Branching-ratio approximation for the self-exciting Hawkes process

Stephen J. Hardiman and Jean-Philippe Bouchaud

Capital Fund Management, 23 rue de l'Université, 75007 Paris, France

We introduce a model-independent approximation for the branching ratio of Hawkes self-exciting point processes. Our estimator requires knowing only the mean and variance of the event count in a sufficiently large time window, statistics that are readily obtained from empirical data. The method we propose greatly simplifies the estimation of the Hawkes branching ratio, recently proposed as a proxy for market endogeneity and formerly estimated using numerical likelihood maximization. We employ our method to support recent theoretical and experimental results indicating that the best fitting Hawkes model to describe S&P futures price changes is in fact critical (now and in the recent past) in light of the long memory of financial market activity.

I. THE SELF-EXCITING HAWKES PROCESS

The Hawkes model [1,2] is a simple and powerful framework for simulating or modeling the arrival of events which cluster in time (e.g., earthquake shocks and aftershocks, neural spike trains, and transactions on financial markets). In one dimension, the model is a counting process $N(t)$ with an intensity $\lambda(t)$ (the expected number of events per unit time) given by a constant term μ and a "self-exciting" term which is a function of the event history:

$$\lambda(t) = \mu + \int_{-\infty}^{t} \phi(t-s)dN(s). \qquad (1)$$

This self-exciting term gives rise to event clustering through an endogenous feedback: past events contribute to the rate of future events. $\phi(s) \geqslant -$ is the "influence" kernel which decides the weight to attribute to events $dN(s)$ occuring at a lag s in the past. The base intensity μ and the kernel shape $\phi(t)$ are parameters to be varied. A popular choice for the kernel is the exponential function $\phi(\tau) = \alpha e^{-\beta\tau}$ [3,4] but in general the kernel to be used should depend on the application or the dynamics of the data to be modeled. Note that for $\phi(t) = 0$ the model reduces to a Poisson process with constant intensity μ.

By taking the expectation of both sides of Eq. (1) and assuming stationarity {i.e., a finite average event rate $E[\lambda(t)] = \Lambda$}, we can express the average event rate of the process as $\Lambda = \mu/(1-n) \geqslant \mu$ where $n = \int \phi(\tau)d\tau$. One can create a direct mapping between the Hawkes process and the well known branching process [5] in which exogenous "mother" events occur with an intensity μ and may give rise to x additional endogenous "daughter" events, where x is drawn from a Poisson distribution with mean n. These in turn may themselves give birth to more "daughter" events, etc.

The value n, which corresponds with the integral of the Hawkes kernel is the branching ratio and determines the behavior of the model. If $n > 1$, meaning that each event typically triggers at least one extra event, then the process is nonstationary and may explode in finite time [6]. However, for $n < 1$, the process is stationary and has proven useful in modeling the clustered arrival of events in a wide variety of applications including neurobiology [7], social dynamics [8,9], and geophysics [10,11]. The Hawkes model has also seen many recent applications to finance [12–14], especially as

a means of modeling the very high-frequency events affecting the limit-order book of financial exchanges [15–19].

One novel application of the Hawkes framework to finance is as a means of measuring market endogeneity or "reflexivity" in financial markets [3,6]. In [3], the authors consider midprice changes in the E-mini S&P futures contracts between 1998 and 2010 and observe that the branching ratio n of the best-fitting exponential kernel model has been increasing steadily over this period, from $n \approx 0.25$ in 1998 to $n \approx 0.65$ in 2010 (see our Fig. 6 below).[1] They argue that this observation implies that the market has become more reflexive in recent years with the rise of high frequency and algorithmic trading and is therefore more prone to market instability and so-called "flash crashes."

In [20], however, we have argued that due to the presence of long-range dependence in the event rate of midprice changes (detectable in both 1998 and 2011) as one extends the window of observation, the best fitting stationary Hawkes model must in fact be critical, i.e., have a branching ratio $n = 1$. This is backed up by theoretical arguments and empirical measurements on market data.

Let us however insist that this conclusion only holds if one believes that Hawkes processes provide an exact representation of the reality of markets. It is very plausible that the dynamics of markets is more complicated [and involves, for example, nonlinearities absent from the Hawkes process defined by Eq. (1)], but that the best way to represent this dynamics within the framework of Hawkes processes is to choose $n = 1$ with a long-ranged influence kernel.

In this article we introduce a simple approximation for the branching ratio of the Hawkes process which allows us to faithfully reproduce the results of [3] which proposes the statistic as a measure of market instability and as a crash prediction metric. The interest of our approximation lies in its great simplicity: one need only estimate the mean and variance of the event count in a sufficiently large time window. The approximation also avoids a number of pitfalls [21] inherent to the significantly more complex approach [4] employed in [3].

———

[1]As in [3] the specific S&P futures contract that we study at any given moment in time is the most liquid expiry. This is the one closest to its rollover date.

The estimator accepts one parameter, a time window size W during which we measure the mean and variance of the event count. We note that when we employ our estimator to midprice changes in the S&P electronic futures market with a fixed window size W, then the branching ratio estimate obtained increases over time as reported in [3]. If, however, we allow the window size to scale appropriately (halving in size every 18 months) to adapt to the decreasing latency of interactions on the market we recover a constant branching ratio estimate as proposed in [20]. This result reiterates the need for a scale-invariant, or at least scale-sensitive means of measuring the "reflexivity" of financial market events.

II. MAXIMUM LIKELIHOOD ESTIMATION

Given observed events (e.g., midprice changes) at times $t_1, t_2, \ldots, t_n$ in an interval $[0, T]$, one can fit the Hawkes model by maximizing the log likelihood [4,22] over the set of parameters θ:

$$\ln L(t_1, \ldots, t_n | \theta) = -\int_0^T \lambda(t|\theta)dt + \int_0^T \ln \lambda(t|\theta)dN(t).$$
(2)

In the case of the exponential kernel $\theta = \{\mu, \alpha, \beta\}$. In practice, the model parameters $\{\hat{\mu}, \hat{\alpha}, \hat{\beta}\}$ which maximize this log likelihood are obtained with numerical techniques due to the lack of a closed form solution.[2] The branching ratio estimate is then $\hat{n} = \hat{\alpha}/\hat{\beta}$.

However, there are a number of pitfalls to using this procedure as a means of estimating the Hawkes branching ratio $n = \int \phi(\tau)d\tau$ [21]. Arguably the most important of which is that any estimate of n made in this manner will be heavily dependent on the choice of kernel model (e.g., exponential, power law, etc.) It may be that the chosen model cannot satisfactorily describe the observed events; hence the meaning of the branching ratio extracted from the maximum likelihood method is questionable.

Care must also be taken when employing this method in the presence of imperfect event data as illustrated in [21]. In one of their figures, the authors present a (negative) log-likelihood surface which features two minima (one local and one global). The global log-likelihood minimum does in fact little more than describe packet clustering inside the millisecond which arises from the manner in which events arriving at the exchange at different times are bundled and recorded with the same timestamp. Subsequent randomization of the timestamps inside a short time interval (in this case, 1 ms) creates a spurious high frequency correlation, that makes the global minimum irrelevant. The local log-likelihood minimum, which is in fact a better fit to the "true" lower frequency dynamics, does a poor job at explaining the spurious high frequency clustering and is punished with a lower log likelihood. Indeed, when the authors of [21] choose to fit this local minimum they corroborate results presented in [20].

---

[2]Note that the method above is not the only method proposed to fit the parameters of the Hawkes process for financial applications; indeed a recent publication [18] proposes a fast albeit still parametric method for fitting the multivariate exponential Hawkes process.

We believe it is therefore essential to have additional checks (such as nonparametric methods [15]) at one's disposal to support results obtained from likelihood maximization. To address the pitfalls in the branching ratio estimation that arise from the model choice we propose a simple model-independent tool for the branching ratio approximation, in the next section, which accurately reproduces previous results of [3] and also indicates the criticality of the relevant Hawkes process which describes the market.

III. A MEAN-VARIANCE ESTIMATOR FOR THE BRANCHING RATIO n

We begin with a general expression relating the Fourier transform of the kernel function to the Fourier transform of the autocovariance $\nu(\tau) = E[\frac{dN(t)dN(t+\tau)}{dt^2}] - \Lambda^2$ of the event rate (see [1,15] for a derivation):

$$\hat{\nu}(\omega) = \frac{\Lambda}{|1 - \hat{\phi}(\omega)|^2}.$$
(3)

Setting $\omega = 0$ we obtain a relation between the branching ratio, the average event rate, and the integral of the autocovariance (in the stationary case $n \leqslant 1$):

$$\int_{-\infty}^{\infty} \nu(t)dt = \frac{\Lambda}{\left|1 - \int_0^{\infty} \phi(t)dt\right|^2} \equiv \frac{\Lambda}{(1-n)^2}.$$
(4)

Therefore, to deduce the branching ratio of the stationary Hawkes process, we need only find the value of Λ and $\int_{-\infty}^{\infty} \nu(t)dt$. Estimating Λ is trivial; it is given by the empirical average number of events per unit time. To estimate $\int_{-\infty}^{\infty} \nu(t)dt$, we consider the variance of the total event count N_W in a window of length W. Theoretically, this is given by

$$\sigma_W^2 = E\left[\int_{t=0}^{W} \frac{dN(t)}{dt}dt \int_{t'=0}^{W} \frac{dN(t')}{dt'}dt'\right] - (\Lambda W)^2$$

$$= \int_{t=0}^{W} \int_{t'=0}^{W} \nu(t - t')dt\,dt'$$

$$= \int_{t=0}^{W} \int_{\tau=-t}^{W-t} \nu(\tau)dt\,d\tau$$

$$= \int_{\tau=-W}^{W} \nu(\tau)(W - |\tau|)\,d\tau \leqslant W \int_{-\infty}^{\infty} \nu(\tau)d\tau.$$
(5)

We then note that for some constant R, provided that
(1) $\nu(t) \to 0$ for $|t| > R$,
(2) $W \gg R$ such that $(W - |t|)/W \approx 1$ for all $|t| < R$, then

$$\sigma_W^2 \approx W \int_{\tau=-\infty}^{\infty} \nu(\tau)d\tau,$$
(6)

and we find, using Eq. (4):

$$n \approx 1 - \sqrt{\frac{\mu_W}{\sigma_W^2}} := \tilde{n},$$
(7)

where $\mu_W = \Lambda W$ is the average number of events falling in a window of size W. The above expression has a very intuitive interpretation. When the variance is equal to the event rate, the process is obeying Poisson statistics and $n = 0$. Any increase in the variance above the event rate indicates
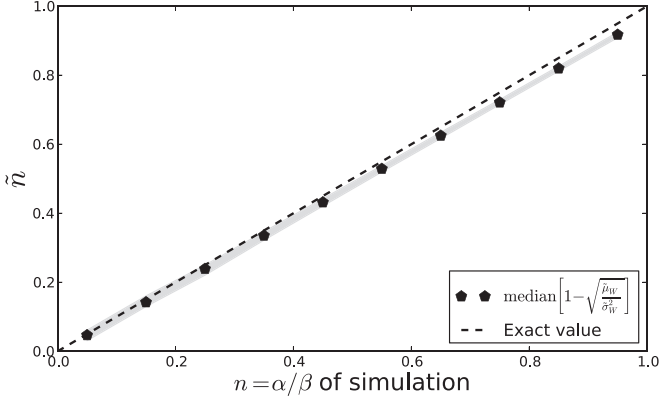
FIG. 1. Applying the mean-variance branching ratio approximation method to a simulated Hawkes processes with an exponential kernel form and scale parameter $\beta = 1.0$. α is varied to decide the branching ratio, and μ is varied to keep the average event rate fixed at $\Lambda = 1.0$. The process is simulated for $T = 100000s$, and the branching ratio estimate is made using Eq. (7) with a window size $W = 20$. The mean and [5%,95%] confidence bands are calculated over an ensemble of 100 processes for each parameter set.
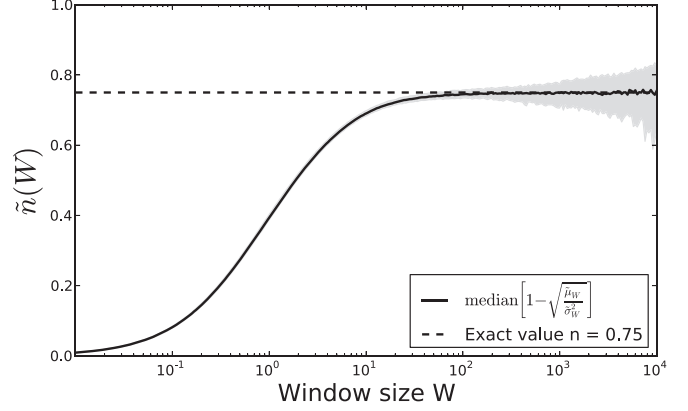

FIG. 2. Applying the mean-variance branching ratio approximation method to simulated data. Shaded area represents the [5%,95%] confidence interval. We note that the branching ratio estimate converges on the expected value of 0.75 as the window size increases. For very large W, we lack a sufficient number of event count observations to estimate the variance of N_W with precision and the confidence interval grows considerably.

some positive correlations and, within a Hawkes framework, a positive branching ratio.

Note from Eq. (5) that σ_W^2/W is a biased estimator for $\int_{-\infty}^{\infty} \nu(\tau)d\tau$ and for a general W will fall short of its theoretical value. This means that Eq. (7) will generally underestimate the branching ratio and only become exact in the limit $W \to \infty$.

In practice, to estimate the branching ratio of an empirical sample of total length $T = mW$ we substitute the mean and variance terms in Eq. (7) by their sample estimates:

$$\tilde{\mu}_W = W \frac{N_T}{T} \equiv \frac{1}{m} \sum_{i=1}^{m} N_W(i), \qquad (8)$$

$$\tilde{\sigma}_W^2 = \frac{1}{m-1} \sum_{i=1}^{m} (N_W(i) - \tilde{\mu}_W)^2. \qquad (9)$$

An estimate obtained in this fashion is asymptotically convergent for large $T \gg W$, i.e., $m \gg 1$. For a fixed window size W we can always ensure that our estimate of the variance of N_W is within a desired maximum error with a desired minimum probability by increasing T and therefore the number of observations m of N_W. Furthermore, we can make Eq. (6) exact by allowing W to be arbitrarily large.

Note, however, that for any finite m, the average of $\tilde{n}$ over all possible realizations of the process is in fact $-\infty$! This is because a value $\tilde{\sigma}_W^2 = 0$ is always possible with a nonzero probability. For this reason we choose to present the median of our estimates.

IV. NUMERICAL SIMULATIONS AND IMPLEMENTATION NOTES

In Fig. 1, we test the estimation procedure described in the previous section on a number of simulated exponential Hawkes processes with a variety of branching ratios. To do this we fix $\beta = 1.0$ but vary $\alpha = n$ in the range $0 \leqslant \alpha \leqslant 0.95$. We choose a base intensity $\mu = 1 - n$ such that each process has the same

average event rate $\Lambda = 1$. We simulate the process for a time $T = 10^5$ using simulation Algorithm 1 described in [23]. This procedure is prone to edge effects so we disregard the initial period of size 10^4 to ensure the process is close to stationarity in the period studied.

We estimate the branching ratio with a window size $W = 20$, significantly larger than the time scale of correlation of our process $\beta^{-1} = 1$. However, the window size chosen is also sufficiently small that we have a significant number of independent observations $m = (0.9 \times 10^5)/20$ with which to make a reliable estimate of the variance of N_W. We see very good agreement between our branching ratio estimates and the input branching ratio of the simulation; see Fig. 1.

Note however, that our approximation systematically underestimates the branching ratio since our finite window size does not completely cover the region of support of the autocorrelation function. This is particularly visible in Fig. 1 for large n. We now investigate the effect of window size on the estimate obtained.

A. Choice of window size W

For Eq. (6) to be accurate, we must choose a large W. However for a finite sample size, large W implies small $m = (T/W)$ and therefore less statistical power with which to estimate the variance of the event count N_W. This compromise is illustrated in Fig. 2 for which we simulate an ensemble of exponential Hawkes processes with parameters $\alpha = 0.75$, $\beta = 1.0$, and $\mu = 0.25$. We note that for increasing window size W, the confidence bands of our estimate converge on the expected value of $n = 0.75$.

However when we make the window size too large, we suffer significant error coming from the estimation of the variance. For the practical implementation of this procedure to empirical data we recommend a choice of window size sufficiently large to capture the bulk of the autocorrelation present in the event rate, but sufficiently small that one can
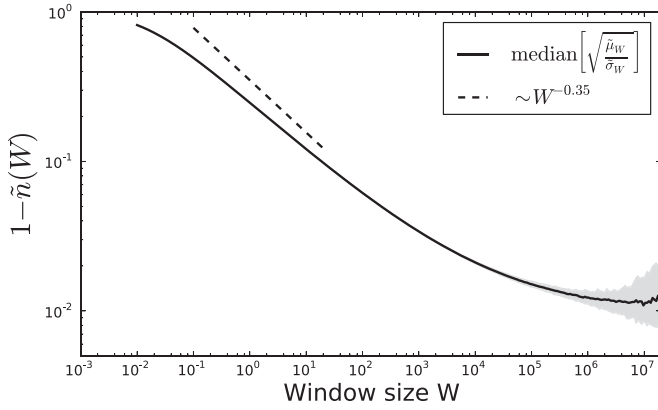
FIG. 3. $1 - \tilde{n}$ for simulated near-critical ($n = 0.99, \epsilon = 0.35$) power-law Hawkes process. The value for $1 - \tilde{n}$ that we obtain approximately scales as $\sim W^{-0.35}$ for large W. Note that the simulated process is only "near critical" (with a branching ratio $n = 1 - 1 \times 10^{-2}$) so for very large W the curve levels off and converges to 1×10^{-2}.

expect to obtain reliable estimates of the variance of the event count in that window.

One can approximate upper and lower confidence intervals on the branching ratio estimate from a single realization of a time series by bootstrap resampling. Indeed the simple variance estimator of Eq. (9) is not optimal and can be improved with the use of overlapping windows or, for example, a Monte Carlo sampling scheme which selects random windows with replacement.

B. Power-law kernel

Finally we test the estimation procedure on a (near) critical power-law Hawkes process, a type suggested in some recent publications as a good fit to financial event data [15,20]. Specifically, we consider a kernel with an Omori law form:

$$\phi(\tau) = \frac{n \epsilon \tau_0^\epsilon}{(\tau_0 + \tau)^{(1+\epsilon)}}. \tag{10}$$

In the critical case of $n = 1$ with $0 < \epsilon < 0.5$ this process will exhibit long-range dependence, with an autocorrelation function $\nu(\tau)$ decaying asymptotically as a power law: $\nu(\tau) \sim \tau^{-(1-2\epsilon)}$ [20,24]. The integral of the autocorrelation function is therefore divergent for large τ's, and the variance of the event count in a window of size W grows with as $\sigma_W^2 \sim W^{1+2\epsilon}$. The $\sqrt{\frac{\mu_W}{\sigma_W^2}}$ term in Eq. (7) will in this case not converge to a finite constant for large W but rather shrink to zero, leading to $1 - \tilde{n} \sim W^{-\epsilon}$.

We have simulated such a process with exponent $\epsilon = 0.35$ and cutoff parameter $\tau_0 = 1.0$. To allow our simulation to attain a stationary state with an average event rate $\Lambda \approx 1$, we make our process very near critical by choosing $n = 0.99$ and $\mu = 0.01$. We simulate for a very long period $T = 1 \times 10^9$ and discard the first 0.9×10^9 to ensure the process is close to stationarity ($\Lambda > 0.99$ at 0.9×10^9). In Fig. 3 we present the results of a branching ratio estimation using Eq. (7) with a variety of window sizes W.
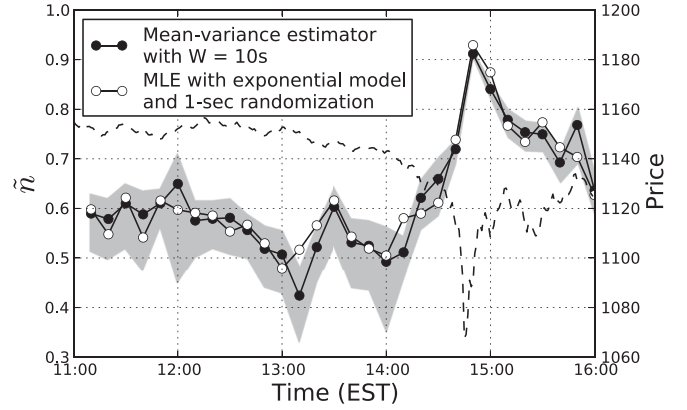

FIG. 4. Reproduction of the flash-crash branching ratio analysis of Filimonov and Sornette using the mean-variance estimator. Our results compare well with those obtained by maximizing the likelihood of the exponential model (with 1-s randomization). The dashed line is the E-mini S&P price. The points are placed at the end of the 10-min period for which the branching ratio estimate is made. The shaded region is a [10%,90%] confidence interval generated by bootstrap resampling.

We note that the branching ratio estimate we obtain is very much dependent on the choice of window size used. To capture all the correlation present in the process and obtain estimates close to the the true input value $n = 0.99$ we must probe the correlation on very large scales, by choosing a very large value for W. The branching ratio obtained indeed varies with window size according to the law $W^{-\epsilon}$, in a similar way to the integral of the kernel $\phi(\tau)$; see Fig. 3.

V. EMPIRICAL APPLICATIONS

A. Flash crash revisited

To demonstrate the effectiveness of this simple estimator we have repeated the flash-crash day branching ratio analysis of Filimonov and Sornette [3]. We consider nonoverlapping periods of 10 min in the hours of trading before and just after the flash crash. For each 10-min period, we calculate the sample mean and variance of the number of midprice changes in the 60 windows of length $W = 10s$ contained. The resulting values are plugged into Eq. (7) to obtain an approximation of the branching ratio for each period. The results are presented in Fig. 4.

Note that this simple estimator is biased, and for a general W, will typically underestimate the value of $\int_{-\infty}^{\infty} \nu(t) dt$ and therefore the branching ratio. Since we consider a window size $W = 10s$ we have systematically underestimated n in Fig. 4 as measurements of $\nu(t)$ on this data have support outside the interval $[-10s, 10s]$—there is still significant autocorrelation in the event rate at scales longer than $10s$.

However, we have identified that a window size W of the order of approximately 10 to 30 s produces estimates of the branching ratio on our data in line with those obtained by maximising the likelihood of the exponential model after intrasecond randomization (the method applied in [3].) Note that we do not fix β in our ML fit but let it settle naturally at the value which maximizes the log likelihood. We observe that this value $\hat{\beta}$ is very much dependent on the period
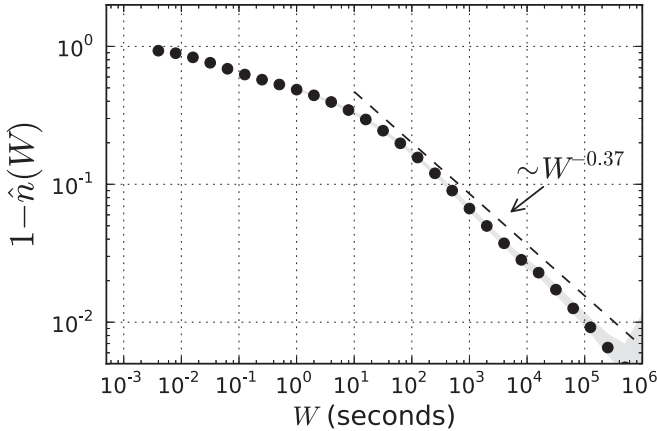
FIG. 5. $1 - \tilde{n}$, the estimate of the branching ratio as a function of window size for E-mini S&P midprice changes in 2010. The mean and variance of N_W are estimated on the full year. The change of power-law behavior occurs around 100 s. Note that we have "stitched" together all 5-min bins of regular trading hours (09:30 to 16:00 EST) that contain at least one event (this solves a problem arising from missing data at half days). We have then detrended the intraday seasonality by dividing each 5-min event count by a normalized average event rate for each 5 min of the trading day calculated on the full year.

of randomization[3] of the timestamps. When we randomize timestamps inside each millisecond we obtain $\hat{\beta}^{-1} \approx 10^{-2}$ for periods in 2010 but randomization at larger intervals (e.g., the intrasecond randomization of Filimonov and Sornette) prevents $\hat{\beta}^{-1}$ from exceeding values of the order of magnitude of the scale of randomization.

Note that since our results with $W = 10s$ correspond well with those obtained using the methods of Filimonov and Sornette [3], their procedure must *also* underestimate the branching ratio. To converge on the true n in expectation, we must take Eq. (7) in the limit of $W \to \infty$. We do just this in Fig. 5 for midprice changes of the E-mini S&P futures contracts in 2010. One notes that as the window size increases, the branching ratio converges to $n = 1$ in a nontrivial way. As reported in [20] for the structure of the kernel at short and long time scales, two regimes are detectable with a transition around five min. The branching ratio asymptotically tends towards $n = 1.0$ with a scaling exponent $\epsilon = -0.37$ compatible with the value of 0.45 estimated in [20] for a 14-yr period. Note that in taking the limit $W \to \infty$ we consider variation in the event rate at significant time scales to be explained by the stationary Hawkes model.

B. Reflexivity: 1998–2013

Using the mean-variance estimator, we can also easily reproduce the result of [3] that claims to demonstrate that reflexivity has been increasing in the S&P futures market since 1998. To do this we set our only parameter $W = 30s$

---

[3]To address the limited precision of the event data in [3], the authors randomize timestamps uniformly inside the second that they are reported.
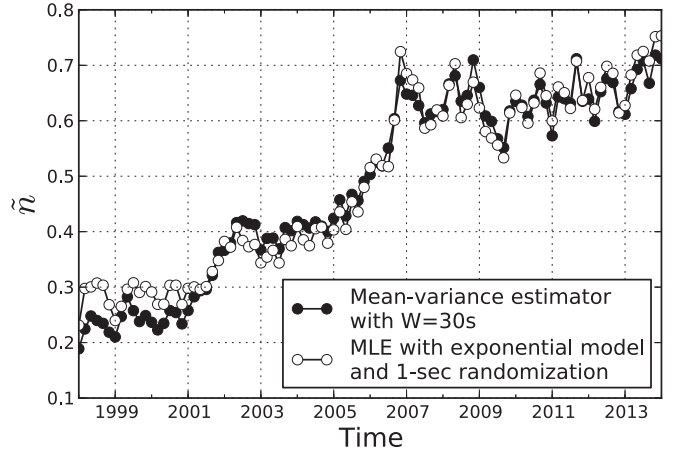

FIG. 6. Estimates of the branching ratio for 15-min windows using the method of Filimonov and Sornette [3] and estimates using the mean-variance estimator with a window size of $W = 30s$. Note that our MLE results differ somewhat from the plot presented in [3]. We attribute this to differences in the data source or identification of midprice changes.

and estimate the branching ratio in periods of 15 min. In Fig. 6 we present the 2-month medians of these estimates beside the median of the branching ratio estimate obtained using the exponential maximum likelihood approach after intrasecond randomization.

Since we expect the minimum time scale of correlation in the data to have decreased over the past decades (due to decreasing latency with advancing technology) we now reperform the experiment but with a window size W_t that follows Moore's law in such a way to keep the average number of events in W_t roughly constant. More precisely the window size W_t halves every 18 months; this describes well the increase in the high frequency activity of markets (see [20]). The results of this experiment are presented in Fig. 7 and confirm that the kernel integral is approximately invariant over time provided that the measurement window W_t is appropriately rescaled to account for the changing speed of interactions in the market. We find this quite remarkable, as this suggests that the amount of self-reflexivity in financial markets is *scale invariant*, and has not significantly increased due to high-frequency trading.

VI. SUMMARY

We have introduced a simple estimator for the branching ratio of the Hawkes self-exciting point process. The method is straightforward to apply to empirical event data since it requires only a rudimentary calculation on the mean and variance of the event rate. The estimator does not suffer from the bias inherent to the contentious question of the choice of kernel in the likelihood maximization approach, and furthermore it avoids the need for complicated numerical minimization techniques.

Despite its simplicity, our estimator can accurately reproduce results obtained for the branching ratio using this prior method. The estimator presented is indeed a biased estimator, but it can be made asymptotically unbiased in the limit of large W and T, for which we observe that the branching ratio
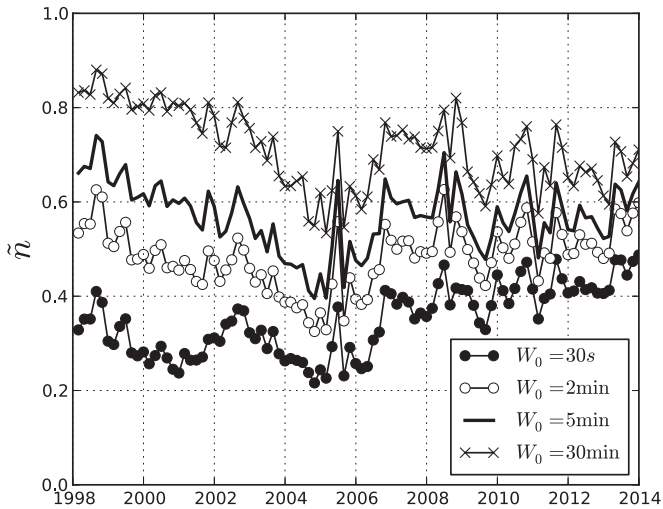
FIG. 7. Estimates of the branching ratio on 2-month periods using the mean-variance estimator with a window size that follows Moore's law: $W_t = W_0 e^{-c(t-t_0)}$ with $c = -\ln(1/2)/(18 \text{ months})$ and $t_0 = 1998$. We again stitch together periods of regular trading hours and detrend the event count by the intraday U shape for each year. When W is appropriately rescaled, the branching ratio estimate is approximately constant through time, for all values of W_0, and tends to $n = 1.0$ for large W_0.

for empirical midprice changes of the E-mini S&P futures contracts approaches unity in line with previous theoretical and empirical results [20].

Furthermore we demonstrate that if the window size is allowed to scale with Moore's law and adapt to the changing speed of interaction in the market over the past fifteen years, then the branching ratio approximation recovered is constant supporting prior observations of the invariance of the Hawkes kernel and branching ratio over time in [20].

Finally, let us reiterate the caveat made above: the Hawkes analysis of the activity in financial markets leads to the conclusion that the process must be critical. However, it might well be that the dynamics of markets is more complicated and requires nonlinearities absent from the Hawkes process defined by Eq. (1). More work on this issue would be highly interesting, and is in progress in our group.

[1] A. G. Hawkes, Biometrika **58**, 83 (1970).

[2] A. G. Hawkes, J. R. Stat. Soc., Ser. B **33**, 438 (1971).

[3] V. Filimonov and D. Sornette, Phys. Rev. E **85**, 056108 (2012).

[4] T. Ozaki, Ann. Inst. Stat. Math. Part B **31**, 145 (1979).

[5] T. E. Harris, *The Theory of Branching Processes* (Dover Publications, Mineola, NY, 2002).

[6] V. Filimonov, D. Bicchetti, N. Maystre, and D. Sornette, J. Int. Money Finance **42**, 174 (2013).

[7] V. Pernice, B. Staude, S. Cardanobile, and S. Rotter, PLOS Comput. Biol. **7**, e1002059 (2011).

[8] R. Crane and D. Sornette, PNAS **105**, 15649 (2008).

[9] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, J. Am. Stat. Assoc. **106**, 100 (2011).

[10] Y. Ogata, Pure Appl. Geophys. **155**, 471 (1999).

[11] A. Helmstetter and D. Sornette, J. Geophys. Res.: Solid Earth **107**, ESE 10-1 (2002).

[12] G. Bormetti, L. M. Calcagnile, M. Treccani, F. Corsi, S. Marmi, and F. Lillo. arXiv:1301.6141.

[13] L. Bauwens and N. Hautsch, in *Handbook of Financial Time Series*, edited by T. Mikosch, J.-P. Kreiss, R. A. Davis, and T. G. Andersen (Springer, Berlin, 2009), p. 953ff.

[14] P. Embrechts, T. Liniger, and L. Lin, J. Appl. Probab. A **48**, 367 (2011).

[15] E. Bacry, K. Dayri, and J.-F. Muzy, Eur. Phys. J. B **85**, 157 (2012).

[16] E. Bacry and J.-F. Muzy, Quant. Finance **14**, 1147 (2014).

[17] I. M. Toke and F. Pomponio, Economics **6**, 1 (2012).

[18] J. Da Fonseca and R. Zaatour, J. Fut. Mark. **34**, 548 (2014), doi:10.1002/fut.21644.

[19] E. Bacry, S. Delattre, M. Hoffmann, and J.-F. Muzy, Quant. Finance **13**, 65 (2012).

[20] S. J. Hardiman, N. Bercot, and J.-P. Bouchaud, Eur. Phys. J. B **86**, 1 (2013).

[21] V. Filimonov and D. Sornette, arXiv:1308.6756.

[22] I. Rubin, IEEE Trans. Inf. Theory **IT-18**, 547 (1972).

[23] J. Møller and J. G. Rasmussen, Adv. Appl. Probab. **37**, 629 (2005).

[24] P. Brémaud and L. Massoulié, J. Appl. Probab. **38**, 122 (2001).