Modelling Market Fluctuations under Investor Sentiment with a Hawkes-Contact Process*†

Junhuan Zhang ‡1, Jiaqi Wen[1], and Jing Chen[2]

[1]School of Economics and Management, Beihang University, China
[2]School of Mathematics, Cardiff University, UK

Abstract

We present a new Hawkes-Contact model that combines a Hawkes process and a finite range contact process in order to model the stock price movements, especially under the impact of news and other information flows that could lead to contagious effects. To fully capture the underlying price process, we take the Hawkes process to track the full pathway of historical prices on their future movements and the contact process to capture the impact from news/investment sentiment. We compare this full model to a univariate Hawkes process that works as a benchmark model through analyzing their statistical properties using both simulated returns and the real five-minute returns of the crude oil index (Wind CZCE-TA). The statistical properties include probability density function (PDF), complementary cumulative distribution function (CCDF), Lempel-Ziv Complex (LZC). Our results show that the real returns' distribution is often far from normal but the simulated returns through the Hawkes or Hawke-contact model can achieve close fit to the real returns and exhibit similar statistical properties. More importantly, the Hawkes-Contact model performs better than the simple Hawkes model in capturing characteristics in the return movements, which indicates that the price evolution is also driven by the news announcements and sentiment created after them.

Keywords: Market Microstructure; Investor Sentiment; Hawkes Process; Contact Process.
JEL Classification Codes: G10

*We greatly appreciate the financial support from National Natural Science Foundation of China (grant number 71801008), and Beihang University (grant numbers KG12113201, ZF202S1876).

†Conflict of Interest: The authors declare that they have no conflict of interest.

‡Corresponding Author. Email: junhuan_zhang@buaa.edu.cn.

1 Introduction

Traditional point processes generally believe that past events have no effects on similar future events. For example, the Poisson process considers the occurrence of events to be incrementally independent. The Markov process also believes that historical information prior to the current event does not have any impact on the occurrence of the future events. However, Hawkes processes suggests that the development of future events are path-dependent on past events and they have been widely used in the analysis of financial time series (Hawkes, 1971a,b). Historical events are believed to increase the intensity of occurrence of the same type of events in the future. Once an event occurs, the intensity function tends to decrease exponentially or follow some sort of power-law decay, until the next event arrives. Alongside the its first application in earthquakes and aftershocks(Hawkes, 1973), many areas have been using Hawkes processes to better model events with contagious features.

In the last couple decades, such applications have been prevailing in finance. For example, Bowsher (2007) introduces a bivariate Hawkes process (mutual-excitation) to model the joint dynamics of trades and quotes. Abergel and Jedidi (2015) model a limit order book using a Markovian Hawkes process and find that the arrival of orders follows an exponential intensity function. Bacry et al. (2015b) apply Hawkes processes to estimate the volatility at the transaction level, subsequently the market stability. Chen et al. (2018) test the success rate of three common information criteria to assist model selection of exponentially-decayed Hawkes processes using high frequency financial data.

The Hawkes processes have great advantages and better accuracy in predicting the occurrence of future events as they take into account of the entire path of the event evolution. Therefore, when events have contagious nature, Hawkes processes will be particularly useful because the underlying processes of financial time series cannot be simple diffusion anymore. In addition, Hawkes processes are flexible regarding the statistical distribution assumptions of time series. Thus, they are often used to study asset price jumps. For instance, Bacry et al. (2013) associate the Hawkes process with positive and negative price jumps by coupling stochastic intensities of upward and downward price changes for multiple assets. Fonseca and Zaatour (2014) provide explicit formulas for the moments and the autocorrelation function of the number of jumps over a given interval in a self-exciting Hawkes process. Bacry et al. (2015a) propose a class of toy models based on Hawkes processes named as Hawkes Impact Model, HIM) to study the transient and decay impact of price updates. Fonseca and Zaatour (2015) construct a multivariate Hawkes process to analyse the trade prices and identify clustering features. Fičura (2015) also find evidence of jumps in exchange rates that exhibit both self- and cross-exciting properties. Zhang (2016) uses power-law Hawkes processes to capture the microscopic structures pertaining to limit order books and reproduce the intensity of price jumps. Clements and Liao (2017) further use Hawkes models to forecast the variance of stock index returns in order to identify jumps and cojumps. Khashanah et al. (2018) specially introduce a slightly depressing process to model the reverse phenomenon of self-exciting mechanisms, with a decline in the intensity of jumps observed in market regimes.

As indicated already, most studies in the literature choose exponential kernels to fit the intensity function of Hawkes processes. Errais et al. (2010) prove that exponential kernel has advantageous properties as it allows one to compute the expected value of arbitrary functions. Bacry et al. (2012) define a numerical method that provides a non-parametric estimation of the kernel shape in symmetric multivariate Hawkes processes. Hardiman et al. (2013) suggest a power-law kernel with a decay parameter between -1.15 and l -1.45. Bacry et al. (2016) develop non-parametric estimation for exponential Hawkes kernel using power law decay in the context of high frequency finance. Recently, Chen et al. (2020) suggest that a Mittag-Leffler type kernel could simplify the power-law based kernels used in some ETAS (Epidemic Type Aftershock Sequence) models such as (Ogata , 1988) in order to obtain the relevant spectral properties.

In recent years, with the advancement of technology and data science, the financial market has changed to be much more complex in nature and the information flows and dynamics in the market are inevitably far more complicated. Great attention has been focused on studying the behaviour of the market and its participants that challenges the classic efficient market hypothesis and modelling approaches. Investment sentiment has been particularly focused on and studies such as Chen et al. (2020) suggest that trading in the market tends to be driven by sentiment, especially after the 2008 crisis. One prominent method to model the investment sentiment is through a finite range contact process. This stochastic process is known for studying interacting particle systems (Zhang and Wang, 2010), for instance, epidemic spreading that mimics the interplay of local infections and recovery of individuals. They have applied such a finite range contact process to model stock prices that could behave like the epidemic outburst under extremely stressed market condition. Zhang et al. (2010) build a finite-range contact system to study the financial return distribution. Yang et al. (2015) further develop a finite-range multi-type contact model to investigate the stock prices, where they imitate the interaction and dispersal of different types of investment attitudes like virus spreading behaviour.

However, there are hardly any studies considering models that can deal with the cognitive factors in the context of a complex financial market. Yang et al. (2018) take the financial market as a bi-variate system and investigate the trading process using a Hawkes model, where the positive/negative returns as well as positive/negative sentiment form self- and cross-excitations flows and drive the trading forward. Although Yang et al. (2018) analyze the dynamics between two types of factors that drive trading, their Hawkes models focus on the types of events (e.g. returns or sentiment) that influence the market but concentrate on the features of contagious effect. In contrast, the contact model proposed by (Zhang and Wang, 2010) can depict the epidemic spreading process of investor sentiment and account for the herding behaviour of investors: both give more description of the characteristics of contagious effects.

Therefore, we propose a Hawkes-Contact model that allows us to model the underlying price process driven by both historical prices and investment sentiment. On one hand, the new model can sufficiently account for contagion and help analyze the stock market microstructure in a complex financial system. On the other hand, using Hawkes and contact processes to model historical prices and investment sentiment separately in one model allow us to better understand price evolvement and trader behaviour individually and collectively at the same time. We consider the Hawkes process as a benchmark model, which indicates the baseline intensity of the entire underlying process; and the contact model is further employed to measure impact from news-driven events, namely investor sentiment.

To robustly test the new Hawkes-Contact model, we fit it to both simulated return series and an intra-day crude oil index (Wind CZCE-TA) at a five-minute sampling frequency. We, then, can compare the statistical properties such as probability density function (PDF), complementary cumulative distribution function (CCDF), Lempel-Ziv Complex (LZC). The results show that both the simulated and real returns have fat-tails when we fit the benchmark and full models. The statistical properties of simulated returned are close to real returns in both models but the full model that introduce the contact process clearly perform better than the bench market model. When the basic intensity increases, the yield fluctuation decreases and the dispersion is more concentrated. We further find that the larger the attenuation coefficient β, the more the simulated returns deviate from the real returns. The complexity increases when the basic intensity increases, while the complexity decreases when the attenuation coefficient, the higher level of complexity. In the Hawkes-Contact model, with Contact Model's weight increasing, the distribution of the simulated returns gradually approach to that of the real returns. This suggests that the consideration of investor sentiment from the contact process can improve the Hawkes price model through including effects of investor sentiment diffusion. Therefore, the full model can represent the real market better.

The rest of the paper is organized as follows. Section 2 introduces the benchmark Hawkes model and the Hawkes-Contact model. Section 3 shows the empirical results of crude oil index, and the simulated returns from Hawkes model and Hawkes-Contact model. Section 4 makes conclusions.

2 Methodology

2.1 The Hawkes process Model

In the stock market, we assume that price changes are only affected by historical price fluctuation event. Then, we present a simple stock price model by a univariate Hawkes process as a benchmark model. This stock price model is given as follows.

The intensity of the stock price changes at the future time t is $\lambda(t)$, which is defined as follows.

$$\lambda(t) = \mu + \int_{u<t} \alpha e^{-\beta(t-u)} dN_u \tag{1}$$

where μ is the basic intensity of the stock price changes at the future time t, time $u \in (0, t)$, α and β are two parameters, N_u is a counting process that records the number of occurrences of stock price fluctuations. It is calculated as the rounding of 10000 times the percentage return of stock price. The percentage return in the calculation represents the ratio of return to stock, which is different from the log-return. The multiplier of return with 10000 ensures that small fluctuations of stock return can also be captured.

In the real financial market, the financial time series are discrete. Therefore, the simple stock model is introduced by a discrete one-dimensional Hawkes process. Then, we have,

$$\lambda(t) = \mu + \sum_{u<t} \alpha e^{-\beta(t-u)} N_u \tag{2}$$

Then, the stock price at time t, P_t, is given as follows.

$$P_t = P_0 \exp\left\{ \mu + \sum_{u<t} \alpha e^{-\beta(t-u)} N_u) \right\} \tag{3}$$

where P_0 is the initial price at time 0.

2.2 The Hawkes-Contact Model

To our knowledge, there are few studies on the application of Hawkes process and contact model on the stock market microstructure modeling by considering both investor sentiment and market price fluctuation. In the paper, we present a new stock price model by considering the factors of the investor sentiment and the historical price. The investor sentiment interaction process is modeled by a finite range contact process. The impact of historical price on the future price is modeled by one-dimensional Hawkes process. That means the stock price is driven by both the historical price and the investor sentiment. We assume that the stock price is driven by both historical price fluctuation and investor sentiment. The investor sentiment diffusion process is inspired by (Zhang and Wang, 2010).

Let's assume that $n+1$ investors are in a line as $\{-2/n, \cdots, -1, 0, 1, \cdots, n/2\} \subset \mathbb{Z}$. Each investor can trade at intraday time s on the t-th day ($t \in (1, 2, 3, \ldots, T)$). There is at most one unit of stock which maybe traded each time. l is the intraday trading time length in a trading day. Let $P_t(s)$ be the stock price at the intraday time s in the tth trading day and $s \in [0, l]$. At time $(s - \Delta s)$, the investor have three kinds of sentiments. They are positive sentiment ($\xi_s = 1$), negative sentiment ($\xi_s = -1$) or neutral sentiment ($\xi_s = 0$) with probability p_1, p_{-1} or $1 - (p_1 + p_{-1})$. ξ_t is a continuous-time Markov process in the configuration $\{0, 1\}^{\mathbb{Z}^d}$. Investors in A ($A = \{x \in \mathbb{Z}^d : \xi_s = 1\}$) have non-neutral sentiments, while other investors are neutral.

The transition processes for $\xi_s(x)$ are given by: (a) $A \to A \backslash \{x\}$ for all $x \in A$ at rate 1, and (b)$A \to A \cup \{x\}$ for all $x \notin A$ at rate $\gamma |y \in A : |y - x| \leqslant R|$, where R is the diffusion range among investors. $|A|$ represents the cardinality of finite set A, and $y - x$ is the minimal length of a path from x to y. Then, at the initial time $(s - \Delta s)$, investors have initial state $\xi_0^A = A$ and initial distribution is ν_θ. From time $(s - \Delta s)$ to s, investors may change their sentiment based on the communication among each other.

For a fixed $s \in [0, l]$ (l is large enough), let $B_t(s)$ be the intensity at which the investors have different sentiments. Then, $B_t(s)$ is given as follows.

$$B_t(s) = \xi_0^\theta |\xi_s^\theta| / n, \quad s \in [0, l] \tag{4}$$

where n depends on the trading day t. We call this process as the investor sentiment diffusion process.

Base on the result of Section 2.1, we get the intensity of the stock price changes at the future time s in the t-th day, $\lambda(t)$, which is defined as follows.

$$\lambda_t(s) = \mu + \sum_{i<s} \alpha e^{-\beta(s-i)} N_t(s) \tag{5}$$

Because we assume that the stock price is driven by the investor sentiment, and historical market price fluctuations. Then, we get the stock price, $P_t(s)$, as follows.

$$P_t(s) = e^{m_1 \lambda_t(s) + m_2 B_t(s)} P_{t-1}(s), \quad s \in [0, l] \tag{6}$$

where t is the trading day, s is the intraday time at t-th day, m_1 and m_2 represent a pair of depth parameters of the market.

Further, we get the price as follows.

$$P_t(s) = P_0 \exp\left\{ \sum_{k=1}^{t} [m_1 \lambda_t(s) + m_2 B_k(s)] \right\}, \quad s \in [0, l] \tag{7}$$

where m_1 and m_2 represent the weights of the Hawkes model and the Contact model.Then, we use the following formula to calculate the rate of return.

The log-return at time s on the t_{th} day is given as follows.

$$r_t(s) = \ln P_t(s + \Delta s) - \ln P_t(s), \quad s \in [0, l] \quad t = 1, 2, \ldots, T. \tag{8}$$

3 Results

In this section, we present the results from studying both the Hawkes and Hawkes-Contact model. The former can also be considered as a base model for the underlying price process while the latter is a full model that further designed to capture the news/investment sentiment impact. For both models, we conduct simulations to obtain the set of parameters that allow the model to approach the real return distribution as closely as possible. We compute the PDF, CCDF and LZC as well as the descriptive statistics of the simulated distributions to compare with the real return distribution of the crude oil index. We also calculate the Mean Squared Error (MSEs) to ensure the robustness of the model fitting (see Sections 3.2 and 3.3) In the Hawkes-Contact model, such procedure is inevitably more complex due to the full model being more comprehensive in nature. A weighting factor (m_1/m_2)that connects the Hawkes and Contact parts and we follow a few steps calibrate this full model: first, we assign a random weight to obtain a good set of parameters of the Hawkes component in the full model, which will allow the full model approach the real return distribution (see Section 3.3.1); next, we take this set of Hawkes parameters to estimate the range of the weighting factor(see Section 3.3.2); and finally, we apply the Genetic Algorithm (GA) to accurately estimate the weight factor m_1/m_2 (see Section 3.4).

3.1 Data

The data used in our paper is the five-minute intraday data for one crude oil index (Wind CZCE-TA Index) for from January 3, 2017 to October 31, 2018. Then, we analyze the statistical properties for both the simulation returns and intraday five-minute frequency crude oil index (Wind CZCE-TA Index). The statistical properties are PDF, CCDF, LZC. There are 31511 observations involved in the analysis. For the returns of the crude oil index, we obtain the mean ($7.560e-06$), variance(1.533E-06), skewness (-0.369), kurtosis (36.906), maximum (0.027) and minimum (-0.028) respectively. It is clear that this financial time series has a high peak and two fat tails, which is not normally distributed (see Table 2).

3.2 The Hawkes Model

This section analyzes the simulations of the Hawkes model. The Hawkes model has three key parameters, μ, α, β. We start the simulation with an initial value and gradually increase them for multiple iterations. To be specific, μ starts from 100 and have increments of 20 for 20 iteration; α starts with from 0.04 and increases by 0.002 for 20 iterations and β begins with 0.007 and go up by 0.05 for 20 iterations. Altogether we are able to obtain 8000 sets of parameter combinations, out of which, for each parameter, we plot the probability density distribution using its the minimum, average and maximum values. In this way, we hope to show the changes occurred in the probability density function when these parameters vary. We could also repeat similar idea to calculate the complementary cumulative distribution functions (CCDFs) and Lemel-Ziv Complex (LZCs). With the PDF, CCDF and LZC exercises, we would be able to obtain a more precise understanding of the optimal selection of the μ, α and β. Later, we also simulate the 5-minute log-returns using these different choices of parameters that echo the 5-minute data of the Wind CZCE-TA index (Wind CZCE-TA) in order to perform further statistical analyses.

First, we compare the probability density distributions between the simulated log-returns and the real returns of the crude oil (see Figures 1 to 3). We also plot the normal distribution and use its fitting legend in the plots to offer another layer of comparison. Both left and right tails are highlighted to show clearly how tail distribution vary when a specific parameter gets adjusted. In general, we find that, through right selection of parameters, the statistic characteristics of simulated returns through the Hawkes model could reach the actual return distribution. However, neither the real or simulated log-returns appear to be close to the normal distribution.

Figure 1 shows that the real data has a high peak and thin tails and do not exhibit a normal distribution. The Hawkes model simulations are clearly able to achieve better fit of probability density that approach to the real return distribution. First, when μ gets bigger from 100, the peaks of the simulated distributions do not vary much while the tails are becoming visibly "thinner" and moving further away from that of the "real" tails. This suggests that a larger μ would result in a more concentrated distribution with "thinner" tails that reflect less price fluctuations; while a smaller μ helps approach the 'real' return tail distributions, which catches 'real' price fluctuations. Second, we use a smaller $\mu = 100$ concluded from Figure 1 and vary α. Again,Figure 2 suggests that real return distribution is far from "normal" and when α goes larger from 0.04, the variation in peakness is not significant while the tails become increasingly "fatter" and deviate further away from the real "tails". In contrast, with a a lower $\alpha = 0.04$, the Hawkes model provides a good probability density distribution that captures the real probability density. We, now, can fix $\mu = 100$ and $\alpha = 0.04$ and alter β. Figure 3 shows that the real return distribution is not normal. When we increase β in order to use the Hawkes model simulate the real returns, the peaks become visibly lower and the tails get 'thinner' and further apart from the real tail distribution. As β adjusts the influence of historical data, a smaller β would imply stronger influence from historical data but relatively less impact on the simulation. Therefore, a larger β is likely to result in a "wide and fat" distribution in the simulation that deviate away from the real distribution. To conclude, we should opt for a relatively small the $\beta = 0.007$.

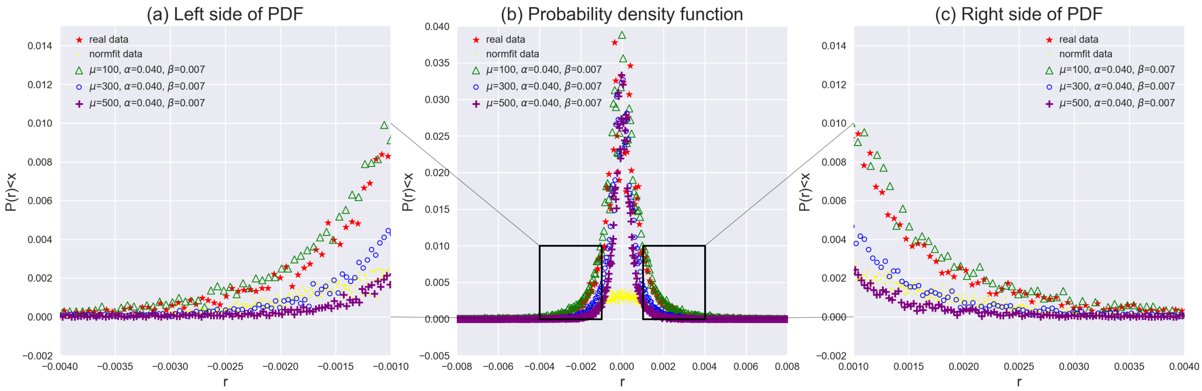

Figure 1: Probability density distributions simulated by Hawkes model using different μ

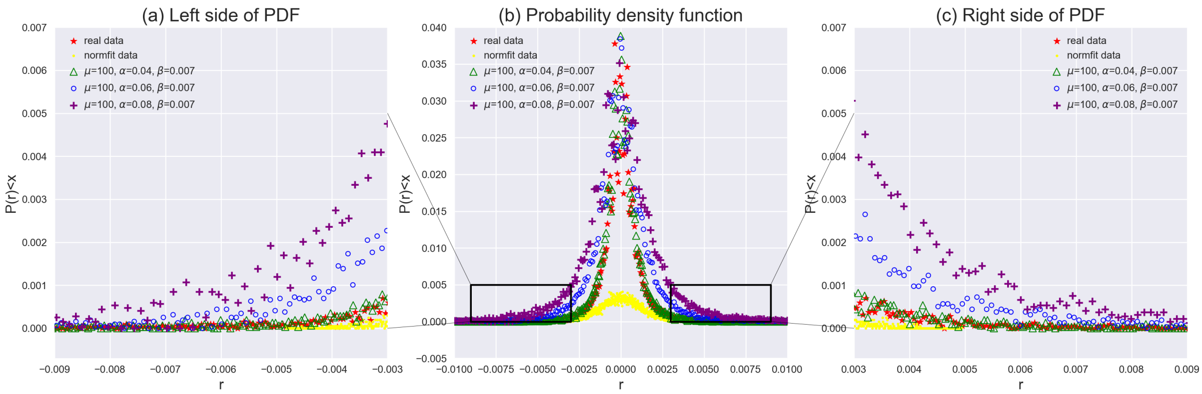

Figure 2: Probability density distributions simulated by Hawkes model using different α

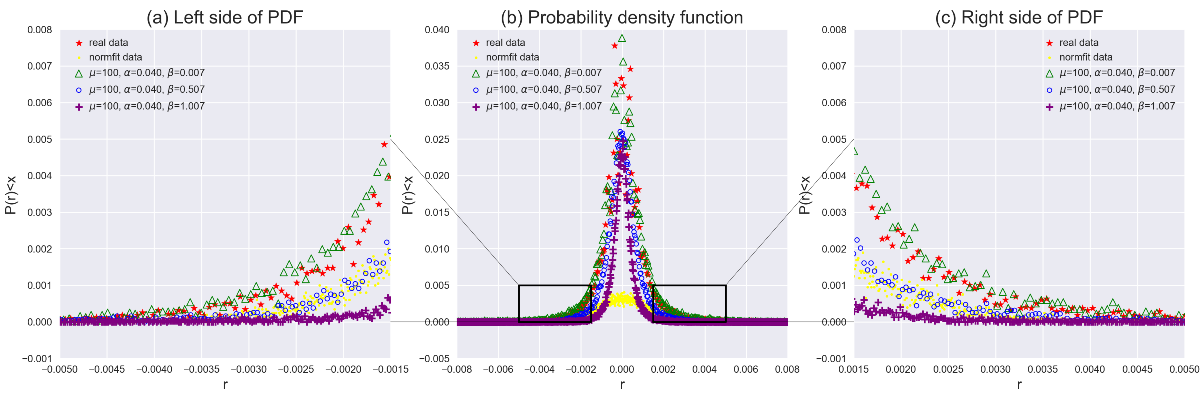

Figure 3: Probability density distributions simulated by Hawkes model using different β

Using the similar procedure, we calculate the CCDFs and LZCs. Figure 4 (a) shows that a smaler $\mu = 100$ would result in a better CCDF that is close to the real returns . From Figure 4 (b), we find that as α increases in the Hawkes model, the CCDF's fit starts to worsen. Therefore, the α should remain low (say 0.04). Finally, Figure 4 (c) indicates that the small $\beta = 0.07$, the simulated CCDF best fits with the 'real' CCDF. Overall, the CCDF analysis suggests that the real returns are not normally distributed and the Hawkes model with the correct parameter setting can accurately describe the real return distribution. Our results show that the parameter choices are consistent with the PDF analysis and that is with $\mu = 100$, $\alpha = 0.04$ and $\beta = 0.07$.

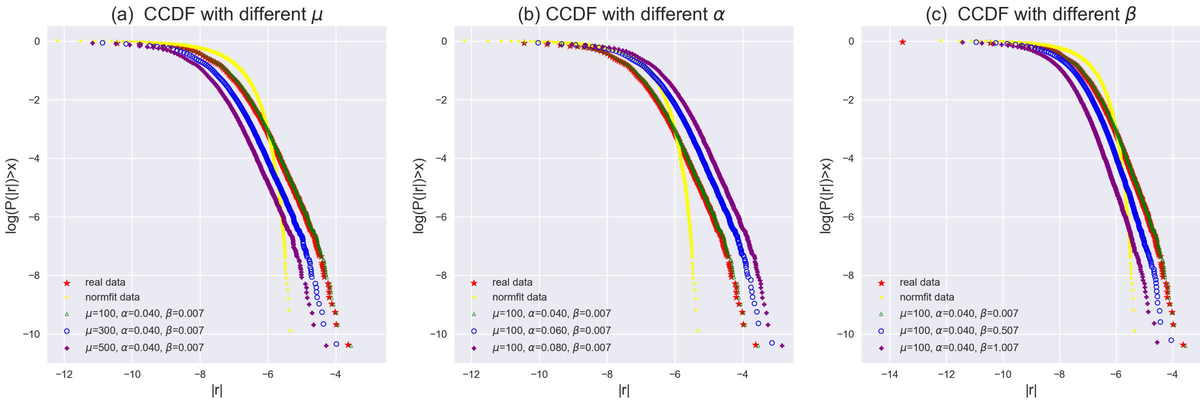

Figure 4: The Complementary Cumulative Distribution Function with different parameters

Finally, we examine the LZC, which is widely used in nonlinear daily studies to describe the rate at which new patterns appear in the time series as the sequence length increases. A larger LZC often indicates a more complicated time series that potentially contain more patterns. Therefore, it is appropriate to use it to examine the complex financial data series. In Figure 5, we examine the LZCs by using different μs, αs and βs. The general conclusion is, again, that the real return distribution is far from the normal distributions. However, the variations in parameter choices do not seem to affect LZC; in fact, from Panels (a) to (c) in Figure 5, it appears that the all Hawkes simulated LZCs fit well with the real data. In order to understand this better, we further compute the Mean Squared Error (MSE) between the simulated and real LZCs (see Table 1)and it is shown that the parameter setting of μ = 100, α = 0. 04 and β = 0.007 provides the smallest MSE. Thus, it further confirms our results from the PDF and CCDF analyses.

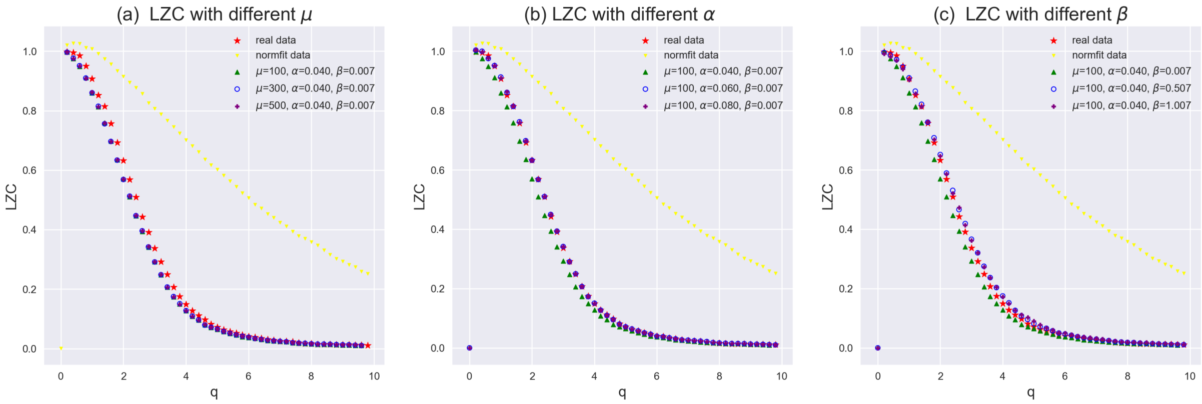

Figure 5: LZC versus q of $|r(t)|^q$ for the real data and simulated data with different μ, α and β

Table 1: The values of MSE for Hawkes model

μ	α	β	MSE
	normfit data		0.155900
100	0.04	0.007	0.000007
300	0.04	0.007	0.000008
500	0.04	0.007	0.000007
100	0.04	0.007	0.000007
100	0.06	0.007	0.000009
100	0.08	0.007	0.000009
100	0.04	0.007	0.000007
100	0.04	0.507	0.000193
100	0.04	1.007	0.000196

After experiment on the parameters of α, μ and β, we analyze the descriptive statistics of the Hawkes model (see Section 2.1) using different sets of parameters as we did in the above analyses. The return statistics including the Mean, Variance, Skewness, Kurtosis, Maximum and Minimum of the real crude oil index are reported in Table 2.We have a few observations: 1) higher μ or β make the kurtosis go down but the higher α make it to go up; 2) increased μ and β lead to higher skewness while increased α cause the skewness to drop; 3) when μ and β become larger, the variance move up before coming back down while the α increment leads to reduced variance; and finally the upper and lower bounds seem to be narrowed when μ and β go up because the maximum values drop as well as the minimum values rise. larger αs have the opposite effect that enlarge the range between the max and min values. Overall, we see that the μ and β increments tend to affect the return statistics to change in the same direction while the αs' impact tend to move to the opposite direction. Looking further, we can notice that the parameter setting of μ = 100, α = 0.04 and β = 0.007 actually provides a very close simulation to the real returns in its mean, variance, skewness, kurtosis, max and min are highly close to their real values - this implies that the Hawkes model can fit the real price-earning distribution well.

Table 2: The return statistics of Hawkes model with different α, μ and β

μ	α	β	Kurtosis	Skewness	Variance	Mean	Max	Min
real crude oil index			36.906	-0.369	1.533e-06	7.560e-06	0.027	-0.028
100	0.040	0.007	37.055	-0.377	1.736e-06	1.977e-07	0.029	-0.029
300	0.040	0.007	36.914	-0.360	7.367e-07	1.289e-07	0.019	-0.019
500	0.040	0.007	36.850	-0.353	4.051e-07	9.557e-08	0.014	-0.014
100	0.040	0.007	37.055	-0.377	1.736e-06	1.977e-07	0.029	-0.029
100	0.060	0.007	37.276	-0.401	3.899e-06	2.960e-07	0.043	-0.044
100	0.080	0.007	37.523	-0.427	6.923e-06	3.940e-07	0.058	-0.059
100	0.040	0.007	37.055	-0.377	1.736e-06	1.977e-07	0.029	-0.029
100	0.040	0.507	27.385	-0.219	8.091e-07	1.734e-08	0.018	-0.018
100	0.040	1.007	26.008	-0.117	3.527e-07	6.043e-09	0.011	-0.011

3.3 The Hawkes-Contact Model

We now simulate the Hawke-Contact model that reflect the stock prices with consideration of investors reactions to news sentiment. Again, we follow the same logic of the Hawkes model analysis in the previous section and examine the PDF, CCDF and LZC before comparing the descriptive statistics under different parameter settings. Referring to the Contact model proposed by Zhang and Wang (2010), we set the number of investors n as 40, the intensity λ as 10 and the parameter of initial investor sentiment state θ as 0.1. We additionally introduce m_1 and m_2 to represent the weights between the Hawkes and Contact parts in the Hawkes-Contact model, where $m_1 > 0$ and $m_2 > 0$ and $m_1 + m_2 = 1$. We need to find the parameters in the Hawkes part first, namely μ, α and β, and then we need to work out the weight setting with the optimal choice of the Hawkes part.

3.3.1 Obtaining the parameters of the Hawkes component in the Hawkes-Contact Model

First, we randomly set the weights $m_1/m_2 = 4/1$ that gives us m_1 and m_2 to be 0.75 and 0.25. Then, we fit the parameters in the Hawkes part, namely μ, α and β. We adjust their values following what we have done for the Hawkes model in Section 3.2: μ ranging from 100 to 500, α ranging from 0.04 to 0.08 and β ranging from 0.007 to 1.007. In this way, we can plot the

PDFs, CCDFs and LZCs (see Figures 6, 7 and 8) and collect return statistics to the simulations so that we can find the optimal parameters for Hawkes part.

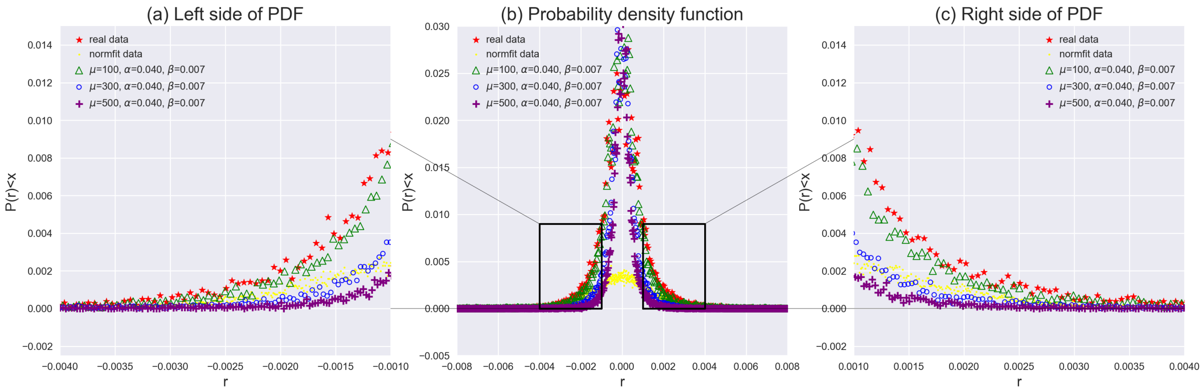

Figure 6: Probability density distributions simulated by Hawkes-Contact model using different μ when m_1/m_2 is randomly set to be 4/1.

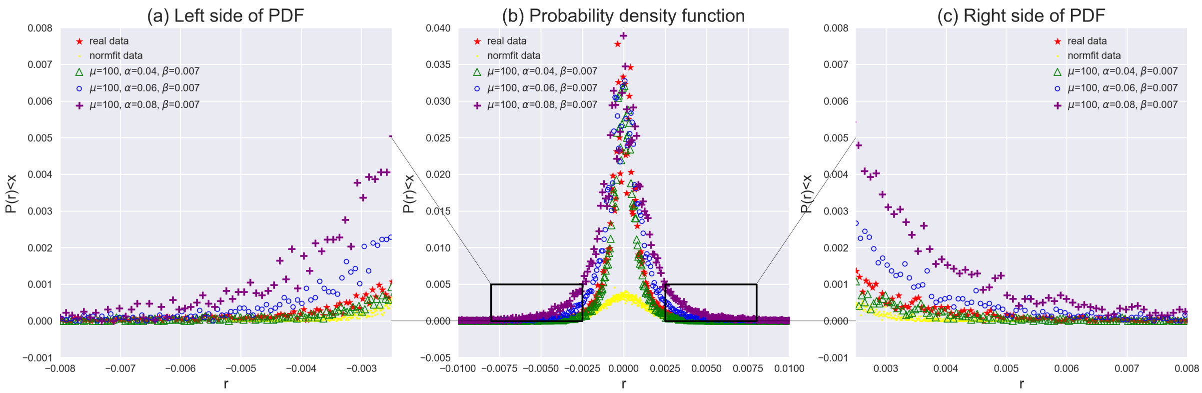

Figure 7: Probability density distributions simulated by Hawkes-Contact model using different α when m_1/m_2 is randomly set to be 4/1.

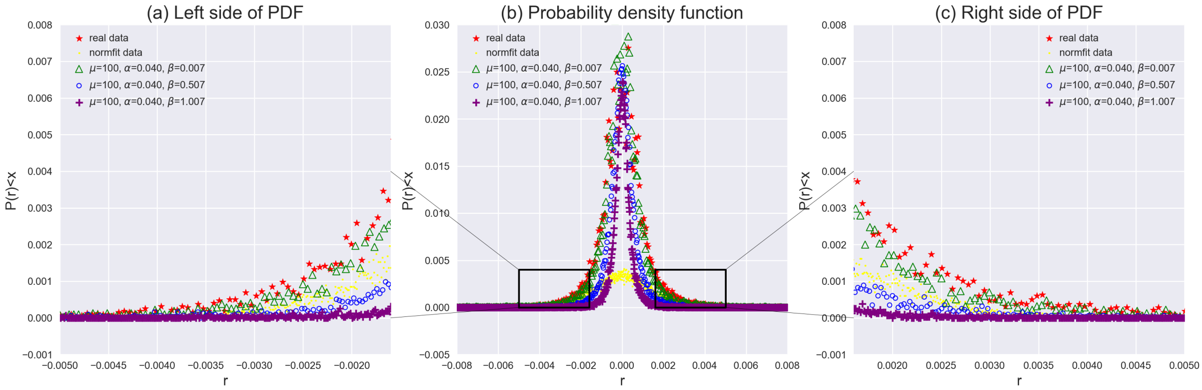

Figure 8: Probability density distributions simulated by Hawkes-Contact model using different β when m_1/m_2 is randomly set to be 4/1.

From Figure 6, the real return distribution is apparently far from the normal distribution while the PDFs of various Hawkes-contact models provide much better fit. When μ gets larger from 100, the peaks do not vary significantly but the tails get thinner and deviate away from the real data. The values of μ mainly influence the tail density under the influence of Contact model - this is consistent with the finding in the Hawkes model that a smaller μ is able to reflect the return behavior that tends to concentrate and fluctuate more at tails. As α increases from 0.04 in the Hawkes-Contact model, the tails of simulated distributions get fatter and gradually deviate from that of the real data. Although a larger α seems to peak more closely to the real peak, overall the peaks under different values of α do not exhibit clear difference (see Figure 7). At last, the PDFs of the Hawkes-Contact model simulation with a lower level of β indicates tighter fit with the real data at both its peak and tails. Overall, we can conclude that $\mu = 100$, $\alpha = 0.04$ and $\beta = 0.07$ under $m_1/m_2 = 4/1$ provide the best PDF that coincides with the real data, which is consistent with the Hawkes model's PDF analysis.

The next is to examine the CCDFs in the Hawkes-Contact model given $m_1/m_2 = 4/1$. In general, we see similar patterns to the Hawkes model when altering the values of μ, α and β. As μ, α 0.04 and β increase from 100, 0.04 and 0.007, the CCDFs start to move away from the "real" CCDF (see Figures 9) or normal distribution. It also shows that the CCDFs simulated by the Hawkes-Contact model have similar statistic characteristics to the CCDFs from Hawkes model, approaching closely to real data.

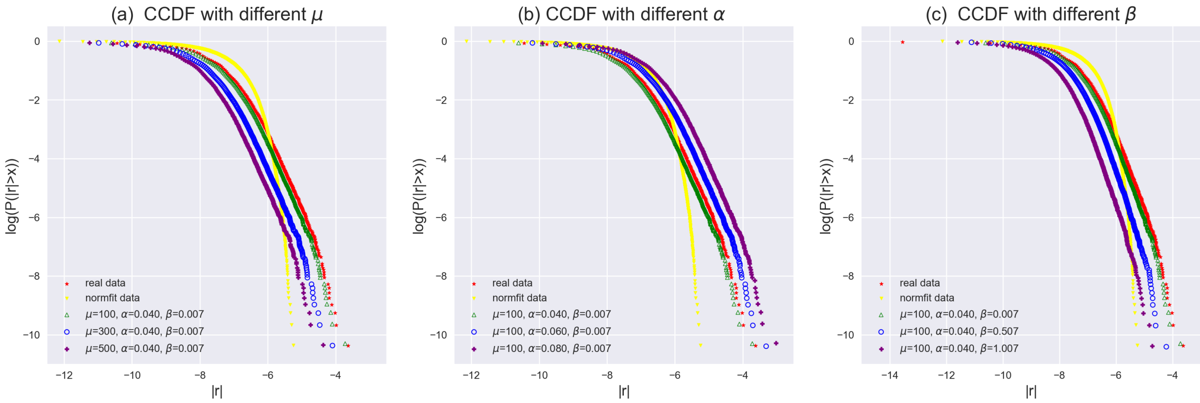

Figure 9: The Complementary Cumulative Distribution Function with different parameters when m_1/m_2 is randomly set to be 4/1.

At last, we examine the LZCs. Similar to the findings in the Hawkes model, all the LZCs fit well with the 'real' LZC and it is hard to differentiate from the different parameter selections (see Figure 10). Therefore, we further calculate the MSEs and prove that when $\mu = 100$, $\alpha = 0.04$ and $\beta = 0.007$, the simulated Hawkes-Contact model provides the best fitting to the real data given random ratio of 4/1 for Hawkes-Contact (See Table 3).

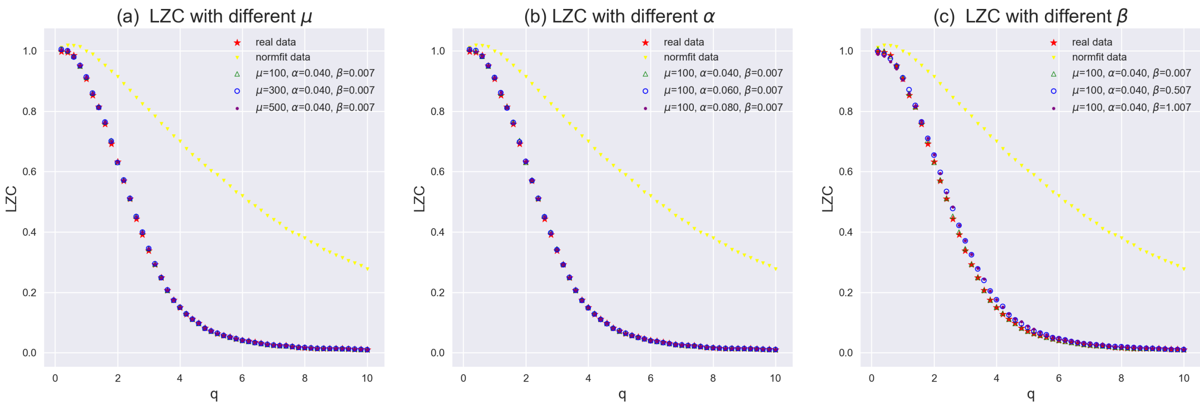

Figure 10: LZC versus q of $|r(t)|^q$ for the real data and simulated data with different μ, α and β when m_1/m_2 is randomly set to be 4/1.

Table 3: The values of MSE for Hawkes-Contact model when m_1/m_2 is randomly set to be 4/1

m_1/m_2	μ	α	β	MSE
	normfit data			0.155900
4/1	100	0.04	0.007	0.000010
4/1	300	0.04	0.007	0.000011
4/1	500	0.04	0.007	0.000012
4/1	100	0.04	0.007	0.000010
4/1	100	0.06	0.007	0.000080
4/1	100	0.08	0.007	0.000070
4/1	100	0.04	0.007	0.000007
4/1	100	0.04	0.507	0.000236
4/1	100	0.04	1.007	0.000281

Given random weight of 4/1 for components of Hawkes-Contact model and different parameter selection, we further report the return statistics including the mean, variance, skewness, kurtosis, max and min in Table 4. This helps us to identify the parameters in the Hawkes-Contact model that are best to describe the real data. Looking closely at Table 4, we observe that μ and β generally affect the return statistics in the same direction. For instance, when they increase, the kurtosis, mean and mean decrease, skewness and min increase, but the variance rise sharply first and fall back slightly afterwards. The α appears to affect these statistics opposite to the μ andβ. However, if comparing to the real crude oil index, the closest fit of return statistics come from the parameter selection of $\mu = 100$, $\alpha = 0.04$ and $\beta = 0.007$, which is consistent with the similar analysis for the Hawkes model. This parameter setting is kept for all the analyses below, where we change the weights of m_1 and m_2, and try to find the variances of simulations.

Table 4: The return statistics of Hawkes model with different α, μ and β when m_1/m_2 is randomly set to be 4/1.

m_1/m_2	μ	α	β	Kurtosis	Skewness	Variance	Mean	Max	Min
	real crude oil index			36.906	-0.369	1.533e-06	7.560e-06	0.027	-0.028
4/1	100	0.040	0.007	36.836	-0.366	1.238e-06	-3.441e-08	0.024	-0.025
4/1	300	0.040	0.007	36.746	-0.354	5.873e-07	-2.371e-08	0.017	-0.017
4/1	500	0.040	0.007	36.701	-0.347	3.416e-07	-1.809e-08	0.013	-0.013
4/1	100	0.040	0.007	36.836	-0.366	1.238e-06	-3.441e-08	0.024	-0.025
4/1	100	0.060	0.007	37.155	-0.386	2.777e-06	3.898e-08	0.037	-0.037
4/1	100	0.080	0.007	37.410	-0.406	4.928e-06	1.122e-07	0.049	-0.050
4/1	100	0.040	0.007	36.836	-0.366	1.238e-06	-3.441e-08	0.024	-0.025
4/1	100	0.040	0.507	27.059	-0.209	5.781e-07	-1.869e-07	0.015	-0.015
4/1	100	0.040	1.007	25.351	-0.103	2.536e-07	-1.891e-07	0.009	-0.009

3.3.2 Setting the Range of the Weighting Factor

After achieving the parameters (μ, α and β)of the Hawkes component in the full model, we now aim to find the range of m_1 and m_2: both m_1 and m_2 need to be positive and $m_1 + m_2 = 1$. Accordingly, if we vary m_1 starts from 0.9 and drop it by 0.1 each time until it reaches 0.1, m_2 will accordingly change from 0.1 to 0.9. We, subsequently, obtain 9 sets of Hawkes-Contact ratios (m_1/m_2) that are 9/1, 4/1,7/3,3/2, 1/1, 2/3, 3/7, 1/4 and 1/9. We take the "best" Hawkes model parameters of μ = 100, α = 0.04 and β = 0.007 (as mentioned in section 3.3.1) to run the Hawkes-Contact simulation and determine the optimal Hawkes-Contact ratio (m_1/m_2). Following similar logic, we first examine the distributions of these simulated returns through the Hawkes-Contact models, namely the PDFs, CCDFs and LZCs. Next, we check the return statistics under different full parameter sets (Hawkes-Contact ratio and Hawkes parameters of μ, α, β).For simplicity, we plot the PDFs, CCDFs and LZCs by using the maximum, average and minimum m_1/m_2 (9/1, 1/1 and 1/9) for simplicity (see Figures 11 and 12) but the full return statistics for the entire 9 sets of m_1/m_2 are presented in Table 4.

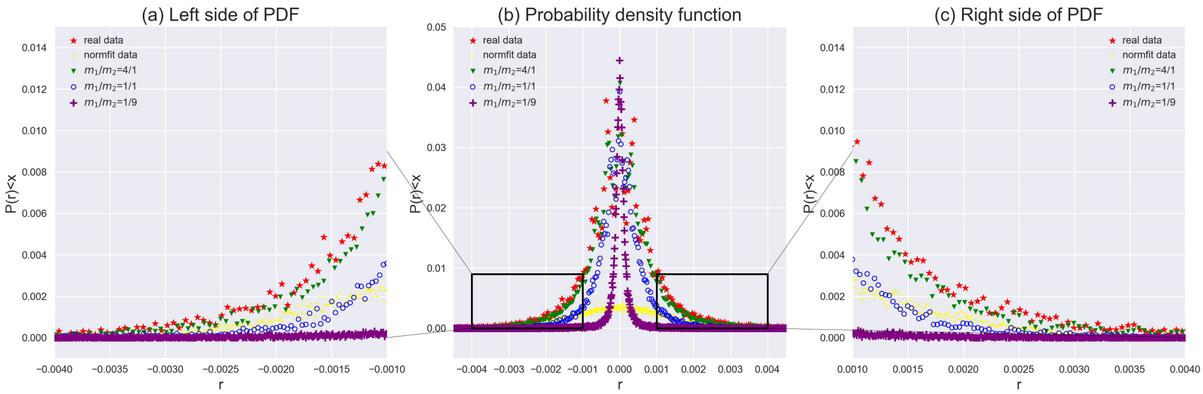

Figure 11: Probability density distributions simulated by Hawkes-Contact model using different m_1/m_2 when the parameters of Hawkes component are set to be μ = 100, α = 0.04 and α = 0.007.

Figure 11 demonstrates the PDFs of simulated Hawkes-Contact model with the maximum, average and minimum Hawkes-Contact weighting (m_1/m_2). It is evident that the real data distribution is not normal and the Hawkes-Contact model is able to fit the real data well, showing the good simulation at both peaks and tails. While the m_1/m_2 decreases, the distribution gets sharply peaked and the tails much "fatter". In comparison to the Hawkes model, these PDFs feature in higher peaks (>0.04)and fatter tails (>0.008)[1] due to the introduction of Contact model. Meanwhile, a high weight of Hawkes-Contact factor (9/1) yields better fit to the real data - this indicates that a properly weighted Contact component help capturing extra features real data. Given the characteristics and function of the contact models, these additional features tend to have clustering features. For financial data, this is greatly useful because either price/return clustering or other events such as investment sentiment are particularly insightful for investors to make trading decisions. Investors who believe in sentiment, comparing those who following the underlying price movement only, could behave differently, leading to sharp price fluctuations. Often, these price activities tend not to be mean-reverting and exhibit features such as skewed distributions with non-normal tails (see the cruid oil return distribution) Therefore, we suggest that the Hawke-contact model can better explain why financial series tend to have high peak and fat tails.

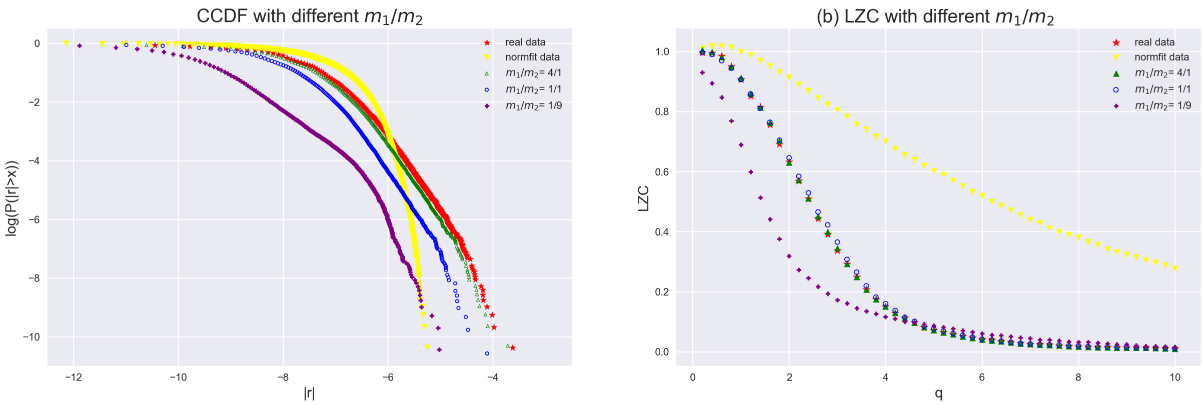

Figure 12: CCDF and LZC of Hawkes-Contact model using different m_1/m_2 when the parameters of Hawkes component are set to be μ = 100, α = 0.04 and α = 0.007.

Table 5: The values of MSE for Hawkes-Contact model when we set the parameters of Hawkes part as μ = 100, α = 0.04 and β = 0.007.

m_1/m_2	μ	α	β	MSE
	normfit data			0.155900
4/1	100	0.04	0.007	0.000008
1/1	100	0.04	0.007	0.000090
1/9	100	0.04	0.007	0.017746

Figure 12 further explores whether the Hawkes-Contact model provides better fit to the real

[1]The PDF for different β in the Hawkes model show similar tail level approaching 0.010

data. Figure 12(a) suggests that the real data are clearly not normally distributed and a high m_1/m_2 (4/1) performs better than an average or low ratio to reflect the real data distribution in CCDFs. Comparing to the CCDFs from the Hawkes model, the introduction of the Contact component differentiate the goodness of fit more visibly. Such visual difference is more evident in the plot of LZCs in Figure 12(b): when m_1/m_2 increases from 1/9 to 4/1, the LZC distinctively get closer to the real data - this indicates the better model choice because such distinction cannot be directly seen in the Hawkes LZCs and we need to lean on further MSE analysis. If we further compute the MSE for different m_1/m_2 and the MSE of the simulations with m_1/m_2=4/1 is the smallest (See Table 5).

Table 6 collects the return statistics of simulations from Hawkes-Contact model with varying Hawkes-Contact weights. These kurtosis, skewness, variance, mean, maximum and minimum are to compare between the simulated and real data. We sort the results by the Hawkes-Contact ratio (m_1/m_2) in a descending order. In general, a m_1/m_2 value larger than 3/2 tend to produce fairly good simulated return statistics that are similar to real returns. In particular, the kurtosis becomes larger and skewness smaller with the increased of m_1/m_2 ratios, which means that the high weight of the Contact component may lead to a lower peak and fatter tails than what the Hawkes model can capture. We also notice that only when the m_1/m_2 is at 9/1, the mean of the simulated series remain negative - this is not at all in line with the real mean. Therefore, we need to keep the Hawkes-Contact ratio high and this makes practical sense because the price fluctuation should be mainly driven by its underlying generating processes, instead of dominantly by clustering effects or sentiment that are associated with the market inefficiency or rare market events. In other words, we believe that financial time series generally move as random walks or in diffusion motions. Though with short-run fluctuations, extreme events such as contagion, bubbles, crisis etc. should not be the norm of the market condition that could disrupt the underlying price process of financial series.

Table 6: The return statistics of Hawkes-Contact model with different m_1/m_2,μ, α and β when the parameters of Hawkes component are set to be μ = 100, α = 0.04 and α = 0.007.

m_1/m_2	μ	α	β	Kurtosis	Skewness	Variance	Mean	Max	Min
statistics of real crude oil index				36.906	-0.369	1.533e-06	7.560e-06	0.027	-0.028
9/1	100	0.040	0.007	37.106	-0.369	1.480e-06	7.261e-08	0.027	-0.027
4/1	100	0.040	0.007	36.836	-0.366	1.238e-06	-3.441e-08	0.024	-0.025
7/3	100	0.040	0.007	36.368	-0.361	1.008e-06	-1.474e-07	0.022	-0.022
3/2	100	0.040	0.007	35.551	-0.354	7.925e-07	-2.668e-07	0.019	-0.020
1/1	100	0.040	0.007	34.109	-0.340	5.966e-07	-3.932e-07	0.016	-0.017
2/3	100	0.040	0.007	31.586	-0.310	4.248e-07	-5.272e-07	0.013	-0.014

Continuted ...

m_1/m_2	μ	α	β	Kurtosis	Skewness	Variance	Mean	Max	Min
3/7	100	0.040	0.007	27.694	-0.230	2.831e-07	-6.696e-07	0.010	-0.011
1/4	100	0.040	0.007	25.856	-0.010	1.786e-07	-8.212e-07	0.007	-0.007
1/9	100	0.040	0.007	45.397	0.504	1.202e-07	-9.830e-07	0.007	-0.005

Based on the above findings from Table 6, we find a rough range of m_1/m_2 from 3/2 to infinity, where m_1 ranges from 0.6 to 1 and m_2 varies between 0.4 and 0 accordingly. We then estimate the most suitable weighting parameters of Hawkes-Contact model within the estimated specific range based on the simulations of stock return series in the next section.

3.4 Estimating the Weighting Factor m_1/m_2

In this section, we estimate the weighting parameters for Hawkes-Contact (full) model using Genetic Algorithm (GA) method that treats the estimation process as an optimization problem through natural selection in order to solve both constrained and unconstrained problems(Papanicolaou and Papantoniou, 2016). Here, we aim to apply the GA method to solve the optimal weights for the Hawkes-Contact model so that this model can most closely fit to the real data. The estimation of weighting parameters m_1 and m_2 is based on their rough range proposed in section 3.3.2. The optimization objective as minimizing the difference of Kullback-Leibler Divergence (KL divergence) ((Kullback and Leibler, 1951))between simulated data and real data We then compare the descriptive statistics (including kurtosis, skewness, mean, variance, maximum and minimum) of the simulated series to the real data under the optimal m_1/m_2 weighting ratio.

The optimization process can be summarized as follows:

$$\underset{x}{argmin}\, f(x)$$
$$s.t. \begin{cases} x = [m_1, m_2], \\ 0.6 \leq m_1 \leq 1, \\ 0 \leq m_2 \leq 0.4, \\ m_1 + m_2 = 1. \end{cases} \tag{9}$$

where $f(x)$ represents the values of KL divergence between the simulation and the real data. x represents the parameters we need to estimate through the optimization process. μ, α and β involved in the optimization are each set as 100, 0.04 and 0.007, which are the same with the parameters in the above section in the aim of comparison. m_1 and m_2 each represent the

[1]The KL divergence, also called relative entropy, measures the difference of one probability distribution from another one, which increases from 0 to 1 with larger difference.

weights of Hawkes component and contact component in Hawkes-Contact model, which should be summed as 1. We set the specific range of m_1 and m_2 based on the findings in section 3.3.2, with m_1 ranging from 0.6 to 1 and m_2 ranging from 0 to 0.4. We then set the parameters needed in the optimization process, where we apply the genetic algorithm and set the maximum iterations as 1000 and the function tolerance as 1e-10.

The estimated m_1 is 0.780, and the estimated m_2 is 0.220. The KL divergence of the optimal simulation and the real data is 0.069. We then compare the simulations from the perspective of different weights of m_1 and m_2 in Table 7

Table 7: The return statistics of Hawkes-Contact model with different m_1/m_2,μ, α and β

m_1/m_2	m_1	m_2	μ	α	β	KL	Kurtosis	Skewness	Variance	Mean	Max	Min
statistics of log return of real crude oil index						0.000	36.906	-0.369	1.533e-06	7.560e-06	0.027	-0.028
39/11	0.780	0.220	100	0.040	0.007	0.069	36.762	-0.365	1.191e-06	-5.651e-08	0.024	-0.024
Inf	1.000	0.000	100	0.040	0.007	0.080	37.055	-0.377	1.736e-06	1.977e-07	0.029	-0.029
9/1	0.900	0.100	100	0.040	0.007	0.078	37.106	-0.369	1.480e-06	7.261e-08	0.027	-0.027
4/1	0.800	0.200	100	0.040	0.007	0.077	36.836	-0.366	1.238e-06	-3.441e-08	0.024	-0.025
7/3	0.700	0.300	100	0.040	0.007	0.076	36.368	-0.361	1.008e-06	-1.474e-07	0.022	-0.022
3/2	0.600	0.400	100	0.040	0.007	0.113	35.551	-0.354	7.925e-07	-2.668e-07	0.019	-0.020

As we can see from Table 7, the values of KL change with different values of m_1 and m_2. When m_1 and m_2 each take the values of 0.780 and 0.220, the simulations can achieve the smallest KL divergence 0.069, while the descriptive statistics approach the statistics of real data. We also find that when the weight of Contact model m_2 increases from 0.000 to 0.400, the value of KL divergence decreases. We already know that the Contact component helps describe important financial phenomena like investor sentiment, with the optimized Hawkes-Contact ratio, it can properly and more precisely suggest the weighting and impact from these events on the price movements through their statistical distribution characteristics. It forms critical guidance for us to understand the price discovery of financial time series, hence, influence the trading strategy, asset allocation, hedging and other important finance decision[2].

4 Conclusion

In the paper, we introduce a new type pricing model, the Hawkes-Contact model, that can account for full information impact on prices. This is associated with the fact that the underlying

[2]Some financial literature has already demonstrated that the trading in the contemporary market can be driven by returns (see (Liu. et al, 2020; Yang. et al, 2018; Chen et al., 2018, 2020; Chen and Liu., 2021))

process is not solely driven by the price movements themselves alone in the contemporary stock market. Instead, more complex information flows such as news sentiment also show strong influence on price updates, thus the new trading and risk strategies, such as sentiment trading start to prevail. The Hawked component forms the base model that tracks the price fluctuations that are driven by the price information flows. This is similar to but better than many classic models that are utilizing the Brownian Motion as this Hawkes pricing model can tolerate contagious effects as the Hawkes processes themselves allow to count the historic path of the price evolvement. The Contact component based a a finite range of contact process is dedicated to capture sentiment impact on prices. Altogether, the full model can fulfill the goal of capturing multiple information flows that jointly drive the underlying price process.

We start from the base (Hawkes) model and analyze the statistical properties of the simulation returns and examine how they behave comparing to the real return distributions of the five-minute crude oil index (Wind CZCE-TA Index). Our results suggest we are able to obtain a set of key parameters that provides a close fit to the real return distribution of the crude oil index and we verify them through looking at the probability density function (PDF), complementary cumulative distribution function (CCDF) and Lempel-Ziv Complex (LZC). The associated parameters include μ indicating the base-line intensity, α indicating the adjustment coefficient of the previous fluctuation to the current price, and β indicating the attenuation coefficient of historical stock prices. For further robustness, we also compare the descriptive statistics of all these simulated cases and the real returns.

We, subsequently, extend the base model to the full Hawkes-Contact model and introduce a weight coefficient m_1, m_2, to indicate the connection between the Hawkes and Contact components. We initially choose a random weighting in order to obtain the key parameters (μ, β and α)for the Hawkes part. Next, we use these parameter values to calibrate the full model in order to study the range of m_1/m_2 and how the weighting change could affect the simulated return series to approach the real returns. Finally, we apply the genetic algorithm (GA) to estimate the accurate weighting factor m_1/m_2 according the range we have obtained in the previous step. Therefore, we suggest that the Hawkes-Contact Model is a fitter model to interpret price fluctuations in a complex financial market.

References

Liu, A., Chen, J., Yang, S. Y., Hawkes, A. G. (2020). The Flow of Information in Trading: An Entropy Approach to Market Regimes. Entropy, 22(9), 1064.

Yang, S. Y., Liu, A., Chen, J., Hawkes, A. (2018). Applications of a multivariate Hawkes process to joint modeling of sentiment and market return events. Quantitative finance, 18(2), 295-310.

Abergel, F., Jedidi, A., 2015. Long-time behavior of a hawkes process–based limit order book. SIAM Journal on Financial Mathematics 6 (1), 1026–1043.

Bacry, E., Dayri, K., Muzy, J.-F., 2012. Non-parametric kernel estimation for symmetric hawkes processes. application to high frequency financial data. The European Physical Journal B 85 (5), 157.

Bacry, E., Delattre, S., Hoffmann, M., Muzy, J.-F., 2013. Modelling microstructure noise with mutually exciting point processes. Quantitative Finance 13 (1), 65–77.

Bacry, E., Iuga, A., Lasnier, M., Lehalle, C.-A., 2015a. Market impacts and the life cycle of investors orders. Market Microstructure and Liquidity 1 (02), 1550009.

Bacry, E., Jaisson, T., Muzy, J.-F., 2016. Estimation of slowly decreasing hawkes kernels: application to high-frequency order book dynamics. Quantitative Finance 16 (8), 1179–1201.

Bacry, E., Mastromatteo, I., Muzy, J.-F., 2015b. Hawkes processes in finance. Market Microstructure and Liquidity 1 (01), 1550005.

Bowsher, C. G., 2007. Modelling security market events in continuous time: Intensity based, multivariate point process models. Journal of Econometrics 141 (2), 876–912.

Chen, J., Hawkes, A. G., Scalas, E., Trinh, M., 2018. Performance of information criteria for selection of hawkes process models of financial data. Quantitative Finance 18 (2), 225–235.

Ogata, Y., 1988. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. Journal of The American Statistical Association 83 (401), 9–27.

Chen, J., Hawkes, A. G., Scalas, E., 2020. A Fractional Hawkes process. Nonlocal and Fractional Operators: Theory and applications to physics, probability and numerical analysis. SEMA SIMAI Springer Series Springer.

Chen, J., Liu,A., 2021. Information transition in trading and its effect on market efficiency: an entropy approach. Proceeding of the First International Academic Forum on Financial Mathematics and Financial Technology. Financial Mathematics and Fintech Springer 59-77.

Clements, A., Liao, Y., 2017. Forecasting the variance of stock index returns using jumps and cojumps. International Journal of Forecasting 33 (3), 729–742.

Dassios, A., Zhao, H., et al., 2013. Exact simulation of hawkes process with exponentially decaying intensity. Electronic Communications in Probability 18 (62), 1–13.

Errais, E., Giesecke, K., Goldberg, L. R., 2010. Affine point processes and portfolio credit risk. SIAM Journal on Financial Mathematics 1 (1), 642–665.

Fičura, M., 2015. Modelling jump clustering in the four major foreign exchange rates using high-frequency returns and cross-exciting jump processes. Procedia Economics and Finance 25, 208–219.

Fonseca, J. D., Zaatour, R., 2014. Hawkes process: Fast calibration, application to trade clustering, and diffusive limit. Journal of Futures Markets 34 (6), 548–579.

Fonseca, J. D., Zaatour, R., 2015. Clustering and mean reversion in a hawkes microstructure model. Journal of Futures Markets 35 (9), 813–838.

Fonseca, J. D., Zaatour, R., 2017. Correlation and lead–lag relationships in a hawkes microstructure model. Journal of Futures Markets 37 (3), 260–285.

Hardiman, S. J., Bercot, N., Bouchaud, J.-P., 2013. Critical reflexivity in financial markets: a hawkes process analysis. The European Physical Journal B 86 (10), 442.

Hawkes, A., 1973. Cluster models for earthquakes-regional comparisons. Bull. Int. Stat. Inst. 45 (3), 454–461.

Hawkes, A. G., 1971a. Point spectra of some mutually exciting point processes. Journal of the Royal Statistical Society. Series B (Methodological), 438–443.

Hawkes, A. G., 1971b. Spectra of some self-exciting and mutually exciting point processes. Biometrika 58 (1), 83–90.

Khashanah, K., Chen, J., Hawkes, A., 2018. A slightly depressing jump model: intraday volatility pattern simulation. Quantitative Finance 18 (2), 213–224.

Kirchner, M., 2016. Hawkes and inar (∞) processes. Stochastic Processes and their Applications 126 (8), 2494–2525.

Seol, Y., 2017. Moderate deviations for marked hawkes processes. Acta Mathematica Sinica, English Series 33 (10), 1297–1304.

Seonwoo, Y., Oh, A., Park, S., 2018. Hierarchical dirichlet gaussian marked hawkes process for narrative reconstruction in continuous time domain. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. pp. 3316–3325.

Yang, S. Y., Liu, A., Chen, J., Hawkes, A., 2018. Applications of a multivariate hawkes process to joint modeling of sentiment and market return events. Quantitative Finance 18 (2), 295–310.

Zhang, C., 2016. Modeling high frequency data using hawkes processes with power-law kernels1. Procedia Computer Science 80, 762–771.

Zhang, J., Wang, J., 2010. Modeling and simulation of the market fluctuations by the finite range contact systems. Simulation Modelling Practice and Theory 18 (6), 910–925.

Zhang, J., Wang, J., Shao, J. 2010. Finite-range contact process on the market return intervals distributions. Advances in Complex Systems 13 (05), 643-657.

Yang, G., Wang, J., Fang, W. 2015. Numerical analysis for finite-range multi-type stochastic contact financial market dynamic systems. Chaos: An Interdisciplinary Journal of Nonlinear Science 25 (4).

Chen, Y., Han, B., Pan, J. 2020. Sentiment Trading and Hedge Fund Returns. Journal of Finance. Forthcoming.

Papanicolaou, C.G., Papantoniou, I.C.. 11 - Optimum design of textile-reinforced concrete as integrated formwork in slabs. Textile Fibre Composites in Civil Engineering, 245–274.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Annals of Mathematical Statistics 22 (1), 79–86.