A. Rolls [1,2]
Melbourne School of Psychological Sciences
University of Melbourne
Victoria 3010, AUSTRALIA
drolls@unimelb.edu.au

Owen D. Jones
Dept. of Mathematics & Statistics
University of Melbourne
Victoria 3010, AUSTRALIA
odjones@unimelb.edu.au

March 13, 2013

[1]corresponding author
[2]+61-3-8344-4768

**Abstract**

We present a new test for the "continuous martingale hypothesis". That is, a test for the hypothesis that observed data are from a process which is a continuous local martingale. The basis of the test is an embedded random walk at first passage times, obtained from the well-known representation of a continuous local martingale as a continuous time-change of Brownian motion. With a variety of simulated diffusion processes our new test shows higher power than existing tests using either the crossing tree or the quadratic variation, including the situation where non-negligible drift is present. The power of the test in the presence of jumps is also explored with a variety of simulated jump diffusion processes. The test is also applied to two sequences of high-frequency foreign exchange trade-by-trade data. In both cases the continuous martingale hypothesis is rejected at times less than hourly and we identify significant dependence in price movements at these small scales.

**Keywords:** continuous martingale hypothesis; foreign exchange; high-frequency data; hypothesis testing; time-changed Brownian motion

**Mathematics Subject Classification (2000):** 60G44, 62G10

# 1 Introduction

In this paper we develop two new tests for the "continuous martingale hypothesis" (Jones and Rolls, 2011). That is, given observations of some process $X$ at some times $t_0, \ldots, t_n$ we test the hypothesis that $X$ is a continuous local martingale. Examples of continuous local martingales used in finance (possibly after taking logs and removing drift) include stochastic volatility models (Heston, 1993; Hull and White, 1987). We demonstrate the higher power of these new tests on simulated datasets using standard alternatives and jump-diffusion processes. An application to foreign exchange (FX) data is also described.

It is known that if $X$ is a continuous local martingale with $X(0) = 0$ and quadratic variation $\langle X, X \rangle$ such that $\langle X, X \rangle_t \overset{a.s.}{\to} \infty$ as $t \to \infty$ then $X = B \circ \theta$ where $\theta = \langle X, X \rangle$ (Dambis, 1965; Dubins and Schwarz, 1965) (See also Revuz and Yor (1999, Theorem V.1.6), and Revuz and Yor (1999, Theorem V.1.7) to relax the requirement that $\langle X, X \rangle_t \overset{a.s.}{\to} \infty$ as $t \to \infty$.) Conversely, it is also known that if $B$ is a Brownian motion and $\theta$ is a continuous, non-decreasing process defined on the same filtration, $X = B \circ \theta$ is a continuous local martingale with quadratic variation $\theta$ (Revuz and Yor, 1999, Theorem V.1.5). In what follows we call $\theta$ a chronometer, meaning a non-decreasing process, possibly dependent on the past but not on the future.

A natural approach to testing the continuous martingale hypothesis has been to estimate $\theta$ using the realised volatility (McAleer and Medeiros, 2008) and then test to see if $\hat{B} = X \circ \hat{\theta}^{-1}$ is a Brownian motion. In particular, one can test if the increments of $\hat{B}$ are independent and identically distributed (i.i.d.) with a Gaussian distribution (e.g., Andersen et al. (2003); Vasudev (2007); Peters and de Vilder (2006); Andersen et al. (2007); Rolls and Jones (2011)), although other statistical tests have also been proposed (Ané and Geman, 2000; Guasoni, 2004). A fundamental problem with using realised volatility is that small scale estimation of $\theta$ is made extremely difficult by the presence of noise, and high-frequency financial data is known to exhibit market micro-structure noise due to effects such as bid-ask bounce, asynchronous trading and price discreteness (McAleer and Medeiros, 2008). A second problem is that an arbitrary parameter must be chosen which provides the length of the increments of $X \circ \hat{\theta}^{-1}$ which are to be tested. Rolls and Jones (2011) show with simulated data that the choice for this parameter alone could mean the difference between accepting and rejecting the continuous martingale hypothesis.

Jones and Rolls (2011) (see also Rolls and Jones (2011)) proposed a new test of the continuous martingale hypothesis using a different approach, focusing not on times but on spatial levels of the process. Their technique relies on a characterization of Brownian motion in terms of first passages (called "crossings") using a concept called the "crossing tree" (Jones and Shen, 2004). A key observation is that for a fixed realization of $B$, the sequence of first passages (but not the first passage times) of $B$ and $B \circ \theta$ are the same, effectively removing the effect of the chronometer. A natural application of a crossing-tree-based test and the tests described here is to tick data in which times are random but the ordinate (e.g., returns, prices, exchange rates) are known exactly. This differs from the quadratic-variation based tests, for which a natural context would have the observations regularly sampled in (calendar) time.

Our approach is similar to that of Rolls and Jones (2011) in that we use an embedded random walk, which inherits i.i.d. increments from $B$. A key difference is that we directly test the i.i.d. properties of the increments, using a binomial test for distribution and a runs test for independence. The tests in Rolls and Jones (2011) were effectively based on its scaling properties. (See Jones and Rolls (2011) for the supporting theory.) Our new approach offers several advantages. They are simpler and more intuitive than using the crossing tree, and do not rely on approximating distributions, especially continuous distributions with unknown parameters, or the asymptotic distribution of a test statistic. The binomial test is shown to be more powerful in the presence of drift (a clear weakness of the crossing tree tests).

The runs test is shown to be more powerful in the presence of mean reversion. Another advantage of the runs test is that it is truly a test that the increments are independent, not just uncorrelated. This is most relevant since the increments are dyadic, not Gaussian. We also show the two tests are independent, so a joint test with specified power is easily described. A joint test for the crossing tree tests has not been described, and may be illusive due to the complicated dependence between those tests. Further, we also illustrate the effect of jumps on the power of our tests in a manner that also provides practical insight to the role of data length.

Our key motivation for testing the continuous martingale hypothesis is to determine at which spatial scales, if any, financial (and other) data can be modelled as a continuously time-changed Brownian motion. So when we ask, “Is X a continuous local martingale?”, we are really asking, “At what spatial scale (if any) does X look like a continuous local martingale?” One application of this result is to fitting Multifractal Embedded Branching Processes (MEBP) (Decrouez, 2009; Decrouez and Jones, 2012). We imagine fitting a Brownian motion subjected to a continuous multifractal time-change $\theta$ (i.e., $X(t) = B(\theta(t))$). A first step is to test at which spatial scales the data can be modelled as a continuously time-changed Brownian motion.

# 2 Statistical Tests

## 2.1 Crossings and Embedded Random Walks

For a continuous process $\{X(t), t \geq 0\}$ with $X(0) = 0$, Jones and Shen (2004) introduced the crossing tree using crossing times (more precisely first passage times) and first passages given by $T_0 = 0$ and

$$T_j = \inf\left\{t > T_{j-1} : X(t) \in \delta\mathbb{Z},\, X(t) \neq X(T_{j-1})\right\},\, j = 1, 2, \ldots, \tag{2.1}$$

$$Y_j = X(T_j) \tag{2.2}$$

for a spatial scale $\delta$. (In fact, they use a range of spatial scales that produce nested sequences of first passages.) Jones and Rolls (2011) (see also Rolls and Jones (2011)) used these to develop tests for continuous local martingales. In this paper we again wish to use the sequence of first passages.

**Proposition 1.** *If $X$ is a continuous local martingale, then the discrete time process given by $Z_j = X(T_j)/\delta$ is a simple random walk.*

*Proof.* From Dambis (1965) and Dubins and Schwarz (1965), we can write $X$ as $B \circ \theta$ for some chronometer $\theta$. For our given $\delta$, let $\{\overline{T}_j\}_{j=0}^{\infty}$ be the crossing times of $B$. Then by the continuity of $B$ and $\theta$ we have that $\overline{T}_j = \theta(T_j)$ for all $j$, a.s. Thus $Z_j = B(\overline{T}_j)/\delta$, from which we immediately have that $Z$ is a sample random walk (see Knight (1981), for example.) $\square$

Unlike our earlier papers we will also estimate power for alternatives with jump discontinuities. Clearly, for a process $X$ with jumps, if a movement of size $\delta$ has been caused by a jump, then immediately following the jump the process will not necessarily lie on the $\delta$ lattice. This will necessitate the following generalisation of the definition of the crossing times of $X$. For any cadlag process $X$ with $X(0) = 0$, let $Y$ be $X$ observed on the $\delta$ lattice such that $T_0 = 0$, $Y_0 = 0$,

$$T_j = \inf\{t > T_{j-1} : |X(t) - Y_{j-1}| \geq \delta\}, \tag{2.3}$$

$$Y_j = \begin{cases} \max\left\{(Y_{j-1}, X(T_j)] \cap \delta\mathbb{Z}\right\}, & X(T_j) > Y_{j-1} \\ \min\left\{[X(T_j), Y_{j-1}) \cap \delta\mathbb{Z}\right\}, & X(T_j) < Y_{j-1} \end{cases}, \text{ and} \tag{2.4}$$

$$J_j = (Y_j - Y_{j-1})/\delta. \tag{2.5}$$

where the sequence $\{J_j, j = 1, \ldots, N\}$ is the signed number of barriers of the $\delta$-lattice crossed between $T_{j-1}$ and $T_j$. Thus, if at time $T_j$ the process $X$ has just diffused across a point on the $\delta$ lattice, then $Y_j = X(T_j)$, but if $X$ has jumped across the lattice then $Y_j$ will be the closest value on the $\delta$ lattice, in the direction the jump came from. This is illustrated in Figure 1 where $(T_2, Y_2)$ and $(T_4, Y_4)$ are crossing points by diffusion and $(T_1, Y_1)$ and $(T_3, Y_3)$ are crossing points by jumps. For a continuous process $X$, the definitions in (2.3)–(2.5) reduce to (2.1) with $|J_j| = 1$ for all $j$, and so Proposition 1 applies equally to continuous local martingales with crossing times defined by (2.3).

====================================

Figure 1 approximately here

====================================

## 2.2 Crossings from Data

Suppose we have data $\{(t_0, x_0), \ldots, (t_n, x_n)\}$ where $t_0 = 0 < t_1 < \ldots < t_n$. Without loss we assume $x_0 = 0$. We imagine a continuous process $X$ such that $X(t_i) = x_i$, $i = 0, \ldots, n$. Our goal is to test (at spatial scale $\delta$) the null hypothesis $H_0$: $X$ is a continuous local martingale against the alternative $H_a$: X is not a continuous local martingale. Note that if only the sequence $\{x_i, i = 0, \ldots, n\}$ are provided our conclusions are unaffected by using $t_i = i$, $i = 0, \ldots, n$.

Let $W$ be the cadlag function

$$W(t) = x_i,\, t \in [t_i, t_{i+1}), i = 0, \ldots, n-1.$$

Fixing $\delta > 0$, and using the definitions in (2.3)–(2.5) for $W$ let the crossing times be $\{\tau_j,\, j = 0, \ldots, \eta\}$, the crossing locations be $\{\omega_j,\, j = 0, \ldots, \eta\}$ and the signed jump sizes be $\{\upsilon_j,\, j = 1, \ldots, \eta\}$, where $\tau_0 = 0$, $\omega_0 = 0$ and $\eta$ is the index of the last observed crossing time. (Note we use Greek symbols to denote quantities from the empirical cadlag process. When necessary we write $\tau_j^\delta$, $\eta^\delta$ and $\upsilon_j^\delta$ to distinguish the quantities are at scale $\delta$.)

Define $S_0 = 0$, $S_k = \sum_{j=1}^k |\upsilon_j|$, and the number $N_W = S_\eta$ of size $\delta$ transitions over the $\delta$-lattice. Construct the binary sequence $\{U_j, j = 1 \ldots, N_U\}$ to capture the signed transitions of $W$ over the $\delta$-lattice with

$$U_j = I_{[\omega_i > 0]} \text{ for } j \in \{S_{\upsilon_i - 1}, \ldots, S_{\upsilon_i}\}, i \in \{1, \ldots, \eta\}. \tag{2.6}$$

where $I_A$ is 1 on the set $A$ and 0 otherwise. Similarly the process $X$ has crossing times $\{T_j,\, j = 0, \ldots, N_X\}$ and an associated binary sequence $\{U_j^X, j = 1 \ldots, N_X\}$ that captures the signed transitions of $X$ over the $\delta$-lattice, although both the crossing times and crossings may be partly unknown.

For testing at scale $\delta$ we assume $N_U = N_X = N$, and $U_j = U_j^X$, $j = 1 \ldots, N$. In other words, we assume all crossing directions of $X$ in $[0, t_n]$ at scale $\delta$ can be inferred from the data. This precludes continuous paths of $X$ with so many unobserved scale $\delta$ crossings that the number of inferred crossings are negligible in comparison, for example.

Under the null hypothesis at scale $\delta$, $\{X(T_j)/\delta,\, j = 1, \ldots, N\}$ is a simple random walk by Proposition 1, and by continuity of $X$, $|X(T_j) - X(T_{j-1})| = \delta$ for $j = 1, \ldots, N$. Since all crossings can be inferred, testing $H_0$ is equivalent to testing that $\{U_j, j = 1 \ldots, N\}$ is an i.i.d. sequence, with a symmetric distribution on 0-1. We test the distribution of $\{U_j, j =$

$1, \ldots, N\}$ using a binomial test. For independence we discuss several known alternatives, and settle on the runs test. Both tests use an exact distribution for the test statistic for short sequences.

### 2.3 Binomial Test of Distribution

Since the increments $\{U_j, j = 1, \ldots, N\}$ are binary by construction, for a test of distribution we need only test $\mathbb{P}(U_j = 1) = 1/2$ for all $j = 1, \ldots, N$, which is equivalent to testing that the proportion of ones is 1/2. Let

$$S_N = \sum_{j=1}^{N} U_j$$

denote the number of up transitions. Under the continuous martingale hypothesis $\{U_j, j = 1, \ldots, N\}$ is an i.i.d. Bernoulli sequence and $S_N$ has a Binomial($N$,1/2) distribution so a binomial test with test statistic $S_N$ can be used . We reject the continuous martingale hypothesis at scale $\delta$ if $S_N$ is too small or too large. As we will show with simulated data in Section 3, in comparison with the crossing tree-based tests, this test is particularly powerful for Brownian motion with drift, for which the crossing tree technique notably lacked power. Strictly speaking there is a dependence component to this test in that the sum of $N$ Bernoulli random variables is binomial *under the assumption of independence*. We will test the binary sequence $\{U_j\}_{j=1}^{N}$ for independence as well, so this is not such a problem.

### 2.4 Runs Test for Independence

At scale $\delta$ we wish to test whether $\{U_j, j = 1, \ldots, N\}$ form an independent (not just uncorrelated) sequence. Since they are binary sequences, there are specialized tests available. For a general binary sequence $X_1, \ldots, X_{m+n}$, define a run to be a subsequence $X_{s+1} = X_{s+2} = \cdots = X_{s+r}$ such that $X_s \neq X_{s+1}$ when $s > 0$, and $X_{s+r} \neq X_{s+r+1}$ when $s + r + 1 < (m + n)$. The well-known runs test of Wald and Wolfowitz (1940) tests whether the number of runs

$R_{m,n}$ in a binary sequence with $m$ zeros and $n$ ones is statistically too small or too large. Too few runs indicates a tendency for values to alternate while too many indicates a tendency to cluster. In particular, jumps in the data will create clusters in $\{U_j, j = 1, \ldots, N\}$ and there will typically be too few runs.

Because of symmetry in the zeros and ones, a test of hypothesis takes the form $R_{m,n} < c_{\alpha,m,n}$ where $c_{\alpha,m,n}$ is the level $\alpha$ critical value for $m$ zeros and $n$ ones. For short sequences $c_{\alpha,m,n}$ can be computed exactly and for long sequences a test based on the Gaussian distribution can be used.

We have implemented a number of other tests (Dixon, 1940; O'Brien, 1976; O'Brien and Dyck, 1985; Larsen et al., 1973) and found by simulation that the runs test has more discriminatory power by far for a range of alternatives, including additional examples not reported here. Moreover, the runs test is the only test which consistently rejects in the situation where small $\delta$ creates interpolation artefacts in simulated data. It also has the advantage of being implemented in readily available software packages. As such, the runs test seems sufficient for our needs and results from other tests are not reported.

## 2.5 Overall (Joint) Test

At each scale $\delta$, for a given significance level it is desirable to have an overall test that combines the results of the binomial and runs tests. In fact, these two tests are independent and so a joint test of given significance is easily described. To show independence under the continuous martingale hypothesis define the indicator variables

$$B_{\alpha,n} = \begin{cases} 1, & \text{the binomial test for } U_1, \ldots, U_N \text{ rejects at level } \alpha \\ 0, & \text{otherwise} \end{cases}$$

for the binomial test, and analogously $R_{\alpha,n}$ for the runs test. Then we have the following result.

**Lemma 1.** *Under the continuous martingale hypothesis, the binomial and runs tests are independent:*

$$\mathbb{P}(R_{\alpha,n} = 1,\, B_{\alpha,n} = 1) = \mathbb{P}(R_{\alpha,n} = 1)\mathbb{P}(B_{\alpha,n} = 1).$$

The proof for this result is straightforward and omitted. For binomial and runs tests at significance levels $\alpha_B$ and $\alpha_R$ respectively, a joint test that rejects if either test rejects has significance given by $\alpha_J = \alpha_B + \alpha_R - \alpha_B\alpha_R$. For example, for $\alpha_B = \alpha_R = 0.05$ the overall significance level is given by $\alpha_J = 0.0975$.

# 3 Power Estimates

## 3.1 Diffusions

To estimate the power of the binomial and runs test against a variety of alternatives we apply them to several simulated diffusion processes. As a comparison, the quadratic volatility (QV) test as implemented in Rolls and Jones (2011) (similar to the test reported in Vasudev (2007) but most favourable to the choice of increment length) and the crossing tree (CT) tests of Rolls and Jones (2011) are also applied. (For Brownian motion we report on drift 1 only; see Jones and Rolls (2008) for comparable results with drift 1.5.) Writing $B$ for Brownian motion, the following test cases were considered.

i) Brownian motion with drift 1 and volatility 1.

$$dX_t = dt + dB_t$$

ii) Ornstein-Uhlenbeck process with drift $-8X_t$ and volatility 1. The Vasicek model (Vasicek, 1977) is an example of its use in finance.

$$dX_t = -8X_t dt + dB_t$$

iii) Feller's square root process with drift $6(0.2 - X_t)$ and volatility $\sqrt{X_t}$, known as the Cox, Ingersoll, Ross (CIR) model in finance (Cox et al., 1985).

$$dX_t = 6(0.2 - X_t)dt + \sqrt{X_t} dB_t$$

These test cases were originally chosen by Vasudev (2007) and were also used in Rolls and Jones (2011), so direct comparisons are possible. Vasudev (2007) imagined data of length 1250 in the time interval [0,5] (i.e., 5 years) and length 5000 in the time interval [0,20] (i.e., 20 years). As in Rolls and Jones (2011), to allow the most direct comparisons, we use a base scale $\delta$ to achieve either an *average* of 1250 crossings in time [0,5] or 5000 crossings in time [0,20]. (See Rolls and Jones (2011) for details on determining $\delta$ to do this.) Thus we use different simulated data for the QV test than for the CT and RW tests, but chosen so that we have the same number of observations, and so that on average the processes are observed over the same time period. Results for 5000 crossings were similar to those for 1250 crossings and omitted for brevity. Similarly, simulation results for Brownian motion with drift 1.5, Ornstein-Uhlenbeck processes with drift $-10X_t$ and Feller's square root process with drift $8(0.2 - X_t)$ are also similar and omitted for brevity.

As noted in the introduction, in practice we are generally interested in the scale at which a particular model (in our case a continuous local martingale) is appropriate. Accordingly, when applying the RW test, a range of spatial scales $\delta$ should be considered. Given crossing points at scale $\delta$, it is natural to then consider scales $2^l\delta$, as the crossings at these scales are *nested* (in the sense of Barlow and Perkins (1988)). That is, all the crossing points at scale

$2^l\delta$ are also crossing points at scale $2^{l-1}\delta$. Given $\delta_0$, following Rolls and Jones (2011) we will refer to a test at scale $2^l\delta_0$ as a test at level $l$. Results reported here use base scales of $\delta_0 = 0.063288$, $\delta_0 = 0.063015$, and $\delta_0 = 0.028163$ for cases (i), (ii) and (iii) respectively. Note that, unlike the crossing tree tests, tests based on the embedded random walk properties at different levels are dependent. Also note, as with the crossing tree tests, each one level increase reduces the number of crossings/transitions by approximately a factor of 4 and power decreases to zero at the largest levels.

Table 1 gives power estimates for each of the cases considered, obtained as the number of paths rejected out of 10,000 simulations in each case. For the RW test the power for both the binomial test and the runs test is shown at a range of levels for the crossings/transitions. Results for the finest resolution are reported starting at level 0. In contrast, results from the crossing tree tests are normally reported in terms of *subcrossing* levels, which start at level 1, but are based on crossings at level 0. We claim this is one way our new test is more intuitive. For the crossing tree tests, Rolls and Jones (2011) described results for 8 different tests across four levels. Here, only the most favourable of the 32 results is reported for each of the three cases. We make no claims about joint significance when all eight tests are used this way. The QV test requires the user to specify the size of the increments of $\hat{B} = X \circ \hat{\theta}^{-1}$ that are tested for normality, and it is sensitive to this choice. Following Rolls and Jones (2011) (see also Jones and Rolls (2008)) we considered a range of possible increment sizes and report on the choice that gave the most power. Clearly this is not possible in practice, but puts the QV test in the best light for comparison with the other tests.

================================

Table 1 approximately here

================================

From the results in Table 1, the binomial test is good at detecting drift. In fact it has power comparable with the most favourable QV test. On the other hand it shows negligible power for the two stationary, mean-reverting alternatives. This makes sense as we would expect the number of up and down crossings to be roughly the same, and so $S_N$ should be close to its expected value of 1/2. On the other hand, the runs test has very little power for Brownian motion with drift 1, which is unsurprising since the crossings are independent. But it does much better with the mean reverting alternatives than any of the CT or QV tests. The further these processes are from their mean the more likely consecutive transitions will move the process back towards the mean. These will look like clustering in the sequence $\{U_j\}$.

## 3.2 Jump Diffusions

In this paper we also report on the effect of jumps on the power of our tests. The introduction of jumps changes things considerably. In particular, jumps across the $\delta$-lattice can also generate crossings. If a jump generates several crossings that will show as a run in the binary sequence $\{U_j, j = 1, \ldots, N\}$. More generally, following a jump that creates the $n$-th crossing, the process need not be on the lattice $\delta\mathbb{Z} + X_0$ (i.e. an overshoot). The overshoot will mean the transition probability to crossing $n+1$ will depend on the *location* of crossing $n$ in relation to the lattice.

To explore the effect of jumps on our tests we consider the doubly-exponential jump diffusion process (DEP) of Kou (2002):

$$X_t = X_0 + \mu t + \sigma B_t + \sum_{i=1}^{N_t} Y_i,\ t \geq 0 \tag{3.1}$$

for constants $\mu$ and $\sigma > 0$, Brownian motion $B_t$, rate $\lambda$ Poisson process $N_t$ and i.i.d. jump sizes $\{Y_i,\ i = 1, 2, \ldots\}$. In the doubly-exponential model the jump sizes have probability density function

$$f_Y(x) = p\eta e^{-\eta x} I_{[x\geq 0]} + (1-p)\theta e^{-\theta x} I_{[x<0]}$$

where $0 < p < 1$, $\eta > 1$, and $I_A$ is 1 on the set $A$ and 0 otherwise.

An attractive feature of the DEP model is that conditional on there being a jump up (or down) beyond upper and/or lower barriers, the distribution of the overshoot beyond the barrier is exponential by the memoryless property. Further, conditional on the overshoot being greater than 0, the first passage time and the overshoot size are independent. In particular, for a jump over the lattice $\delta\mathbb{Z}+X_0$ that creates crossings, the number of crossings will be geometrically distributed. Extension to hyperexponential jump sizes (the HEP model) has also been proposed (Cai, 2009; Cai et al., 2009). Thus, the DEP (and the HEP) are well-suited to considering first passage times analytically (Kou and Wang, 2003; Cai, 2009) and pricing single and double barrier options (Cai et al., 2009).

Most relevant here are results of Cai et al. (2009) for computing the Laplace transform $\mathbb{E}^0[\exp(\alpha\tau)f(X_\tau)]$, where $f$ is any non-negative measurable function, $\tau$ is the first passage time to flat barriers and $\mathbb{E}^x$ is the expectation when process $\{X_t\}$ starts at $X_0 = x$. In particular, this allows computation of the probabilities of up and down crossings whether by diffusion or by jump, for any $x$ in $[(n-1)\delta, n\delta]$ for any $n \in \mathbb{Z}$, and direct simulation of sequences of crossings. To simplify bookkeeping, denote the crossing as type 1 (down by diffusion), type 2 (up by diffusion), type 3 (down by jump) and type 4 (up by jump) (i.e., even types for up crossings; larger types for crossings by jump). For $x \in (-\delta, \delta)$ the probabilities the next scale-$\delta$ crossing is in the up direction by diffusion and by a jump are

$$p_2(x) = \lim_{\alpha\searrow 0} \mathbb{E}^x[\exp(-\alpha\tau)I_{[X_\tau=\delta]}] \text{ and} \tag{3.2}$$

$$p_4(x) = \lim_{\alpha\searrow 0} \mathbb{E}^x[\exp(-\alpha\tau)I_{[X_\tau>\delta]}] \tag{3.3}$$

respectively. The probabilities $p_1(x)$ and $p_3(x)$ are found similarly. It follows that $\sum_{i=1}^4 p_i(x) = 1$ for all $x \in (-\delta, \delta)$. The time to the next crossing has Laplace transform $\mathbb{E}^x[\exp(-\alpha\tau)]$ from which moments can be computed by differentiating at $\alpha = 0$.

By routine calculations, the first and second moments of the jump sizes are $\mathbb{E}[Y_i] = p/\eta - q/\theta$ and $\mathbb{E}[Y_i^2] = p/\eta^2 + q/\theta^2$, the variance of the return at time $t$ is

$$\mathrm{Var}(X_t) = \sigma^2 t + \lambda t \mathbb{E}[Y_i^2]$$

and the contribution of the jump variability to return variability is

$$\frac{\lambda \mathbb{E}(Y_i^2)}{\sigma^2 + \lambda \mathbb{E}[Y_i^2]}.$$

To simulate the sequences of crossings from the DEP model, evaluation of the four transition probability expressions (e.g., (3.2) and (3.3) for the up direction) is needed to sufficient precision across the range of parameters used here. All simulations used the following fixed values: $\mu = 0$, $\sigma^2 = 2 \times 10^{-5}$ and $\delta_0 = 5.6 \times 10^{-5}$. These are roughly based on values found for the GBP-USD data considered below. In particular, for $t$ measured in days, if we imagine crossings of mean length 0.2 minutes (as reported in Rolls and Jones (2011)), then 100,000 crossings is about 2.4 simulated weeks at 138 hours per week (i.e., 6 AM Monday to 12 AM Saturday). Indeed, with these parameters the mean time between crossings in the simulations is about 0.2 minutes. No drift component has been included to consider the effects of jumps alone.

In calculating transition probabilities, numerical instability was found to be an issue for small values of $\sigma^2$. A consistently quick and reliable method was to evaluate the limit symbolically (which creates long, unwieldy formulas) and then evaluate for values of $X_0 = x_0$ in the range $(-\delta, \delta)$. Figure 2 (left panel) shows the up and down transition probabilities by diffusion (as lines) using this method. Parameter values were $\lambda = 1$, $\eta = 1000$, $p = 1/2$ and $\delta = 2^4 \delta_0$ (corresponding to case (i) below at level 4.) For comparison, estimated probabilities (with 95% confidence intervals) found by Monte Carlo simulation of 100,000 crossings are shown for a range of values of $x_0$. Each crossing was obtained by simulating the process in

(3.1) on a temporal grid with size $\Delta t = 10^{-6}\sqrt{2^4\delta_0}$, and other parameters as described above. Figure 2 (right panel) shows similar results for the up and down transition probabilities by jumps. While the agreement is quite good, Monte Carlo simulation to obtain the transition probabilities is time-prohibitive across a sufficient range of values for $x_0$ and the various other parameters considered here. For example, at level 0, even 100,000 crossings is not enough so that confidence intervals for transition probabilities by jumps do not include zero. From now, all simultions from the DEP model will simulate crossings directly using the four transition probability expressions (e.g., (3.2) and (3.3) for the up direction.)

===================================

Figure 2 approximately here

===================================

Unlike the results in Section 3.1 and in Rolls and Jones (2011), power estimates are not reported across a number of levels per se, because it combines an effect from explicitly varying a parameter with the effect of decreased data length (which falls by a factor of four per level for Brownian Motion). Instead these effects are separated by holding the number of crossings fixed at either $N = 10^2$, $N = 10^4$ or $N = 10^6$. Particularly for the latter, these datasets are orders of magnitude larger than for results previously reported. In practice, if one had $10^6$ crossings at level 0 they would have about 15 crossings at level 8.

Four cases are reported here.

i) Effect of increasing $\delta$. Fix $\lambda = 1$, $\eta = 1000$, $p = 1/2$, then take crossings of size $\delta = 2^l\delta_0$ for $l = 0, 1, \ldots, 8$ (i.e., levels $0 - 8$).

ii) Effect of increasing $\lambda$. Fix $\eta = 1000$, $p = 1/2$, $\delta = \delta_0$ , and take $\lambda = 0, 0.5, 1, 2, 5, 10, 20$.

iii) Effect of increasing $p$. Fix $\lambda = 1$, $\eta = 1000$, $\delta = \delta_0$ , and take $p = 0.5, 0.6, \ldots, 0.9$.

iv) Effect of increasing mean jump size $(1/\eta)$. Fix $\lambda = 1$, $p = 1/2$, $\delta = \delta_0$ , and $\eta = 10^6, 10^5, \ldots, 10^1$.

For each case, 1000 simulations were performed, and the percentage of tests that were significant at the 5% level (with 95% confidence intervals) is reported.

Figure 3 (left panel) shows the effect of increasing $\delta$ on runs test power. Jump contribution to variability is constant at 9.1% since the parameters of the process do not depend on $\delta$. As expected, power increases with longer datasets. For 100 crossings the test does not reject beyond its significance level, while for $10^6$ crossings the test has almost 100% power at most levels. In between are the datasets with $10^4$ crossings. The shape of that curve exactly mirrors the curve of level $l$ versus the average number of runs, and is an interesting consequence of competing effects. (The number of up and down crossings are roughly equal due to symmetry.) For smallest $l$, jumps are infrequent but create large runs. As $l$ increases there are more jumps between crossings but jumps are less likely to cause a run or even a crossing ($n = 0$: about 1.5 jumps per path, mean jump size $17\delta$; $n = 4$: 388 jumps per path, mean jump size $0.08\delta$; $n = 8$: 94,000 jumps per path, mean jump size $< 0.0003\delta$.) Due to symmetry in the data, the binomial test offers no useful information.

==================================

Figure 3 approximately here

==================================

Figure 3 (right panel) shows the effect of increasing jump arrival rate $\lambda$. Jump contribution to variability increases from 0% at $\lambda = 0$ to 9.1% at $\lambda = 1$ to 66.7% at $\lambda = 20$. As expected, more paths are rejected with increasing data size $N$. Due to symmetry in the data, the binomial test offers no useful information.

Figure 4 shows the effect of increasing $p$, which controls the ratio of up to down jumps. Since the mean up and down jump sizes are equal, $p \neq 1/2$ breaks the symmetry in jump direction but not jump size and jump contribution to variability is constant at 9.1%. Figure 4 (left panel) shows results for the runs test, which rejects about 75% at $N = 10^6$ crossings, but does not reject below the level of the test for smaller $N$. Figure 4 (right panel) shows results for the binomial test. Again, only $N = 10^6$ crossings shows a considerable fraction of paths rejected. As expected, as $p$ becomes more extreme the data is increasingly less symmetric due to the bias for up direction jumps, and binomial test power increases.

=====================================

Figure 4 approximately here

=====================================

Figure 5 shows the effect of increasing mean up and down jump sizes on the runs test (left panel) and the binomial test (right panel) by increasing $\mu_{jump} = 1/\eta = 1/\theta$ from $10^{-6}$ to 0.1, or roughly from $0.18\delta_0$ to $1786\delta_0$. Jump contribution to variability increases from $10^{-5}\%$ to 99.9%. The runs test rejects a considerable number of paths when the jump contribution to variability is at least 91% ($N = 10^4$) and 9.1% ($N = 10^6$), while the binomial test shows less power by rejecting a considerable number of paths when the jump contribution to variability is at least 91% ($N = 10^4$ and $N = 10^6$). The increased power in the binomial test may be surprising, but it can be explained by noting how many crossings are created by a single jump for those values of $\eta$ and $\theta$. This makes the test sensitive to even a single jump. An investigation into the value of the test statistic shows across the 1000 sample paths it is much more variable than under the null hypothesis, on account of this effect. Comparing this to case (ii), the test is better at detecting many small jumps than a few large jumps, given that the proportion of variability due to jumps is fixed.

=====================================

Figure 5 approximately here

=====================================

Note that cases (ii), (iii) and (iv) only consider level 0 crossings. In line with case (i) we would expect the power to increase markedly at higher levels. For detecting jumps the runs test performed uniformly better than the binomial test, though its performance was sensitive to the crossing size.

## 4 Analysis of Foreign Exchange Rates

We applied our RW tests to two sequences of high frequency foreign exchange rate trade-by-trade data obtained from the Securities Industry Research Centre of Asia-Pacific (SIRCA): Australian Dollar to US Dollar (AUD-USD) and UK Pound to US Dollar (GBP-USD). Each series represents a different exchange rate, for the period January to December 2003. This is a subset of the sequences used in Rolls and Jones (2011); some summary statistics for these sequences are reported there. As is common for considering financial data, we work with log-transformed data.

Table 2 shows the results and summary information for the AUD-USD data. Results are reported at spatial scales $\delta = 2^l \delta_0$ for levels $l = 0, 1, \ldots, 6$ and $\delta_0 = 0.000155569$. Test results are shown as $p$-values. A result indicating a test rejects the continuous martingale hypothesis at a level of 5% is also shown in boldface. A significance level of 5% was used for all of them, with the understanding that the overall test, rejecting by either the binomial or runs test, would have 0.0975% significance. For this data the binomial test rejects at all levels. More will be said about this below. The runs test also rejects at levels 0 to 2 indicating the number of sequences of up or down transitions is not typical of a random walk.

==================================

Table 2 approximately here

==================================

The last three rows of the table provide summary information about the embedded random walks. The fourth and fifth rows give the number of up and total transitions, respectively. For example, at level 0 there are 320,681 transitions in the embedded random walk, and about 50.29% of those are up transitions. The last row shows an estimate of the mean crossing length at that level to give a sense of the average timescale. They are found by dividing 6,396 hours per year for currency trading by the number of crossings in each level. Using these average times as a rough guide to the timescale equivalent to each spatial scale, there is evidence of micro-structure noise at timescales up to roughly 15 minutes.

To test whether drift provides the explanation for the consistent binomial test rejections, the log-transformed data was detrended and RW tests performed again. The detrending operation simply subtracted a linear component with slope given by the sample mean of the increments. No attempt was made to account for any non-linear component in the log-transformed data. After detrending, the embedded random walks were created again, using the same base scale $\delta$ used initially. After this transformation the binomial test did not reject at any level. In fact, the number of up transitions was exactly 1/2 the total at all levels. The runs test results remained unchanged, rejecting at the first three levels and never above. This represents complete agreement with the CT tests, which rejected the first three levels only, and is known to be less powerful in the presence of drift. Thus, a linear drift component is the likely cause of the rejections by the binomial tests and reveals a key advantage of the embedded random walk approach over the crossing tree for testing.

Table 3 shows the results for the GBP-USD data using $\delta_0 = 0.0000563285$. In this case the binomial test was not significant at all, but the runs test was significant at levels 0–4, at the 5% level. This could be due to jumps or dependence between crossing directions. To distinguish between the two we went back to the original data and looked for evidence of jumps. At level 3 crossings have size $8\delta_0$, and at this scale an absolute change in the log price of $16\delta$ or more can be considered a jump. This happened in 98 out of 34,288 crossings at this scale, or 0.286% of the time, and so had a negligible impact on the runs test statistic. At level 4 there were 24 out of 8,362 (0.287%) of the crossings with potential jumps, which again had a negligible impact. Thus we conclude that the runs test is not rejecting the null hypothesis because of jumps. Rather there is significant evidence of dependence in foreign exchange rate movements at scales up to 45 minutes. The same check performed with the

AUD-USD data gave similar results.

=====================================

Table 3 approximately here

=====================================

# References

Andersen, T., Bollerslev, T., Diebold, F., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71:579–625.

Andersen, T., Bollerslev, T., and Dobrev, D. (2007). No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics*, 138:125–180.

Ané, T. and Geman, H. (2000). Order flow, transaction clock, and normality of asset returns. *Journal of Finance*, 60:2259–2284.

Barlow, M. T. and Perkins, E. A. (1988). Brownian motion on the Sierpiński gasket. *Probab. Theory Related Fields*, 79(4):543–623.

Cai, N. (2009). On first passage times of a hyper-exponential jump diffusion process. *Oper. Res. Lett.*, 37(2):127–134.

Cai, N., Chen, N., and Wan, X. (2009). Pricing double-barrier options under a flexible jump diffusion model. *Oper. Res. Lett.*, 37(3):163–167.

Cox, J. C., Ingersoll, Jr., J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.

Dambis, K. È. (1965). On decomposition of continuous submartingales. *Teor. Verojatnost. i Primenen.*, 10:438–448.

Decrouez, G. (2009). *Generation of multifractal signals with underlying branching structure*. PhD thesis, University of Melbourne.

Decrouez, G. and Jones, O. D. (2012). A class of multifractal processes constructed using an embedded branching process. *Journal of Applied Probability*. (To Appear).

Dixon, W. J. (1940). A criterion for testing the hypothesis that two samples are from the same population. *Ann. Math. Statistics*, 11:199–204.

Dubins, L. and Schwarz, G. (1965). On continuous martingales. *Proc. Nat. Acad. Sci. USA*, 53:913–916.

Guasoni, P. (2004). Excursions in the martingale hypothesis. In *Stochastic Processes and Applications to Mathematical Finance*, pages 73–95. World Sci. Publ., River Edge, NJ.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *J. Finance*, 42:281–300.

Jones, O. and Shen, Y. (2004). Estimating the Hurst index of a self-similar process via the crossing tree. *Signal Processing Letters*, 11:416–419.

Jones, O. D. and Rolls, D. A. (2008). Looking for continuous local martingales with the crossing tree (working paper). arXiv: 0911.5204v2.

Jones, O. D. and Rolls, D. A. (2011). A characterisation of, and hypothesis test for, continuous local martingales. *Electronic Communications in Probability*, 638–651.

Knight, F. B. (1981). *Essentials of Brownian Motion and Diffusion*, volume 18 of *Mathematical Surveys*. American Mathematical Society, Providence, R.I.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Manage. Sci.*, 48:1086–1101.

Kou, S. G. and Wang, H. (2003). First passage times of a jump diffusion process. *Adv. in Appl. Probab.*, 35(2):504–531.

Larsen, R. J., Holmes, C. L., and Heath Jr., C. W. (1973). A statistical test for measuring unimodal clustering: a description of the test and of its application to cases of acute leukemia in metropolitan Atlanta, Georgia. *Biometrics*, 29:301–309.

McAleer, M. and Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, 27:10–45.

O'Brien, P. C. (1976). A test for randomness. *Biometrics*, 32:391–401.

O'Brien, P. C. and Dyck, P. J. (1985). A runs test based on run lengths. *Biometrics*, 41:237–244.

Peters, R. T. and de Vilder, R. G. (2006). Testing the continuous semimartingale hypothesis for the S&P 500. *J. Bus. Econom. Statist.*, 24(4):444–454.

Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition.

Rolls, D. A. and Jones, O. D. (2011). Testing for continuous local martingales using the crossing tree. *Australia & New Zealand Journal of Statistics*, 53(1):79–107.

Vasicek, O. (1977). An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188.

Vasudev, R. (2007). *Essays on Time Series: Time change and applications to testing, estimation and inference in continuous time models*. PhD thesis, Rice University.

Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population. *Ann. Math. Statistics*, 11:147–162.

| | QV test (most favourable increment length) | CT tests (most favourable of 8 tests and 4 levels) | binomial test (levels) | | | | runs test (levels) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 3 |
| BM with drift 1 | 60 | 10.5 | 59 | 59 | 50 | 31 | 5 | 4 | 3 | 1 |
| Ornstein-Uhlenbeck | <1 | 18.5 | 0 | 0 | 0 | 0 | 7 | 15 | 36 | 43 |
| Feller's square root | < 1 | 14.5 | 0 | 0 | 0 | 0 | 9 | 17 | 31 | 32 |

Table 1: Power estimates. Percentage of 10,000 simulations rejected by tests at 5% significance level. Quadratic variation (QV) results use the most favourable increment length. Crossing tree (CT) results reported as most favourable of 32 results (eight different tests at each of four levels).

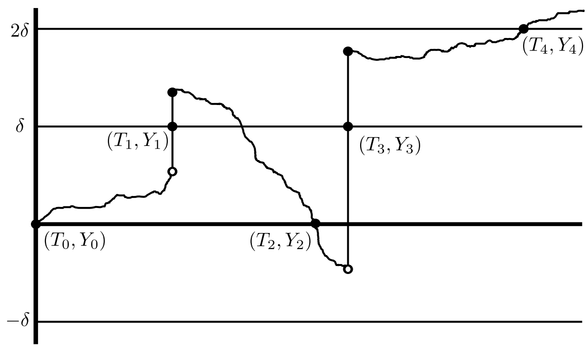


Figure 1: Illustration of crossings by diffusion and by jump for a process $X$. For crossings by diffusion across the $\delta$ lattice, the crossing point is on the path and $Y_i = X(T_i)$. For a crossing $i$ due to a jump, $Y_i$ is the closest value on the $\delta$ lattice in the direction the jump came from.

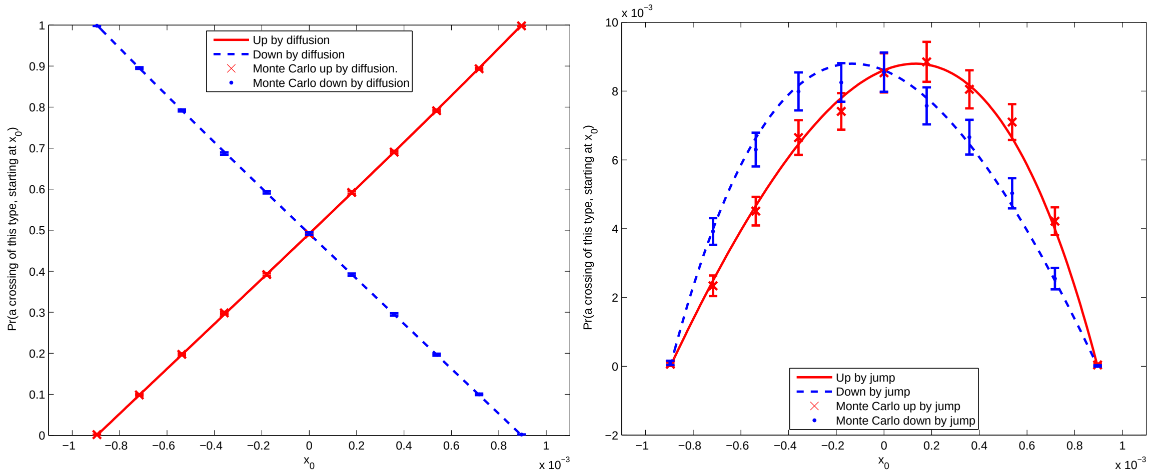


Figure 2: Probabilities of up and down crossings by diffusion (left panel) and jump (right panel) from $X(0) = x_0$. Solid lines show values found using Laplace transforms such as (3.2) or (3.3). Vertical bars show estimated values (with 95% confidence intervals) from simulating 100,000 crossings for the jump diffusion starting at $x_0$.

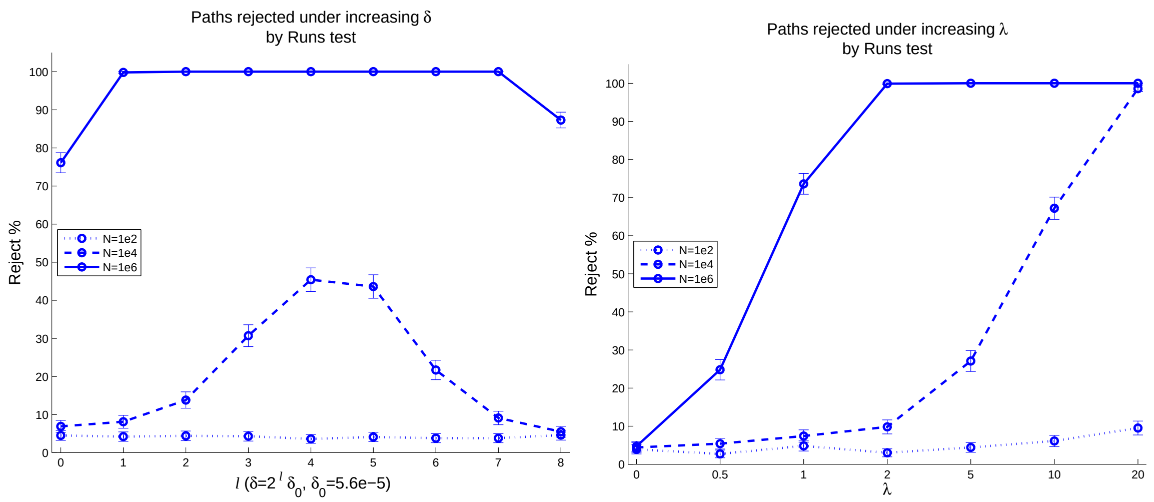


Figure 3: Effect of increasing spatial scale ($\delta$) (left panel) and jump arrival rate ($\lambda$) (right panel) on runs test power. Percentage of 1,000 sample paths rejected for $N$ crossings with approximate 95% confidence intervals. For both cases, due to symmetry in the data the binomial test (not shown) has no power.

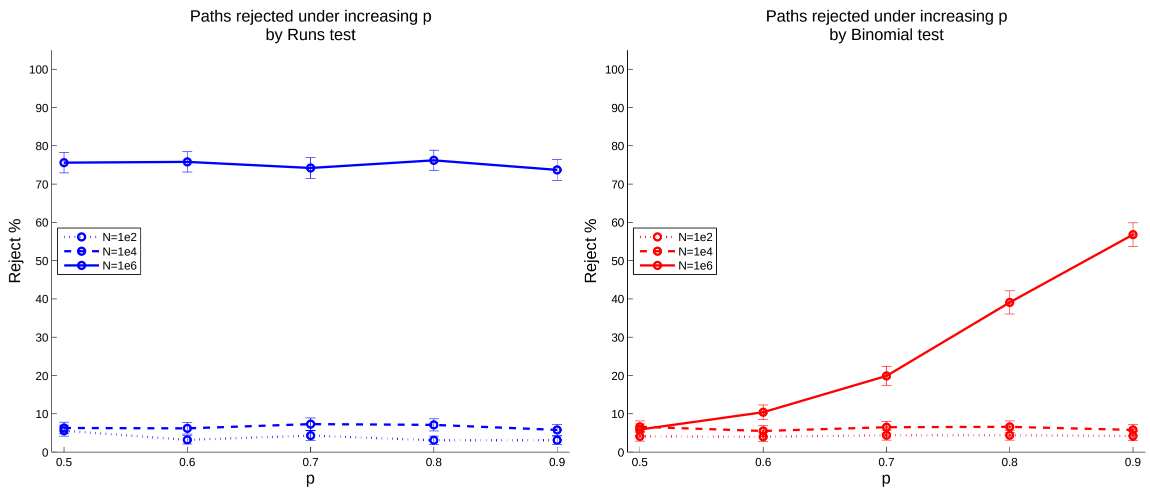


Figure 4: Effect of increasing jump asymmetry ($p$) on test power for the runs test (left panel) and binomial test (right panel). Percentage of 1,000 sample paths rejected for $N$ crossings with approximate 95% confidence intervals.

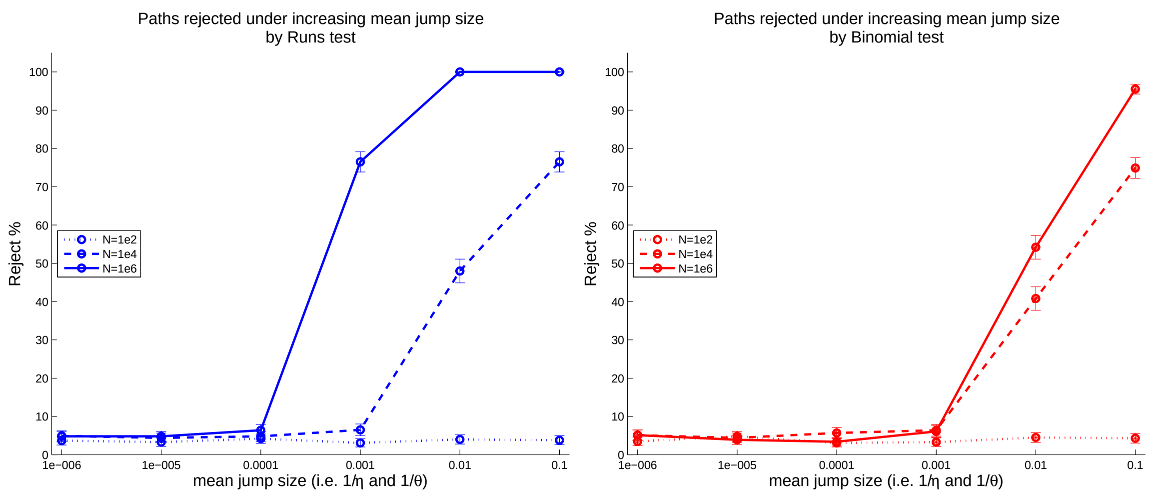


Figure 5: Effect of increasing mean jump size ($1/\eta$) on test power for the runs test (left panel) and binomial test (right panel). Percentage of 1,000 sample paths rejected for $N$ crossings with approximate 95% confidence intervals.

| xing levels ($l$) | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| **binomial** | **0.001** | **0.003** | **0.004** | **0.004** | **0.005** | **0.006** | **0.008** |
| **runs** | **0.000** | **0.000** | **0.002** | 0.120 | 0.883 | 0.886 | 0.343 |
| **# up** | 161275 | 49981 | 13392 | 3451 | 909 | 242 | 71 |
| **# up or dn** | 320681 | 99027 | 26316 | 6668 | 1701 | 426 | 113 |
| **mean xing length** | 1.2 min | 3.9 min | 14.6 min | 57 min | 3.8 h | 15.4 h | 63 h |

Table 2: AUD-USD test results. $p$-values for the RW tests applied to 2003 AUD-USD exchange rate trade-by-trade data. Level $l$ corresponds to a crossing scale of size $\delta = 2^l \times 0.000155569$. The presence of drift in the data illustrates the power of the binomial test, which rejects at all levels.

| xing levels ($l$) | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| **binomial** | 0.242 | 0.287 | 0.267 | 0.220 | 0.217 | 0.212 | 0.221 |
| **runs** | **0.000** | **0.000** | **0.000** | **0.000** | **0.012** | 0.727 | 0.153 |
| **# up** | 1221554 | 368482 | 84569 | 17258 | 4238 | 1034 | 257 |
| **# up or dn** | 2441279 | 736049 | 168681 | 34288 | 8362 | 2011 | 486 |
| **mean xing length** | 0.2 min | 0.5 min | 2.3 min | 11.2 min | 46 min | 3.1 h | 12.7 h |

Table 3: GBP-USD test results. $p$-values for the RW tests applied to 2003 GBP-USD exchange rate trade-by-trade data. Level $l$ corresponds to a crossing scale of size $\delta = 2^l \times 0.0000563285$.