Process Systems Engineering

Travelling traders' exchange problem: stochastic modelling framework and two-layer model identification strategy

Chunbing Huang, Patrick M. Piccione, Federica Cattani, and Federico Galvanin

Just Accepted

Travelling traders' exchange problem: stochastic modelling framework and two-layer model identification strategy

Chunbing Huang[a], Patrick M. Piccione[b], Federica Cattani[c] and Federico Galvanin[a,*]

[a]*Department of Chemical Engineering, University College London, Gower Street, London WC1E 6BT (UK)*
[b]*Process Studies Group, Technology & Engineering, Syngenta, Breitenloh 5, CH-4333 Münchwilen, Switzerland*
[c]*Process Studies Group, Technology & Engineering, Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire, RG42 6EY (UK)*
[*]*f.galvanin@ucl.ac.uk*

Abstract

The Travelling Traders' Exchange Problem (TTEP) is formalised, aiming at studying the collision-exchange systems found in various research areas. As an example of the TTEP models, a 1-D model is developed and characterised in detail. The computational stochastic simulation of the 1-D TTEP model relies on a stochastic simulation algorithm implemented based on the Monte Carlo method. A model identification framework is proposed where the money distribution in the system obtained from the stochastic model is characterised in terms of (a) standard deviation of the money redistribution; (b) its probability density function. Results indicate that the expressions of the estimated functions for (a) and (b) are tightly related to the system input conditions. The example of curve fitting on the probability density function shows how the variation of money redistribution in

2

the system in time is driven by different values of the parameters describing the interaction mechanism.

Keywords: Travelling traders' exchange problem; collision-exchange system; Monte Carlo method; stochastic simulation

1. Introduction

Collision-exchange processes play a prominent role in a variety of natural phenomena and scientific applications. Collision refers to the consequence of particle movement in the system. Once two or more particles come across, the exchange process may occur and result in quantity transfers between involved particles. Systems in which these processes occur can be defined as collision-exchange systems. Stochastic collision-exchange systems are found in a number of relevant applications including chemistry[1], physics[2] and epidemiology[3]. In all these studies, computational approaches based on stochastic simulations[4] are frequently employed for the study and characterisation of systems in which uncertainty needs to be accounted for. This method is widely employed when it is difficult to describe and analyse the system in a deterministic way. In particular, Bansal et al.[3] studied a stochastic simulation of a compartmental model in epidemiology. Stochastic modelling was discussed for thermal conductivity in harmonic lattices.[2] Stochastic simulation approaches with various extensions and modifications for chemical reaction processes are presented by a number of researchers.[1,5,6]

As an important factor, the migration of particles in the system is sometimes non-negligible. Random migration does make a significant contribution to collision-exchange processes and eventually influences model behaviour. Stochastic models of random migration of molecules based on stochastic differential equations (SDEs)[7], sometime called Smoluchowski equations, are illustrated by Erban et al.[8]. This work presented an example on how a single diffusing molecule

randomly moves in a bounded (one-dimensional) region. Andrews and Bray[9] demonstrated a stochastic simulation of chemical reaction systems by treating each molecule as a point-like particle which freely diffuses in a three-dimensional space. They considered spatial diffusion and collisions of the molecules which may result in a forward/backward chemical reaction. In the stochastic modelling for bimolecular reaction-diffusion process by Erban and Chapman[10], an off-lattice model based on the simulation of Brownian motion of individual molecules and their collisions was discussed. These reports considering random migration of molecules (i.e. diffusion, Brownian motion) in systems give a more comprehensive understanding of the stochastic modelling. However, higher complexity as reflected by a larger number of variables required (i.e. "binding radius" and "unbinding radius"[9]) makes simulations computationally exacting. In stochastic models involving molecular diffusion in bounded space, boundary conditions are crucial for the simulation. Several examples of boundary conditions and their implementation are presented in the literatures[8,11,12].

In order to establish a comprehensive understanding on specific collision-exchange systems to increase predictive power for relevant applications, *The Travelling Traders' Exchange Problem* (TTEP) is formalised, aiming at studying the nature of the stochasticity in collision-exchange systems – and, eventually, linking average results to the understanding behind the laws of stochastic exchange. Regarding the stochastic systems, the Monte Carlo method[13], flexible and easy to implement in very general settings[14], is a powerful tool for modelling and simulating the systems in which collision-exchange processes occur and for which explicit differential equations are not available. Within this theoretical framework, a stochastic collision-exchange system is simulated via the discrete and sequential updating process, known as the stochastic simulation algorithm (SSA) or the Gillespie algorithm[15,16] and consequently a theoretical scheme to produce and analyse both the kinetics of ensemble quantities and the final state of the system is presented.

As uncertainties exist in systems, a stochastic simulation with the same input conditions would hardly produce exactly the same two realisations. That is, discrepancy among realisations is inevitable and it relies on the model complexity and nature of system uncertainties. So the

identification on the stochastic simulation results with respect to the entire set is of more significant importance. The visualisation of system changes in time would give a clear insight and understanding of stochastic collision-exchange systems with more details.

The goal of this paper is to develop a first stochastic model for describing the proposed TTEP system as well as a general procedure for the investigation of the computational simulation results of the stochastic model by parametric studies and regression model identification. In model identification, the standard deviation of money and money distribution in the system are the key variables to investigate how the system status changes over time. The standard deviation quantifies the variation and dispersion of amount of money in the system and the money distribution presents the relationship between frequency and amount of money held by traders.

This paper is structured as follows. Section 2 illustrates the stochastic modelling and simulations procedures with detailed introduction on TTEP stochastic model and identification strategy on the simulation results. Section 3 discusses the results of the case study on 1-D Model of TTEP system. Section 4 presents the conclusions.

2. Stochastic Modelling and Simulation

In order to investigate and set the scene for model identification of a stochastic system, a mathematical model is to be created, exact and computer-aided SSA based on certain simulation techniques are to be established, and computational simulation results for further data analyses needs be produced. Figure 1 shows the model building and simulation procedures employed for a stochastic system.
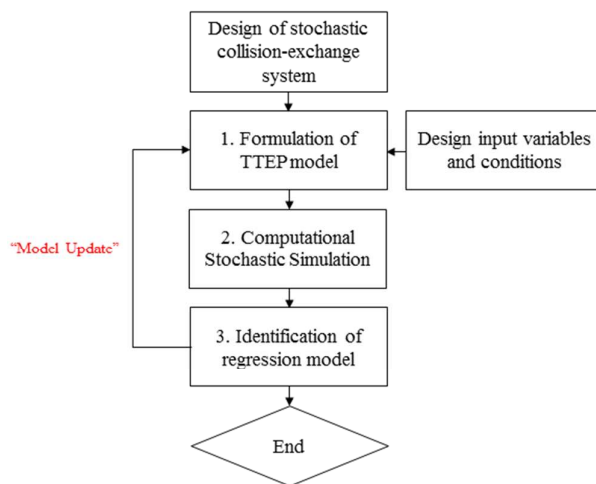
Figure 1. *Schematic flow chart of model building and stochastic simulation procedures of the stochastic collision-exchange system.*

The main stochastic modelling and simulation procedures comprise three fundamental steps. In the first step, the formulation of models for the TTEP is developed mathematically according to the nature of collision-exchange process in systems. Several general assumptions for the TTEP system are listed in Section 2.1. According to the formulation of TTEP, a number of candidate models can be built depending on the different levels of complexity that one would like to achieve. Pre-set variables and initial conditions generated during model construction will be considered as input conditions for the corresponding model in stochastic simulation. In the second step, given the specific model of TTEP system, the stochastic simulations are realised computationally with different scenarios of input variables and conditions. In the third step, after the stochastic simulation, data analysis is a vital procedure in which results obtained are studied with several techniques and methods. Given the simulation results, parametric studies illustrate the impact of operating variables on the simulation behaviour. Regression models are employed to numerically analyse the model evolution over time with different input scenarios. Each simulation result at fixed input scenario presents an observable prediction on the system evolution. If the regression is repeated for all the potential input scenarios, eventually, a predictive model of the TTEP system would be developed.

2.1. Formulation of models for the travelling traders' exchange problem (TTEP)

A test case of quite general applicability for the studies on collision-exchange systems is proposed in this paper, namely the Travelling Traders' Exchanges Problem (TTEP)[17]. The following assumptions are introduced for describing the TTEP problem, in order to identify the system by analysing a population of N traders who may randomly move in space and exchange money with each other:

1. The traders initially start with a certain amount of money M_0;

2. The trader population size is set as N according to the number of traders studied;

3. During a trading season lasting τ the traders may be considered free to move over a bounded territory in the most general framework, or fixed on the sites;

4. Each time the traders meet they may exchange money in various ways to be defined mathematically;

5. The total amount of money is conserved.
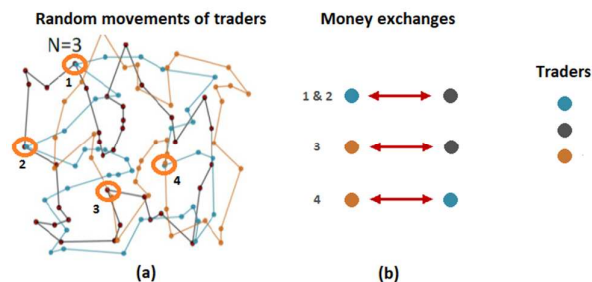
Figure 2. *General sketch of the TTEP problem when number of traders, N = 3.*

Figure 2 illustrates the sketch of the TTEP problem with number of traders $N = 3$. In Figure 2(a), three traders randomly migrate over time in the system with their trajectories recorded. In Figure 2(b), four examples of money exchanges are illustrated with different pairs of traders. This paper studies how the TTEP system dynamically evolves in the given processing time from a number of travelling traders starting with initially given amounts of money M_0 to the terminal

money redistribution of $\boldsymbol{M}_{T_{\text{total}}}$ through a series of interactions among traders. A number of vital model parameters such as time scale, random generators and nature of the collision processes do affect and drive the money redistribution, and give different degrees of contributions. The TTEP system will be introduced as the simplest paradigm of systems with collision-exchange processes.

2.2. Computational stochastic simulation on the 1-D Model

The TTEP presents a general model formulation for systems in which one can have interactions between traders and where the system evolution is determined by the way in which traders interact with each other. In physical applications, these traders refer to carriers or sites that are able to store or carry materials, of which amount or property vary in time and affect the system homogeneity. The complexity of specific TTEP model used to represent different industrial and experimental processes is based on corresponding assumptions leading to model simplification. Compared to stochastic models with multi-dimensions, it is more convenient to investigate the TTEP problem in a computationally tractable way using a simple one dimension model. The term "dimension" here refers to the dimension in which traders interact with each other, rather than the spatial movement of traders in the system. Since an important aspect of the work is systematically building up an understanding of how aggregate measures of stochastic models respond to changes in operating conditions, the present contribution focuses exclusively on the 1-D version of the interaction problem so as to gain initial insights on the emerging properties of the TTEP system.

One-dimension position-fixed linear model (or 1-D Model) is a fundamental model created to mathematically describe the TTEP system with the collision-exchange processes taking place. For simplification, not all rules listed above are applied. In particular, in the 1-D Model, $\boldsymbol{N} = (N_1, N_2, \cdots, N_N)$ traders, numbered from 1 to N, are sequentially placed in the following one-dimensional sequence (see Figure 3):
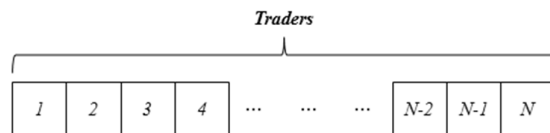
Figure 3. *One-dimension sequence of traders in 1-D Model*

The position of each trader remains unchanged during the simulation process, that is, the spatial change of each trader is not considered in this model and, consequently, the position of each point-like element (i.e. trader) is stationary and time-independent. The term "fixed-position" refers to the fact that each trader is specifically fixed in the sequence and its relative position does not change over time. Note that this restricts only the relative positions of traders while the absolute movements between two neighbouring traders are still considered in the model. Motion of traders is thus allowed so far as collisions are needed and dispersion of carriers (i.e. traders) can be considered as a further level of complexity. The traders are assigned an initial amount of money $\boldsymbol{M}_0 = (M_{01}, M_{02}, \cdots, M_{0N})$. The money exchanges are considered occurring sequentially in the system between arbitrary pair of traders who are next to each other. In this work, random motion of traders is simplified into instantaneous interactions between neighbouring traders. The time gap between every two adjacent exchange events, i.e. events S_i and S_{i+1}, is denoted by Δt

$$\Delta t = t_{i+1} - t_i, \quad \forall\, i \in \mathbb{Z}^+ \tag{1}$$

Further assumptions are:

(a) Only one money exchange takes place at each exchange event. This assumption is reasonable because the likelihood of simultaneous collisions of more than two traders are negligible since realistic "trader densities" over a large territory are relatively low. Two or more money exchanges occurring simultaneously are not considered in this model. Any two or more simultaneously-occurring money exchanges are regarded as separate and sequential exchange events, i.e. events S_i, S_{i+1} and S_{i+2} occurring simultaneously are considered as occupying three separate simulation time points, see Figure 4.

(b) For simplicity, the time gap Δt between two adjacent exchange events is assumed constant, see Figure 4. In reality Δt is a physical time gap which is not constant and does affect the collision-exchange processes. However, in this study, the focus is on the simulation behaviour of 1-D Model based on Monte Carlo method and Δt is regarded as a timer instead of a real physical time gap.

(c) Only two traders are involved in each money exchange event and multi-collisions involving three or more traders are not considered, i.e. binary collision only.

(d) Money exchanges occur only between two adjacent traders. That is, the traders at the boundary positions N_1 and N_N can only exchange money with the trader next to them, i.e. N_2 and N_{N-1}, respectively. The rest of traders have two exchange directions, forward and backward.
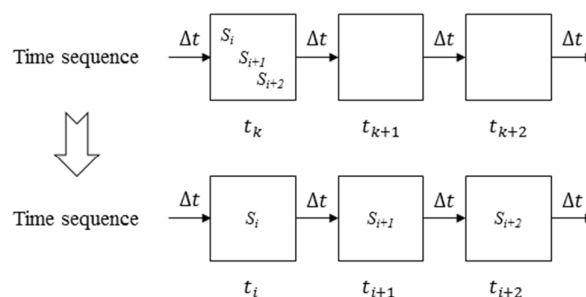
Figure 4. *Sketch of the allocation of simultaneously-occurring money exchange events into separate and sequential money exchange events in simulation*

With assumptions proposed above, the money exchange process between arbitrary two neighbouring traders can be simulated in time. For simulation implementation, some model variables and relative glossary of the 1-D Model are introduced and defined for the SSA construction.

- *Forward exchange propensity*, P_F – the probability that the active trader would like to execute a money exchange event "forward" to the next trader, i.e. $P_F = 0.5$ represents that if trader *n* is selected as active trader in one event, traders *n-1* and *n+1* have the same chance to be determined as passive trader.

- *Active trader* – the trader determined by the location L_i of the exchange event S_i. A random number generator from a discrete uniform distribution is applied to randomly select an event location and, consequently, an active trader in the model. The command ***unidrnd(*N*)*** in MATLAB® generates an independent positive integer $X_L \in \{x | 1 \leq x \leq N, \mathbb{Z}^+\}$ to determine the location L_i.

- *Passive trader* – the trader determined by the active traders' exchange direction, D_i, which relies on the forward exchange propensity P_F. A random number generator from a continuous uniform distribution is applied. The command ***rand(*1*)*** in MATLAB® returns an independent random value $X_D \in [0, 1]$. The exchange direction is then confirmed according to the equation below:

$$D_i = \begin{cases} forward, & X_D \leq P_F \\ backward, & X_D > P_F \end{cases} \tag{2}$$

The forward exchange propensity P_F is used to characterise the exchange direction in each exchange event. Denoting with P_B the corresponding backward propensity one has

$$P_F + P_B = 1 \tag{3}$$

The interaction mechanism of traders is crucial in TTEP systems and the money redistribution is the consequence of money exchange processes. When the exchange couples are determined, the money exchange occurs. Each money exchange in the 1-D Model comprises two money transfers: E_A denotes the money transferred from active trader to passive trader, E_P denotes the money transferred reversely from passive trader to active trader. Consider the following interaction mechanism:

$$E_A = k_1^A M_{S_{i-1}}^A + k_2^A \left(M_{S_{i-1}}^A\right)^2 + k_3^A \left(M_{S_{i-1}}^A\right)^3 + \cdots \tag{4}$$

$$E_P = k_1^P M_{S_{i-1}}^P + k_2^P \left(M_{S_{i-1}}^P\right)^2 + k_3^P \left(M_{S_{i-1}}^P\right)^3 + \cdots \tag{5}$$

Here $M_{S_{i-1}}^A$ and $M_{S_{i-1}}^P$ are the amount of money owned by active and passive traders before money exchange, respectively; k_i^A ($i = 1,2,3,\cdots$) is the money transfer coefficient of i-th order term of

active trader; k_i^P ($i = 1,2,3,\cdots$) is the money transfer coefficient of i-th order term of passive

trader.

After each exchange, the amount of money for active and passive traders can be obtained as

$$M_{S_i}^A = M_{S_{i-1}}^A - E_A + E_P \tag{6}$$

$$M_{S_i}^P = M_{S_{i-1}}^P - E_P + E_A \tag{7}$$

Where $M_{S_i}^A$ and $M_{S_i}^P$ are the amount of money owned by active and passive traders after money

exchange, respectively.

By substituting E_A and E_P in Equation (6) and (7), the relation between $M_{S_{i-1}}$ and M_{S_i} can be

written as

$$
\begin{aligned}
M_{S_i}^A = M_{S_{i-1}}^A \\
- \left[k_1^A M_{S_{i-1}}^A + k_2^A \left(M_{S_{i-1}}^A\right)^2 + k_3^A \left(M_{S_{i-1}}^A\right)^3 + \cdots \right] \\
+ \left[k_1^P M_{S_{i-1}}^P + k_2^P \left(M_{S_{i-1}}^P\right)^2 + k_3^P \left(M_{S_{i-1}}^P\right)^3 + \cdots \right]
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
M_{S_i}^P = M_{S_{i-1}}^P \\
- \left[k_1^P M_{S_{i-1}}^P + k_2^P \left(M_{S_{i-1}}^P\right)^2 + k_3^P \left(M_{S_{i-1}}^P\right)^3 + \cdots \right] \\
+ \left[k_1^A M_{S_{i-1}}^A + k_2^A \left(M_{S_{i-1}}^A\right)^2 + k_3^A \left(M_{S_{i-1}}^A\right)^3 + \cdots \right]
\end{aligned} \tag{9}
$$

For simplification, in the 1-D Model, the interaction between active and passive traders is

considered as linear. In other words, each trader exchanges a fixed fraction of what they own during

each exchange. Note that the assumption of linear interaction is plausible only for certain systems.

Even though the linear interaction mechanism is not always suitable for every real stochastic system

found in industrial or experimental processes, it represents the simplest mechanism to simulate a

money exchange process since it introduces a lower number of money exchange parameters and

therefore reduces the number of factors that impact simulation results. As for the real process, its

12

interaction mechanism needs to be adjusted and verified by further investigations and analysis. Thus, higher order money transfer processes are not considered:

$$k_i^A = k_i^P = 0, \quad \forall\, i \geq 2 \tag{10}$$

Finally, the relations between M_{S_i} and $M_{S_{i-1}}$ in an exchange event are given by

$$M_{S_i}^A = M_{S_{i-1}}^A - k_1^A M_{S_{i-1}}^A + k_1^P M_{S_{i-1}}^P \tag{11}$$

$$M_{S_i}^P = M_{S_{i-1}}^P - k_1^P M_{S_{i-1}}^P + k_1^A M_{S_{i-1}}^A \tag{12}$$

Rearranging equation (11) and (12),

$$M_{S_i}^A = \left(1 - k_1^A\right) M_{S_{i-1}}^A + k_1^P M_{S_{i-1}}^P \tag{13}$$

$$M_{S_i}^P = \left(1 - k_1^P\right) M_{S_{i-1}}^P + k_1^A M_{S_{i-1}}^A \tag{14}$$

In a stochastic simulation, the state of the 1-D Model evolves dynamically over time and the system keeps updating continuously due to the sequential money exchanges. T_{total} denotes the total simulation time. As the time step Δt is constant, the total number of exchange events is obtained from the termination time, i.e. number of exchange events is $T_{\text{total}}/\Delta t$.

By setting input variables and initial conditions, the money redistribution of the 1-D Model can be simulated as a sequence of discrete money exchange events. Figure 5 schematically shows the simulation process of the 1-D Model.
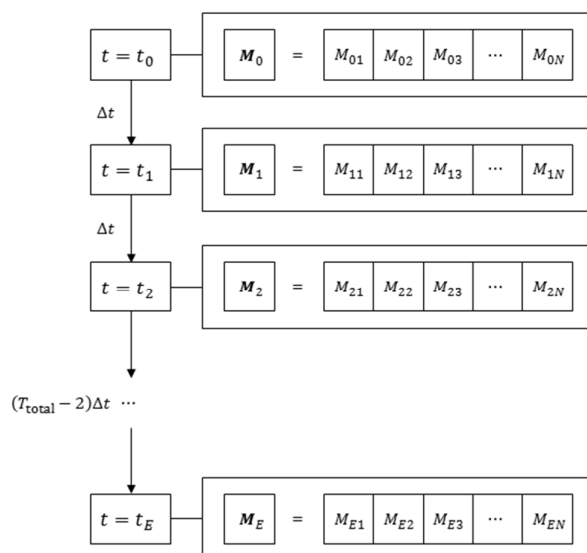
Figure 5. *Time evolution of trader sequence of the 1-D Model in stochastic simulation. The vector M_i represents the money state of the model at $t = t_i$. M_{ij} represents the amount of money held by the trader N_j at $t = t_i$. E represents the terminal time of the simulation at the end.*

According to the model and its interaction mechanism introduced above, the Gillespie SSA of the 1-D Model can be implemented and shown in Figure 6.
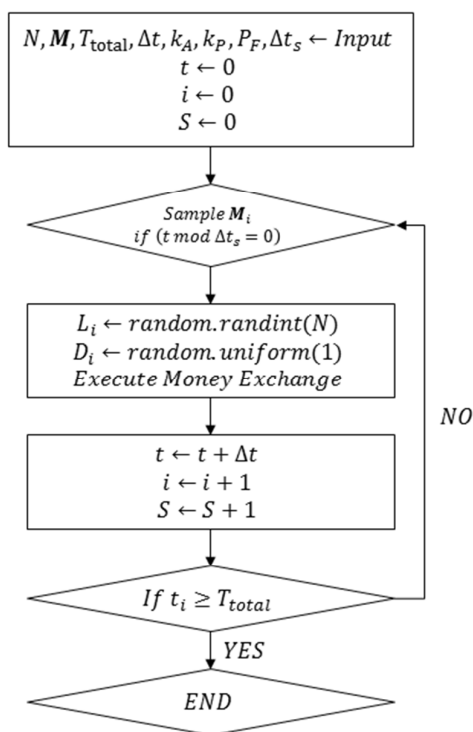
Figure 6. *Flow chart of the stochastic simulation algorithm of the 1-D Model.*

2.3. Identification of regression model

In this paper, the standard deviation of money and money distribution in the system are the selected response variables to investigate the system status changing in time. The former quantifies the amount of variation and dispersion of amount of money in the system and the latter presents the frequency of amount of money held by traders. A curve fitting analysis is applied to estimate the curves of standard deviation evolution and money distribution by looking for an analytical expression. Two or more candidate expressions may give similar goodness of fit for the same simulation realisation but with completely different complexities. Thus, it is crucial to balance the goodness of fit and complexity of the expression form. A sketch of the proposed model identification procedure, namely *two-layer model identification strategy*, for the representation of realisations from the 1-D Model stochastic simulation is shown in Figure 7. This strategy is applied

for describing the time evolution of standard deviation and probability density function of money

distribution in time through nonlinear regression models.

For the probability density function of money distribution, each candidate model has the

general form which can be fitted by an appropriate regression model.

$$y = f(N, \boldsymbol{M}_0, T_{\text{total}}, \Delta t, k_1^A, k_1^P, P_F, t) = f(\boldsymbol{\varphi}, t) \tag{15}$$

$$\hat{y} = g(\boldsymbol{\theta}, t) \tag{16}$$

In Equation (15), $\boldsymbol{\varphi} = [N, \boldsymbol{M}_0, T_{\text{total}}, \Delta t, k_1^A, k_1^P, P_F]^{\text{T}}$ is the N_φ-dimensional set of operating

variables used in the stochastic simulation and in Equation (16), $\boldsymbol{\theta}$ is an N_θ-dimensional set of

regression parameters. In the *first layer*, the PDF curves extracted at each sampling time point are

defined and fitted by candidate models in the form

$$\hat{y}|_{t=t_1} = g(\boldsymbol{\theta}_{t_1}, t_1)|_\varphi = g_1(\theta_{1,t1}, \theta_{2,t1}, \theta_{3,t1}, \cdots)|_\varphi$$

$$\hat{y}|_{t=t_2} = g(\boldsymbol{\theta}_{t_2}, t_2)|_\varphi = g_2(\theta_{1,t2}, \theta_{2,t2}, \theta_{3,t2}, \cdots)|_\varphi \tag{17}$$

$$\cdots$$

$$\hat{y}|_{t=t_n} = g(\boldsymbol{\theta}_{t_n}, t_n)|_\varphi = g_n(\theta_{1,tn}, \theta_{2,tn}, \theta_{3,tn}, \cdots)|_\varphi.$$

In the *second layer*, a nonlinear curve fitting is applied to find a parameterised function to estimate

how time-dependent parameters $\boldsymbol{\theta}_i$ in Equation (17) will evolve over processing time. Optimisation

can be carried out computationally to find a curve fitting function with the best trade-off between

Residual Sum of Squares (RSS) (model adequacy) and number of parameters (N_θ) (evaluating

model complexity) through the Akaike information criterion[18] for the evaluation of the minimum

Akaike information index $AIC = 2N_\theta - \log RSS$.

For the time evolution of standard deviation of money, Equation (15) still applies the

general form. However, the identifications in the first and second layer are of slightly difference. In

the *first layer*, the time-dependent STD curves are defined and fitted by the candidate models, in the

form

$$y|_{\boldsymbol{\varphi}_1} = f(\boldsymbol{\varphi_1}, t)$$

$$y|_{\boldsymbol{\varphi}_2} = f(\boldsymbol{\varphi_2}, t)$$

$$\dots \tag{18}$$

$$y|_{\boldsymbol{\varphi}_n} = f(\boldsymbol{\varphi_n}, t)$$

And,

$$\hat{y}|_{\boldsymbol{\varphi}_1} = g(\boldsymbol{\theta_1}, t)$$

$$\hat{y}|_{\boldsymbol{\varphi}_2} = g(\boldsymbol{\theta_2}, t)$$

$$\dots \tag{19}$$

$$\hat{y}|_{\boldsymbol{\varphi}_n} = g(\boldsymbol{\theta_n}, t)$$

In the *second layer*, a nonlinear curve fitting is applied to find a parameterised function to estimate how time-dependent parameters $\boldsymbol{\theta}_i$ in Equation (19) will evolve over processing time. Same optimisation on RSS and N_θ will be carried out.

In general, the identification of regression model on the standard deviation of money requires a number of simulation results with different input scenarios $\boldsymbol{\varphi_1}, \boldsymbol{\varphi_2}, \dots, \boldsymbol{\varphi_n}$ to estimate the functions of $\boldsymbol{\theta}$, while that on the money distribution requires only one simulation results with specific input variables and condition $\boldsymbol{\varphi}$ to estimate the time evolution of $\boldsymbol{\theta}$.
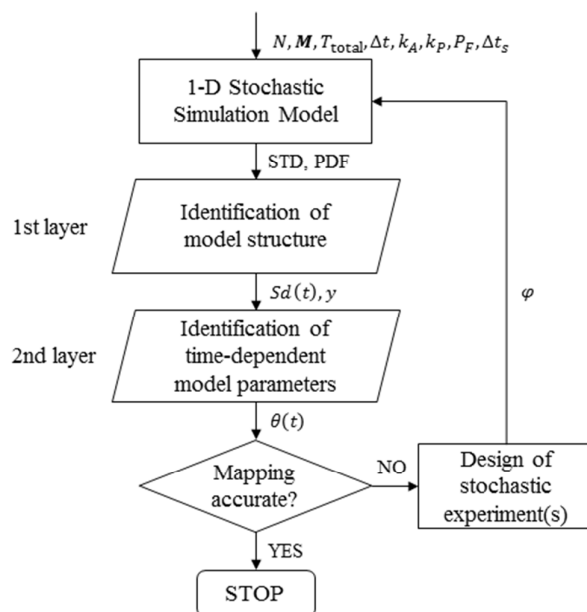
Figure 7. *Sketch of the two-layer model identification strategy for stochastic model simulation of*

TTEP problem.

3. Results and Discussion

Both the analyses of standard deviation evolution and of money redistribution consist of two

sections, parametric study and regression model identification. For clarity, the fixed values of model

input variables, φ, and initial conditions are given at the beginning of section 3.1 The values of

specific variables studied in the parametric study will be explicitly mentioned in figures or captions.

3.1. Simulation realisations from the 1-D Model

The preliminary stochastic simulations on the 1-D Model are presented by using fixed

values of input variables φ which are listed in Table 1. Five realisations of SSA using the same

input conditions in Table 1 are plotted with different colours in Figure 8(left). Each curve represents

the temporal evolution of the standard deviation of the money distribution function. Variability is

the natural consequence of the stochasticity of the process.[8] Thus, fluctuations among dynamic

trajectories starting from the same conditions are to be expected. However, as will be shown below, this variability is of limited importance compared to the effect of changing the values of input conditions. The subsequent sections will present how changing the values of input conditions affects the realisation of stochastic simulation. Figure 8(right) presents the probability density function (PDF) of the money distribution in the 1-D Model at four simulation time points and the last picture also illustrates the final money distribution of the system at the termination time $T_{total} = 1 \times 10^4$. The computational domain, $[M_{min}, M_{max}]$ (i.e. [0,3000] in Figure 8(right)) is divided into $K = 60$ compartments of width $w = 50$. The height of the bar in each compartment represents how many traders fall into the corresponding range of money.

Table 1. *Fixed values of model input variables and initial conditions for the stochastic simulation in the 1-D Model.*

Variable/condition	Symbol	Fixed value
(1) Number of traders	N	1000
(2) Initial amount of money for each trader (i.e. $M_{01} = M_{02} = M_{03} = \cdots = M_{0N}$)	$\boldsymbol{M}_0$	1000
(3) First order money transfer coefficient of active trader	k_1^A	0.5
(4) First order money transfer coefficient of passive trader	k_1^P	0.25
(5) Termination time	T_{total}	1×10^4
(6) Exchange event time gap	Δt	1
(7) Exchange direction propensity	P_F	0.5
(8) Sampling time gap	Δt_s	100
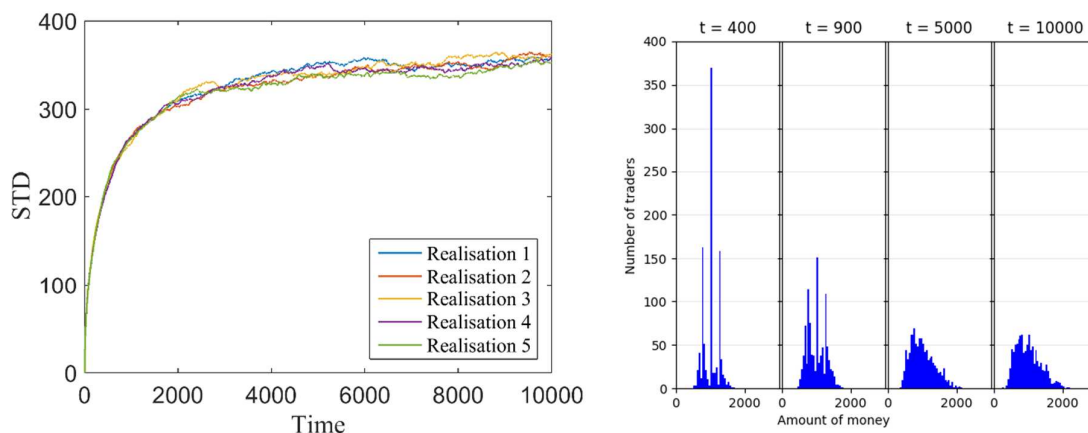
Figure 8. *Left: Five realisations of SSA for the 1-D Model with input variables and condition in Table 1, each colour representing one individual stochastic simulation; right: One realisation of money distribution change over simulation time with input variables and condition in Table 1.*

3.2. Parametric study

3.2.1. Number of traders and exchange direction propensity

The size of the population of traders involved in the exchange system has a significant influence on the time evolution of the standard deviation. Figure 9 graphically illustrates the plots of the standard deviation on both symmetric ($P_F = 0.5$) and asymmetric ($P_F \neq 0.5$) systems. In Figure 9(a) the blue curve with $N = 1000$ reaches steady fluctuation region within a short processing time and the green line with $N = 20000$ is the last one to leave launch region. Launch region is an initial transient region in which the curve has a clear increasing trend. It takes longer time for the system with larger trader populations to equilibrate. This is obvious because the system with larger population size requires higher number of money exchange events, which can affect the majority of the traders in the system, to approach steady fluctuation region. In the steady fluctuation region, which comes after launch region and shows the fluctuation around certain average value, the blue curve with the smallest population of traders has the highest degree of fluctuation while the green curve evolves smoothly. This is because when the system has a relatively small population, every single money exchange event would be an important contribution to the system so that the state of

the system would give a fast reaction within a short period of processing time and present a significant change. Figure 9(b) illustrates the time evolution of standard deviation in the asymmetric system. The blue curve with $N = 1000$ comprises three different regions clearly (i.e. launch, steady fluctuation and reacceleration), within a process time $T_{\text{total}} = 3 \times 10^5$. Reacceleration region appears when the system has a small trader population and/or unbalanced exchange direction propensity. In this region, the curve does not fluctuate around the average value any longer but starts showing an increasing trend again (see blue curve after time $t = 5 \times 10^4$ in Figure 9(b)). In the reacceleration region, the trajectory slowly deviates from the original range of fluctuation. This behaviour can also be found for $N = 2000$.

The appearance of a reacceleration region is due to the overall difference between the money transferred in forward and backward exchanges and, consequently, money accumulation in the asymmetric systems. For the systems with small trader populations and $P_F \neq 0.5$, money accumulation easily happens and presents a growing trend of standard deviation of money in the system as the processing time of stochastic simulation increases, which produces the reacceleration region. A comparison plot on the system with $N = 1000$ and $P_F = 0.25$ has been presented in Figure 9(c). The blue curve presents the time evolution of standard deviation with respect to all traders in the system while the orange curve presents that with respect to the traders from 101 to 900 only, i.e. exclusive of the first and last 100 traders. As expected, the time evolution of standard deviation of these 900 traders in the bulk still stays in the steady fluctuation region. According to the simulation result, the average amount of money on the 900 traders in the bulk fluctuates between 998.4 and 1003.7, extremely close to the initial mean of 1000, which means the asymmetric exchange direction propensity would have little effect on the traders in the bulk within the process time $T_{\text{total}} = 3 \times 10^5$. However, this effect gradually spreads to all the trades as process time increases. On the other hand, a clear money accumulation is observed in the simulation result. The average amount of money held by the first 100 traders gradually increases over time while that

held by the last 100 traders gradually decreases. At the termination, the largest amount of money

held by trader No. 4 is 9610.8 and the smallest one held by trader No. 1000 is 1.92. For asymmetric

systems, the trader holding the largest amount of money at the termination is not always the same

one due to the stochasticity in the simulations but the trader holding the largest amount of money

can usually be observed at the boundary region in model, so as the trader holding the smallest

amount of money.
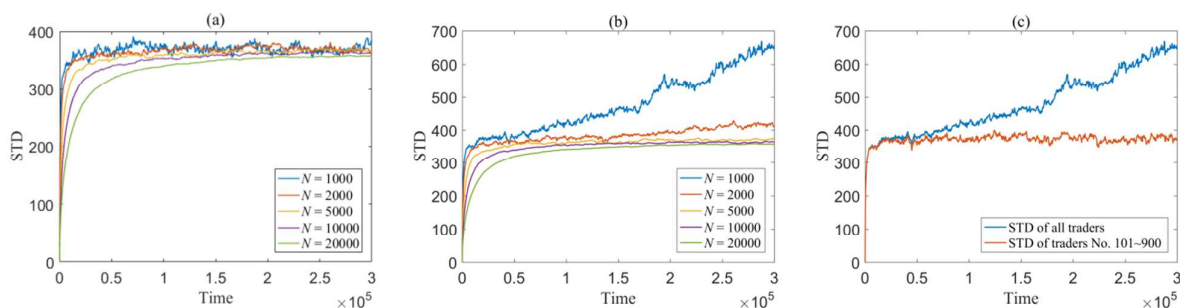
Figure 9. *Stochastic simulation results of the 1-D Model with different N ($\mathbf{M}_0$, k_1^A, k_1^P, Δt at the fixed values in Table 1, $T_{\text{total}} = 3 \times 10^5$, $\Delta t_s = 300$). (a) Symmetric system for $P_F = 0.5$; (b) Asymmetric system for $P_F = 0.25$; (c) The comparison between the time evolutions of STD with respect to all and partial traders in the system with $N = 1000$ and $P_F = 0.25$.*

A further illustration on the effect of exchange direction propensity P_F is presented in Figure

10. As all traders have the same initial amount of money at $t = 0$, with all other input variables and

condition fixed, the influence of exchange direction propensity P_F is solely based on the interval

$P_F \in [0, 0.5]$, i.e. $P_F = 0.3$ and $P_F = 0.7$ basically give similar model behaviours. The importance

of exchange direction propensity can be neglected in the initial period of simulation process and this

period changes according to system configuration, especially population size. In Figure 10(a),

$N = 1000$, the curves separate approximately at $t = 2 \times 10^4$ and the stochastic simulations with

smaller P_F gives a faster money accumulation in the model. For systems with larger populations, i.e.

$N = 5000$, higher number of model updates is required for the STD curve to leave the steady

fluctuation region. According to the results of simulation (not shown here), for the system with

$N = 5000$, a clear separation of the curves occurs at about $t = 3 \times 10^5$, which is of great

difference from what is observed in the system with $N = 1000$. Money accumulation is inevitable in systems with asymmetric exchange direction propensity and its appearance is a matter of time for systems with large population size. The traders at two ends of the model sequence, i.e. first and last trader, have only one exchange direction. When exchange direction propensity is strongly asymmetric, especially $P_F = 0$, the money accumulated on the boundary is extremely hard to be transferred back to the traders in the bulk.
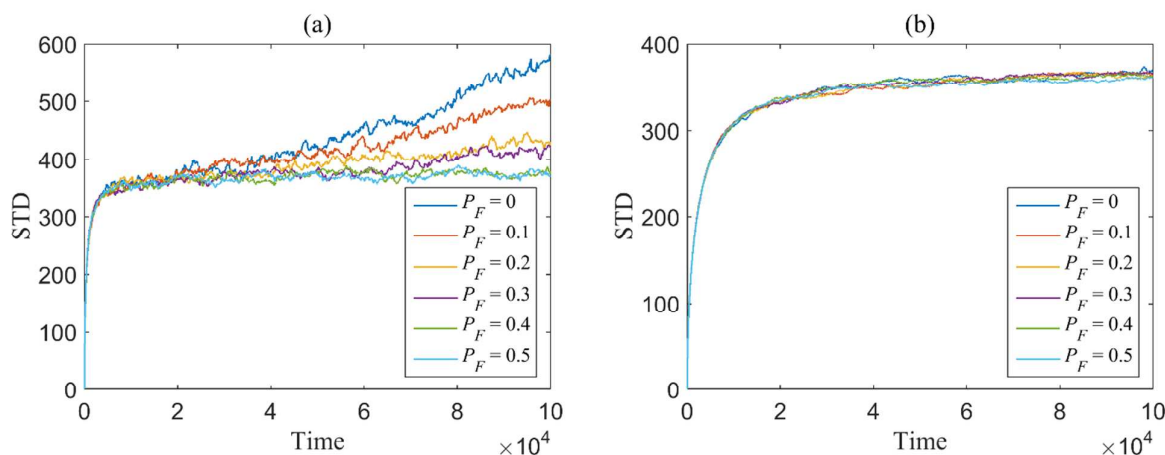
Figure 10. *Stochastic simulation results of both (a) small and (b) large population 1-D Model with different exchange direction propensity ($\boldsymbol{M}_0$, k_1^A, k_1^P, Δt at the fixed values in Table 1). (a) $N = 1000$, $T_{\text{total}} = 1 \times 10^5$, $\Delta t_s = 100$; (b) $N = 5000$, $T_{\text{total}} = 1 \times 10^5$, $\Delta t_s = 100$.*

Figure 11 illustrates the PDF of the money distribution after long simulation time $T_{\text{total}} = 3 \times 10^5$ in both (a) extremely asymmetric and (b) symmetric situations and it visualises the money accumulation from another perspective. According to the figure, the money distributions in two situations resemble each other in bulk, i.e. domain [250,2500]. Most of the traders drop in the compartments of [250,2500] with few outliers. The obvious difference between 11(a) and 11(b) is the number of traders in the compartments from $M = 0$ to $M = 250$, especially for the first compartment of [0,50]. In Figure 11, two sub-figures use the same axis scale for amount of money [0,3000]. The total numbers of traders included in the histograms are 995 and 1000, respectively. The traders with amount of money more than 3000 are not visualised in the histograms. As

expected, the asymmetric P_F in system, causing an unbalanced money transfer, gives some traders

big amount of money after simulation and on the other hand, results in a few traders with almost no

money.
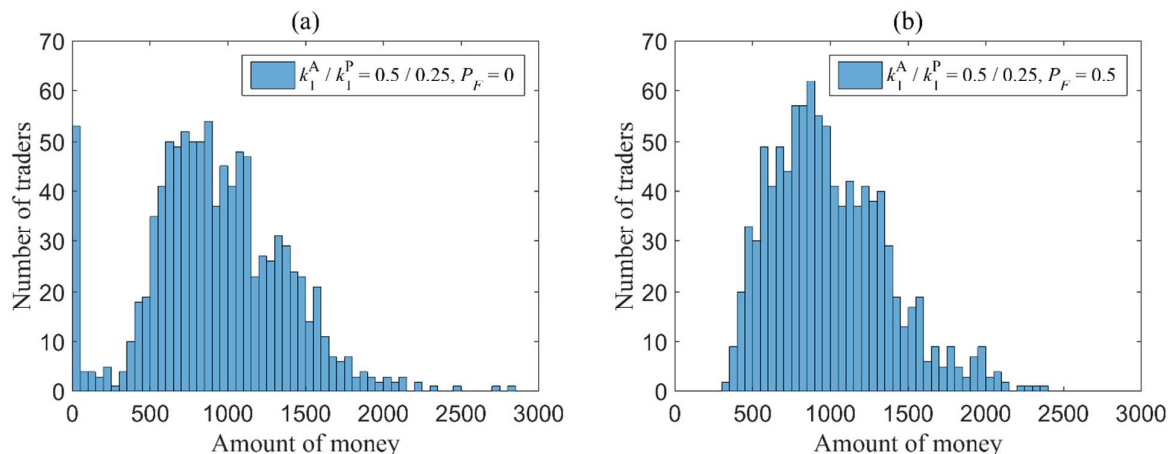
Figure 11. *Stochastic simulation results of the 1-D Model with different exchange direction propensity (N, $\boldsymbol{M}_0$, k_1^A, k_1^P, Δt for fixed values in Table 1, $T_{\text{total}} = 3 \times 10^5$). (a) $P_F = 0$; (b) $P_F = 0.5$*

3.2.2. Initial amount of money

Figure 12(a) are symmetric systems with $P_F = 0.5$. The trader populations for the

simulations are $N = 1000$. In the figure, the initial amount of money for the traders mainly

determines the level of the standard deviation in the steady fluctuation region. From the view of

mathematical point, when two systems with the same population size but different initial amount of

money, M_1 and M_2, undergo the computational simulations with exactly the same money exchange

process, i.e. each pair of active and passive trader from one system matches the other, as the

standard deviation σ satisfies,

$$\sigma^2 \propto \sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{20}$$

And,

$$\frac{x_{i_1}}{x_{i_2}} = \frac{\bar{x}_1}{\bar{x}_2} = \frac{M_1}{M_2} \tag{21}$$

The standard deviation ratio of two systems at one time would be,

$$\frac{\sigma_1(t)}{\sigma_2(t)} = \frac{M_1}{M_2} \tag{22}$$

According to Equation (22), as expected, the initial amount of money linearly affects the scale of standard deviation of the system at not only the steady fluctuation region but also other regions. The plot of standard deviation at steady fluctuation region to initial amount of money is presented in Figure 12(b).
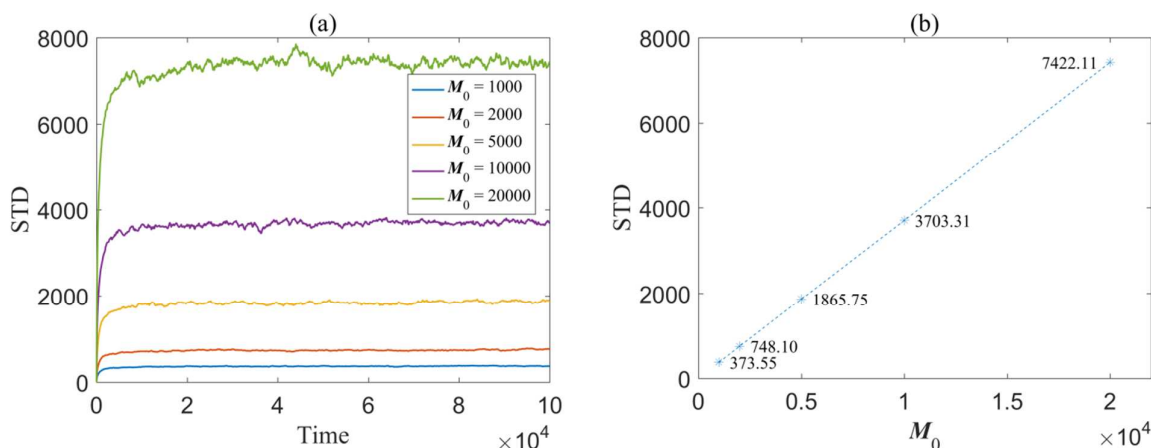
Figure 12. *Stochastic simulation results of the 1-D Model with different initial amount of money for the traders (k_1^A, k_1^P, Δt of the fixed values in Table 1, $N = 1000$, $P_F = 0.5$, $T_{total} = 1 \times 10^5$, $\Delta t_s = 100$). (a) Time evolution of standard deviation with different M_0; (b) STD $-$ M_0 plot.*

3.2.3. Money transfer coefficient

Coefficients k_1^A and k_1^P represent the fraction of money extracted from active and passive traders in each exchange event, respectively. In symmetric systems they are equivalent and have the same physical meaning. Hence the stochastic simulations results for $k_1^A = 0.5$, $k_1^P = 0.25$ (as shown in Table 1) and $k_1^A = 0.25$, $k_1^P = 0.5$, with other input conditions set the same, are the same.

Figure 13(a) implies that when considering a given k_1^A, for example $k_1^A = 0.5$, and two k_1^P

are symmetric with respect to k_1^A (i.e. $k_1^P = [0.2, \ 0.8]$, $k_1^P = [0.3, \ 0.7]$, and $k_1^P = [0.4, \ 0.6]$), their

realisations will be the similar. And the similarity increases as the values of k_1^P pairs approach to

0.5. As illustrated by Figure 13(b), the level of standard deviation curve of money in steady

fluctuation region is affected by the difference between k_1^A and k_1^P. The values of k_1^P for the curves

in Figure 13(b) from top to bottom are $k_1^P = 1$, 0.7, 0, 0.4 and 0.3. The differences between k_1^A

and k_1^P from top to bottom are $|k_1^A - k_1^P| = 0.75$, 0.45, 0.25, 0.15 and 0.05. According to Figures

13(c) and 13(d), when $k_1^A = 0.3$, two curves with different k_1^P, even being symmetric with respect

to k_1^A, do not give the similar realisations of stochastic simulation. The results show that in each

pair of k_1^P (being symmetric with respect to k_1^A), when $k_1^A < 0.5$, the larger k_1^P would give a smaller

standard deviation in steady fluctuation region (see Figure 13(c)). On the contrary, when $k_1^A > 0.5$,

the smaller k_1^P in each pair of k_1^P (being symmetric with respect to k_1^A) would give a smaller

standard deviation in steady fluctuation region (see Figure 13(d)). In general, three rules can be

summarised to identify the level of standard deviation in steady fluctuation region with fixed k_1^A

and different k_1^P.

    i.   The larger difference between k_1^A and k_1^P would give higher standard deviation in steady

                fluctuation region. This big difference will cause an unbalance money transfer between

                active and passive traders involved in each exchange event. After long simulation time,

                some traders will be holding big amount of money while some other traders will be

                having little money. This huge money difference will actually increase the standard

                deviation of money in the system.

    ii.   When $k_1^A = 0.5$, two k_1^P being symmetric about k_1^A would give similar standard

                deviation evolution behaviours and the similarity increases with the difference between

                two k_1^P decreasing. When two k_1^P are symmetric about $k_1^A = 0.5$, they would present the

                same coefficient difference $|k_1^A - k_1^P|$, and the values of money transfer from active and

passive traders at each exchange event would be of little discrepancy, resulting in the similarity in standard deviation. When $k_1^A \neq 0.5$ and two k_1^P are symmetric with respect to k_1^A, the smaller standard deviation in steady fluctuation region would be given by the k_1^P that is closer to 0.5. In this situation, the similarity of two standard deviation evolutions decreases since the impact of coefficient difference $|k_1^A - k_1^P|$ becomes complicated in the system.

Moreover, money transfer coefficients k_1^A and k_1^P have a clear influence on the nature of the PDF Figure 14 presents four realisations of stochastic simulations with different couples of k_1^A and k_1^P, giving four different PDFs on the computational domain of [0,6000].. In Figure 14, the forms of PDF of money distributions are (a) exponential, (b) transition between log-normal and exponential, (c) normal distribution with skewness or log-normal and (d) normal distribution. The form of money distribution at terminal simulation time is significantly affected by the values of k_1^A and k_1^P. The discrepancy between the values of k_1^A and k_1^P would make contribution to the money reallocation to the traders in the system and eventually, the different forms of terminal money distribution are given.

By varying the values of k_1^A and k_1^P from 0 to 1, respectively, with step change equal to 0.05, the k_1^A-k_1^P map is produced in Figure 15, showing the form of the PDF at steady state in the 1-D Model. The k_1^A-k_1^P map presents a clearly symmetric behaviour. This is because with all other input variables and condition fixed, the values of money transfer coefficients for active and passive traders can be swapped and they will give the same simulation behaviour of the model with respect to the aggregate descriptors considered here, i.e. $k_1^A = 0.5$, $k_1^P = 0.25$ and $k_1^A = 0.25$, $k_1^P = 0.5$ are equivalent for the model, The step change of k_1^A and k_1^P applied to produce the map in Figure 15 is 0.05. A smaller step change would give a map with higher precision, especially for the points on the boundaries. Actual regressions and goodness of fit are required, especially to determine the behaviour at the boundaries of the different regions since it is very hard to identify the form of PDF

solely by vision. The map presented in Figure 15 is obtained at specific input condition. From author's point of view, however, this map can still be used to determine the final form of the PDF at steady state of the 1-D Model with other input conditions. This is because, as discussed before, the population of traders and initial amount of money mainly impact the smoothness and oscillation of the curves or plots and the interaction mechanism is the key force that dynamically drives the system evolution within simulation time.
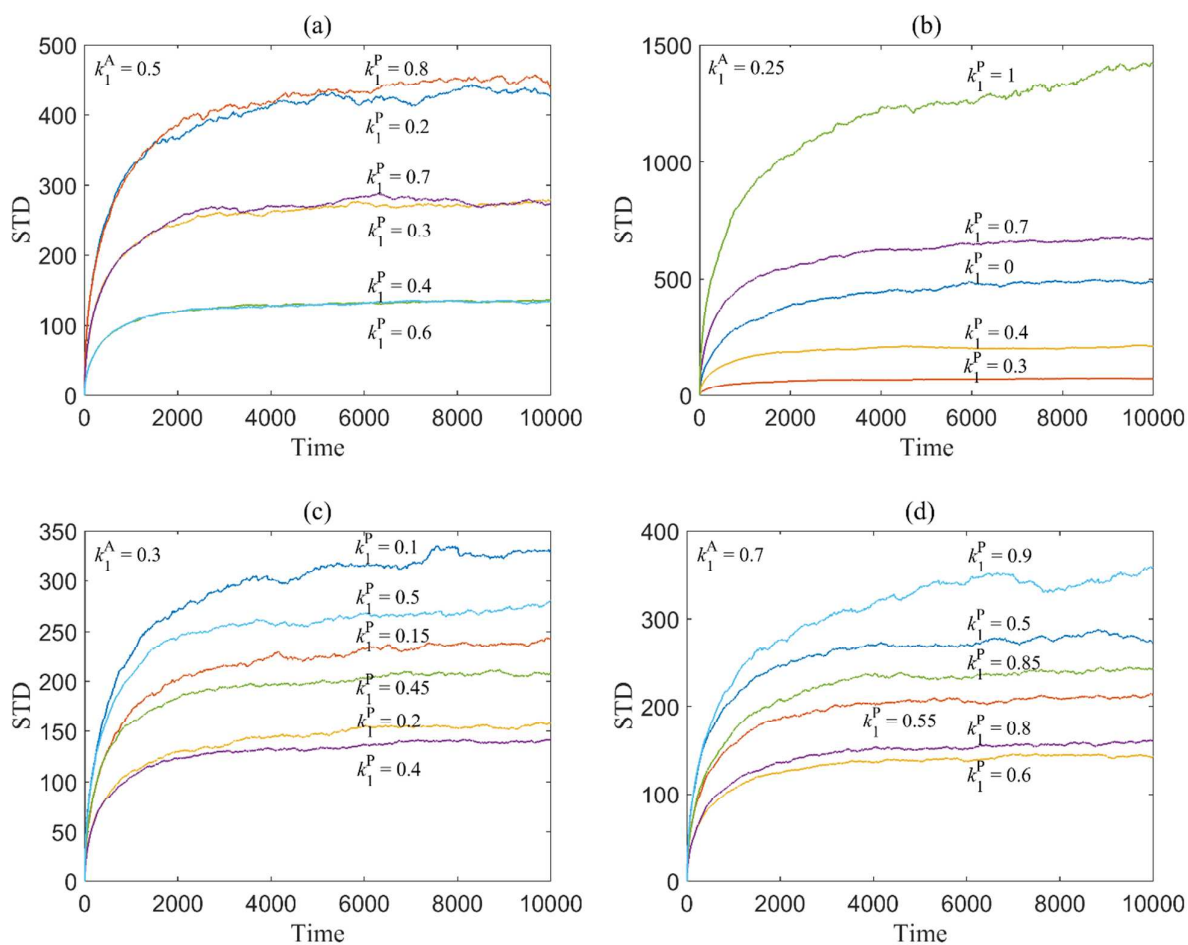
Figure 13. *Stochastic simulation results of the 1-D Model with different money transfer coefficients of active and passive traders (N, M_0, T_{total}, Δt, P_F of the fixed values in Table 1, $\Delta t_s = 10$, see legends for k_1^P). (a) $k_1^A = 0.5$; (b) $k_1^A = 0.25$; (c) $k_1^A = 0.3$; (d) $k_1^A = 0.7$.*
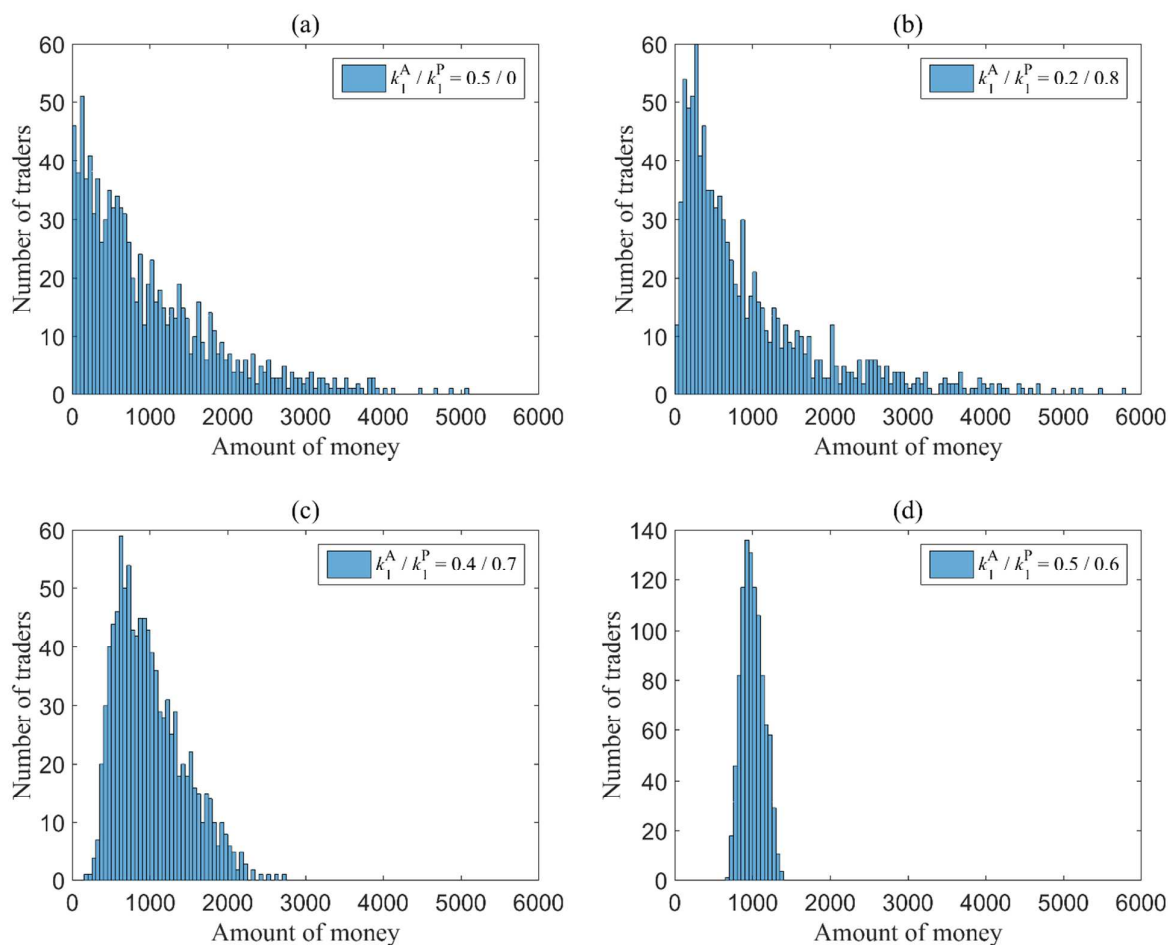
Figure 14. *Stochastic simulation results of the 1-D Model with different money transfer coefficients (N, $\boldsymbol{M}_0$, Δt, P_F for fixed values in Table 1, $T_{\text{total}} = 1 \times 10^5$). (a) $k_1^A = 0.5$, $k_1^P = 0$; (b) $k_1^A = 0.2$, $k_1^P = 0.8$; (c) $k_1^A = 0.4$, $k_1^P = 0.7$; (d) $k_1^A = 0.5$, $k_1^P = 0.6$.*
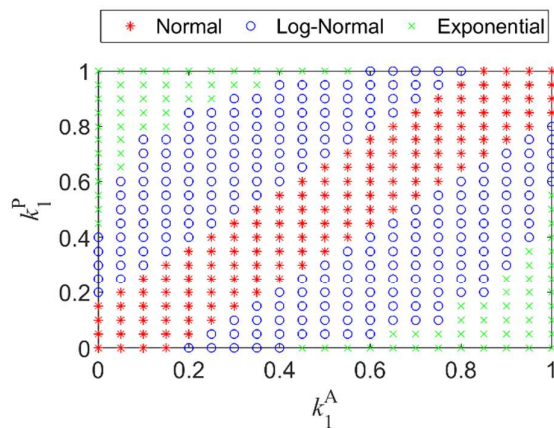
Figure 15. *The k_1^A-k_1^P map of the 1-D model with $N = 1000$, $\boldsymbol{M}_0 = 1000$, $T_{\text{total}} = 1 \times 10^5$, $\Delta t = 1$ and $P_F = 0.5$.*

3.3. Regression model identification

In this section, the simulation results are investigated to build regression models. It is important to select suitable input conditions when simulating the model as inappropriate input conditions may cause issues in the regression, listing below:

- The small trader population in the system may bring strong curve oscillation in steady fluctuation region, see Figure 9, which does not benefit the curve fitting. Improper large trader population needs significantly large process time to get the system adequately simulated, making the simulation computationally expensive.

- In this section, the model behaviour after steady fluctuation region, i.e. reacceleration region, is not considered, so that P_F is fixed at 0.5 to avoid money accumulation in the simulation.

- If the simulation time was too long simulation would not be computationally efficient because when the system reaches steady state the further simulation does not significantly affect the system. If the simulation time was too short the traders in the model would not be adequately simulated.

Considering the issues above, the 1-D Model is simulated with input conditions in Table 1 with $N = 5000$ and $T_{\text{total}} = 1 \times 10^5$. The data of this realisation contains 1,001 points sampled from $t = 0$ to $t = 1 \times 10^5$ with sampling time step $\Delta t_s = 100$. In the first layer of the model identification strategy, a nonlinear curve fitting of the STD based on least square method is carried out on simulation realisation. Table 2 summaries the results of nine candidate functions used for curve fitting on the data of realisation.

Table 2. *Summary of the curve fittings on the time evolution of STD. Relevant values of RSS, COD, N_θ and AIC are shown in boldface.*

Function	Expression	RSS	COD	N_θ	AIC
Asymptotic	$y = a - bc^x$	75633	0.954	3	17.23
Exponential	$y = y_0 + Ae^{R_0 x}$	75633	0.954	3	17.23
Hill	$y = V_{max}\dfrac{x^n}{k^n + x^n}$	6194	0.996	3	14.73
Logistic	$y = \dfrac{A_1 - A_2}{1 + \left(\dfrac{x}{x_0}\right)^p} + A_2$	**5229**	**0.997**	4	16.56
Michaelis-Menten	$y = \dfrac{V_{max}x}{K_m + x}$	14401	0.991	**2**	**13.58**
Bradley	$y = a\ln(-b\ln x)$	111439	0.927	2	15.62
Poly	$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^{n-1}$	18889	0.989	10	29.85
Langmuir EXT	$y = \dfrac{abx^{1-c}}{1 + bx^{1-c}}$	6194	0.996	3	14.73
Nelder	$y = \dfrac{x + a}{b_0 + b_1(x + a) + b_2(x + a)^2}$	5298	0.997	4	16.58

All the curve fittings presented have converged and reached the Chi-Sqr tolerance value of $\varepsilon_{\chi^2} = 1 \times 10^{-9}$. According to Table 2, estimating functions with relatively complicated expressions return the best coefficients of determination (COD). However, Michaelis-Menten function presents the lowest value for AIC, offering a good curve fitting on the simulation results (COD = 0.991) with the lowest model complexity, and is consequently selected for the second layer, where parameters V_{max} and K_m are studied by using curve fitting analysis on a number of stochastic simulation realisations with different values in $\boldsymbol{\varphi}$. The investigation mainly focuses on the crucial variables which have significant impacts on simulation results, i.e. N, $\boldsymbol{M}_0$, k_1^A and k_1^P. Results are given in Figure 16 to 18.

It is worth noting that, in a Michaelis-Menten function, V_{max} (part of numerator) positively influences the function value while K_m (part of denominator) correlates inversely with the function

value. Larger populations of traders increase the value of K_m and, on the contrary, decrease the value of V_{max}. According to Michaelis-Menten expression, when x is small (x can be substituted by simulation time t) the y value is determined together by V_{max} and K_m and produces several launch region forms of STD curve (Figure 9). When x is large, the y value is mainly determined by V_{max} while K_m has little contributions to that. That is, when the simulation time increases, V_{max} becomes much more important than K_m and the standard deviation approximates to the value of V_{max}.

In Figure 16, both V_{max} and K_m show a linear dependency on the number of traders N. The V_{max} value shows slight change from 370 to 366 as trader population increases from 2000 to 20000. It implies that the trader population has slight impact on the standard deviation at steady state, and this is consistent with the fact that, in Figure 9, the standard deviation curves with different trader population have the similar steady fluctuation region. In Figure 17, the linear fitting is excellent on $V_{max} - M_0$ but poor on $K_m - M_0$. The goodness of $V_{max} - M_0$ is consistent with Equation (22) that initial amount of money affects the scale of standard deviation.
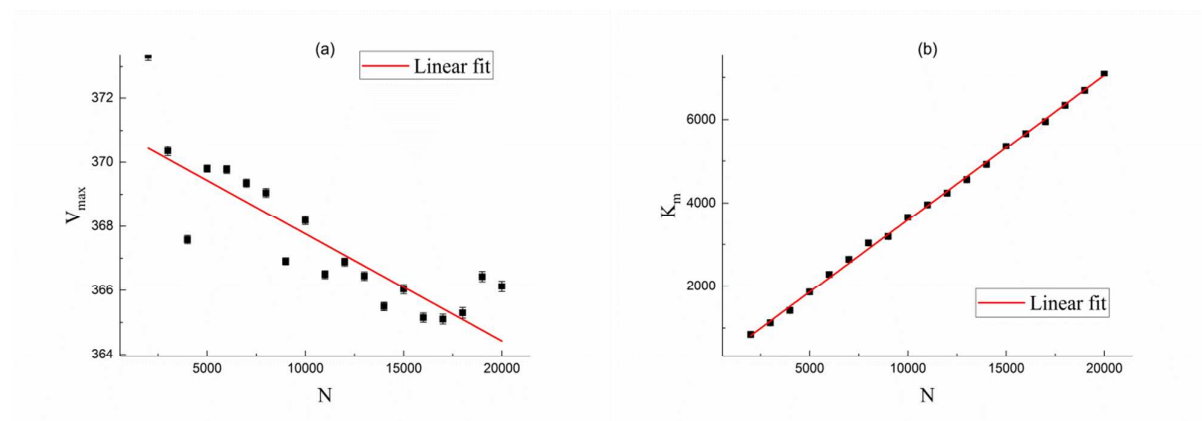
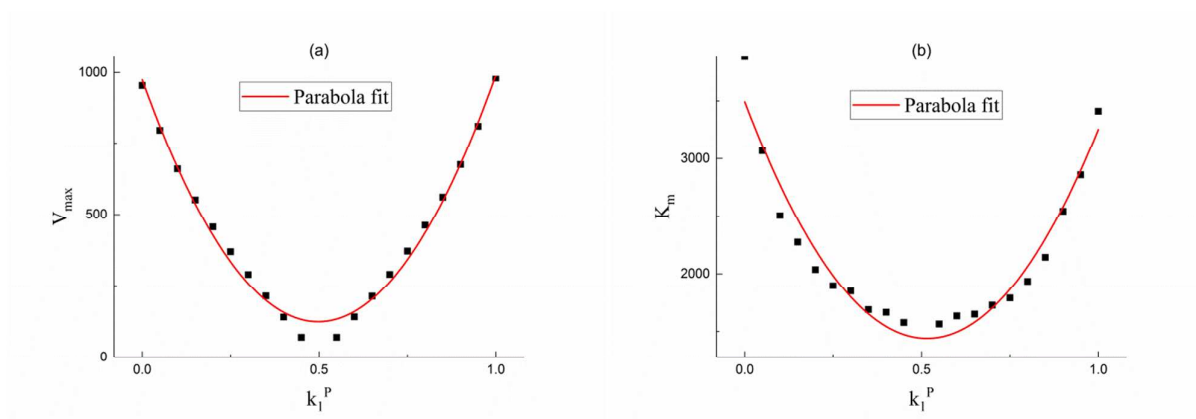

Figure 16. *Curve fitting analysis on the impact of N on STD function for $\mathbf{M}_0 = 1000$, $k_1^A = 0.5$, $k_1^P = 0.25$, $P_F = 0.5$ and $T_{\text{total}} = 3 \times 10^5$. (a) V_{max} vs N; (b) K_m vs N.*
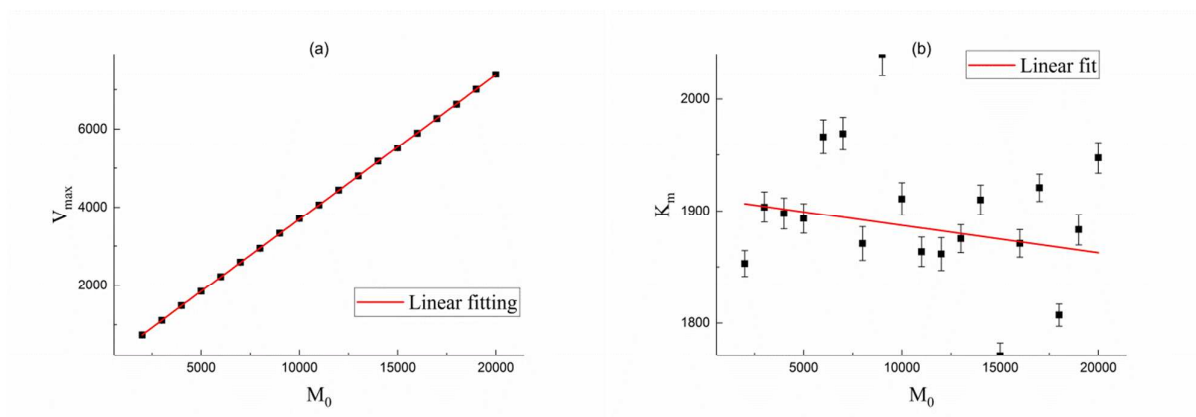
Figure 17. *Curve fitting analysis on the impact of M_0 on STD function for $N = 5000$, $k_1^A = 0.5$,*

$k_1^P = 0.25$, $P_F = 0.5$ and $T_{\text{total}} = 3 \times 10^5$. *(a) V_{max} vs M_0; (b) K_m vs M_0.*

**Figure 18.** *Curve fitting analysis on the impact of k_1^P on STD function for $N = 5000$, $\boldsymbol{M}_0 = 1000$,*

$k_1^A = 0.5$, $P_F = 0.5$ and $T_{\text{total}} = 3 \times 10^5$. *(a) V_{max} vs k_1^P in Poly fit; (b) K_m vs k_1^P in Poly fit.*

Figures 18 shows how the values of V_{max} and K_m change with different values of k_1^P with fixed

$k_1^A = 0.5$. It is found in Figure 18(a) that the values of V_{max} are symmetric about $k_1^P = 0.5$ and the

small values of V_{max} are given by the simulation in which k_1^P is near 0.5. This symmetry is

consistent with the results of the parametric study on k_1^A and k_1^P given in Section 3.2.3, see Figure

13(a). Figure 18(b) presents an approximate symmetry about $k_1^P = 0.5$. This symmetry ensures the

similar evolutions of standard deviation in launch region.

By simulating the 1-D Model with input conditions in Table 1 with $k_1^P = 0.45$ and $T_{\text{total}} = 2500$, the PDF of money redistribution is obtained for further investigation by using the two-layer model identification strategy. According to the $k_1{}^A$-$k_1{}^P$ map, the final form of PDF would be the normal distribution. The Gaussian function shown below is applied in the first-layer curve fitting for the plots extracted at each sampling time.

$$y = y_0^G + A^G e^{-\frac{(x-x_c^G)^2}{2w^{G^2}}} \tag{23}$$

The 1-D Model was simulated for $R = 100$ times to obtain 100 R-by-K arrays of different time evolutions of money distribution. K denotes the number of compartments in the money domain. Then, the average time evolution of money distribution, "average array", is derived by 100 R-by-K arrays. The samples used for curve fitting are extracted from the "average array". In the first layer, 26 samples from $t = 0$ to 2500 are non-linearly fitted by the Gaussian function to obtain the time-dependent parameters y_0^G, x_c^G, w^G and A^G. In the second layer, these parameters are non-linearly fitted again by several candidate functions, i.e. asymptotic function, logistic function. Figure 19 graphically illustrates the results of curve fitting on parameters in Gaussian function. No functions are used to fit the parameter y_0^G because (i) it is hard to find a function with simple expression to precisely describe the plot of the y_0^G shown in the figure and (ii) according to the scale of y-axis, the values of y_0^G are quite close to zero which is very small compared to the values of other parameters (i.e. x_c^G, w^G and A^G). Consequently, the curve fitting on parameter y_0^G is not necessary and parameter y_0^G in Gaussian function can be neglected. However, a rigorous analysis via confidence interval is required to further justify the neglecting procedure. The other sub-figures present the good fit on the estimated data obtained in the first-layer with corresponding functions.
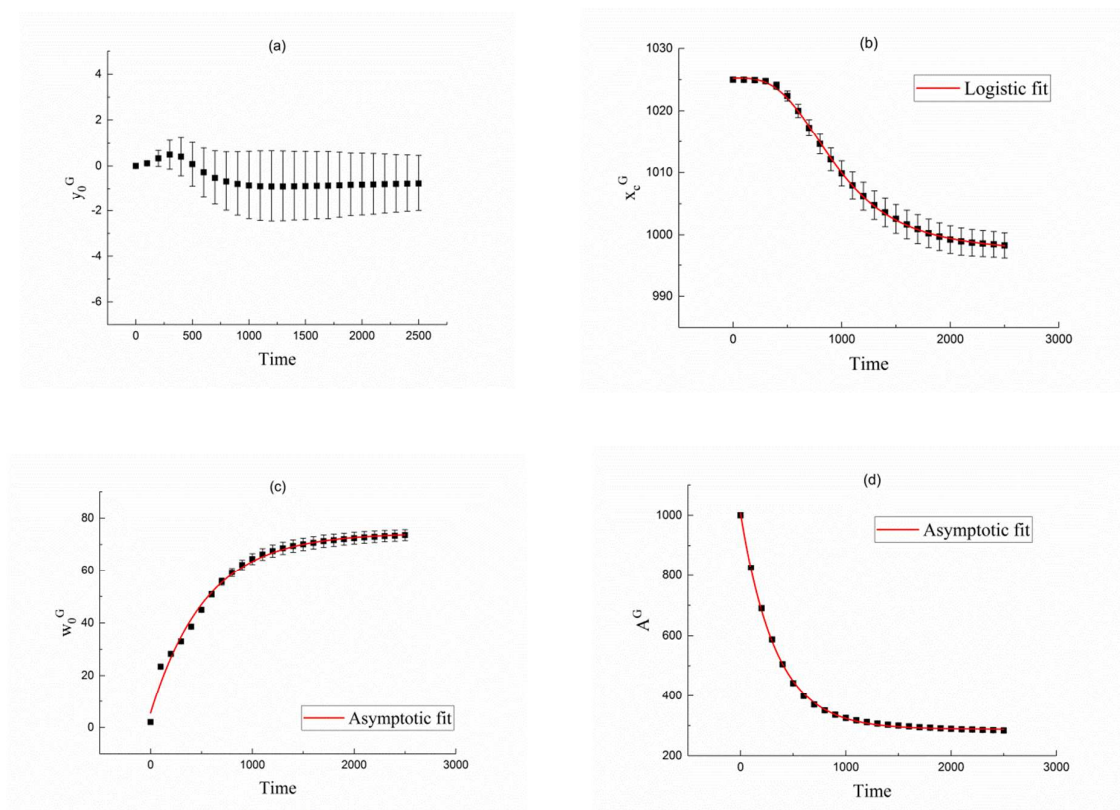
Figure 19. *Second-layer curve fitting on the time-dependent parameters (b)* x_c^G, *(c)* w^G *and (d)* A^G

in Gaussian function with input variables and conditions in Table.

The estimated sample is averaged from 100 simulation runs. The expressions determined in

the second layer for $x_c^G(t)$, $w^G(t)$ and $A^G(t)$ are asymptotic. As simulation time t increases the

expression terms containing independent variable t can be eliminated, showing,

$$\lim_{t \to \infty} x_c^G(t) = x_{c0}^G \tag{24}$$

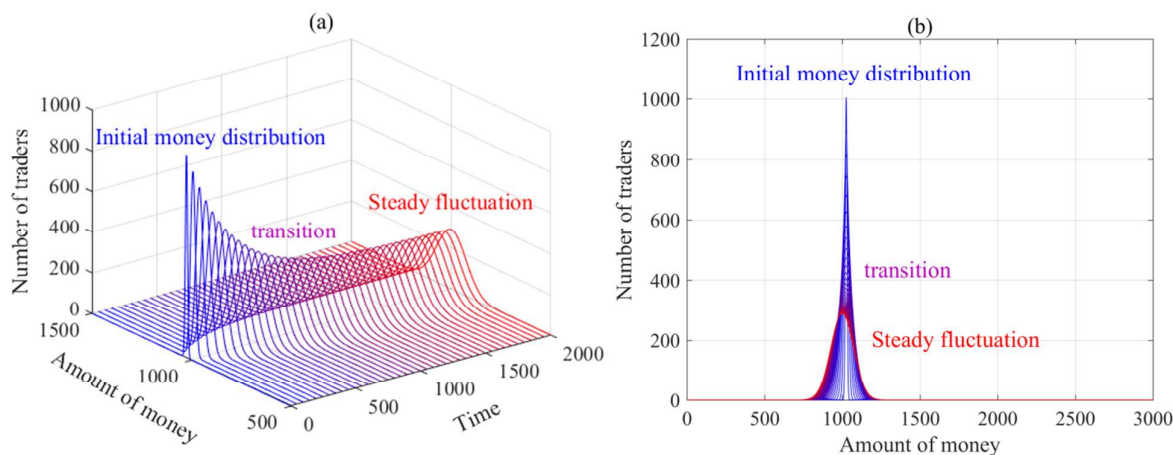$$\lim_{t \to \infty} w^G(t) = w_0^G \tag{25}$$

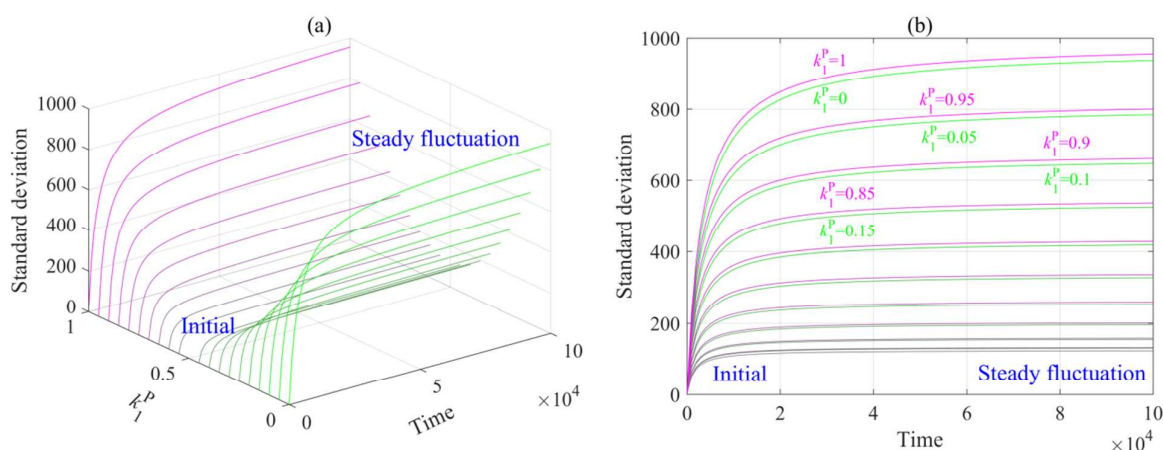$$\lim_{t \to \infty} A^G(t) = A_0^G \tag{26}$$

Then the Equation (23) can be written as,

$$y = A_0^G e^{-\frac{(x - x_{c0}^G)^2}{2 w_0^{G^2}}} \tag{27}$$

Here x is the amount of money of the traders.

Figure 20. *3D plot of the time evolution of money redistribution in the 1-D Model based on the estimated Gaussian function obtained by using two-layer model identification strategy. (a) Normal view; (b) Time-axis view.*

Figure 21. *3D plot of the time evolutions of standard deviation with different values of k_1^P while $k_1^A = 0.5$ in the 1-D Model based on the estimated Michaelis-Menten function obtained by using two-layer model identification strategy. (a) Normal view with step change $\Delta k_1^P = 0.05$; (b) Standard deviation vs time plot with different values of k_1^P. For clarity, the values of k_1^P applied are listed only for the top four pairs of curves.*

The time evolution of estimated Gaussian function is presented in terms of 3D plots in Figure 20. The height of the curve gradually decreases and approaches to the steady state and it is observed that the axis of symmetry of Gaussian curve slowly deviates from its initial position (see

blue and red curves in Figure 20(b)). Similarly, the time evolutions of estimated Michaelis-Menten function with different value of k_1^P are presented in terms of 3D plots in Figure 21.

The regression model analysis on the money distribution evolution in time is carried out at fixed input variables and condition, φ. Although the result is valid only at this specific scenario, the analysis procedure presents a good approach to study the money distribution evolution by looking for candidate regression functions and estimating the evolutions of function parameters. In order to have the regression model validated and predictive on the stochastic simulation, the regression model identifications on different input scenarios, i.e. $\varphi_1, \varphi_2, \dots , \varphi_n$, are required to formulate the data set inclusive of various situations as many as possible. A full regression model analysis on one fixed input scenario comprises two sections, i.e. simulation and identification, which could be computationally expensive and intensive.

4. Conclusion

In this paper, a new problem, namely the Travelling Traders' Exchange Problem (TTEP) is formulated and proposed based on several assumptions to investigate collision-exchange systems of interest in a wide range of physical applications. The purpose of TTEP is to characterise the distribution of money over time related to a population of traders which can interact with each other. TTEP models offer a novel general mathematical formulation to investigate a wide range of physical phenomena (molecular and particle interactions in chemical reaction processes[1], thermal conduction in lattice[2], disease transmission mechanisms in epidemiology studies[3] are only some examples). Different assumptions on the interaction mechanisms and dimension can be introduced in the detailed construction of these models according to the specific features of real processes. In this paper, a 1-D stochastic model of TTEP has been developed to explore the behaviour of the TTEP requiring the definition of (i) initial allocation of amount of money to each trader; (ii) interaction mechanism of any two traders who collide. The exchange events are mathematically

expressed via a number of operating variables including money exchange location, money exchange direction and money transfer coefficients. A simple linear interaction mechanism is postulated to minimise the number of model parameters, preserving at the same time the significance and tractability of the mathematical problem.

A two-layer model identification strategy has been employed to develop nonlinear regression models from 1-D stochastic simulations capable of representing the observed distribution of money exchange in time as a function of operating variables. Stochastic modelling and simulation based on Monte Carlo method represent suitable approaches to reveal the nature of the TTEP. The 1-D Model has been discussed from two different aspects, time evolution of standard deviation and probability density function of money redistribution. Several illustrative examples are used to compare the results of different stochastic simulations on the 1-D Model with various input scenarios. Parametric study on input variables qualitatively illustrates and visualises the impact of these variables on the simulation results of the 1-D Model. In regression model identification section, a Michaelis-Menten function is selected in the first layer to be further refined in the second-layer identification. A symmetric k_1^A-k_1^P map is obtained to present different terminal forms of PDF of money distribution at steady state conditions in the 1-D Model. This map represents a useful tool to characterise the TTEP system with different money transfer coefficients. The Gaussian function applied to estimate the sample data gives smooth fit results in the second layer. With the help of k_1^A-k_1^P map and the two-layer identification strategy, a regression model can be validated and predict the probability density function of money distribution in time the 1-D TTEP model. Future work will aim at automating the regression model identification procedure and to extend.

The study on the TTEP via the simulation of the 1-D Model gives visible and accessible insights about the emergent behaviour of stochastic processes and evolution over time of the distribution of traders' accounts. The results obtained by regression model identification reveal the relation between the money redistribution and operating conditions. These phenomena would benefit from further optimisation studies on operating conditions with the goal to achieve target

money distribution properties. The study on the TTEP system becomes extremely valuable when input variables in the stochastic simulation can directly be linked to the real industrial or experimental operating variables. Future work will aim at providing a feasible and systematic approach to process optimisation where these operating variables will be manipulated to achieve tunable distribution properties while minimising the overall processing time.

Acknowledgements

List of symbols

General symbols

A^G	Parameter in Gaussian function
A_0^G	Value of A^G at large simulation time
D_{ec}	Difference between money transfer coefficients for active and passive traders
D_i	Exchange direction of event S_i with respect to active trader
E	Terminal simulation time point
E_A	Amount of money transferred from active to passive trader in exchange event
E_P	Amount of money transferred from passive to active trader in exchange event
K	Number of compartments in computational domain $[M_{min}, M_{max}]$
K_m	Parameter in Michaelis-Menten function
k_i^A	Money transfer coefficient of i-th order term of active trader
k_i^P	Money transfer coefficient of i-th order term of passive trader
L_i	Exchange location of event S_i with respect to active trader

M_{ij}	Amount of money held by the trader N_j after the exchange event S_i
$M_{\max}$	Upper bound of the computational domain
$M_{\min}$	Lower bound of the computational domain
$M_{S_i}^A$	Amount of money held by involved active trader after exchange event
$M_{S_i}^P$	Amount of money held by involved passive trader after exchange event
$M_{S_{i-1}}^A$	Amount of money held by involved active trader before exchange event
$M_{S_{i-1}}^P$	Amount of money held by involved passive trader before exchange event
N	Number of traders (population of traders) in the system
N_θ	Dimension of set of operating variables
N_φ	Dimension of set of regression parameters
P_B	Backward money exchange propensity
P_F	Forward money exchange propensity
R	Number of simulation runs
S_i	i-th money exchange event
T_{total}	Termination time
t	Time
t_i	Time at exchange event S_i
V_{max}	Parameter in Michaelis-Menten function
w	Width of compartments in computational domain $[M_{\min}, M_{\max}]$
w^G	Parameter in Gaussian function
w_0^G	Value of w^G at large simulation time
x_i	x-axis value of i in the sample
x_c^G	Parameter in Gaussian function
x_{c0}^G	Value of x_c^G at large simulation time

X_D	Output of random number generator
X_L	Output of random number generator
y	Stochastic simulation results
$\hat{y}$	Estimated function of corresponding simulation results
y_i	y-axis value of i in the sample
y_0^G	Parameter in Gaussian function
$\mathbb{Z}^+$	Set of positive integers

Vectors and Matrices [dimension]

$\boldsymbol{M}_0$	Vector of initial amount of money in the 1-D Model *[N]*
$\boldsymbol{N}$	Vector of trader sequence in the 1-D Model *[N]*
$\boldsymbol{\theta}$	Set of candidate function parameters $[N_\theta]$
$\boldsymbol{\varphi}$	Set of model operating variables $[N_\varphi]$

Greek Letters

Δt	Time gap between two adjacent exchange events
Δt_s	Sampling time step
τ	Total time for trading season

Acronyms

COD	Coefficients of determination
PDF	Probability density function
RSS	Residual Sum of Squares
SSA	Stochastic simulation algorithm

SDEs	Stochastic differential equations
STD	Standard deviation
TTEP	Travelling traders' exchange problem

References

1. Gibson, M. A.; Bruck, J. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A.* **2000**, *104*, 1876-1889.

2. Basile, G.; Bernardin, C.; Jara, M.; Komorowski, T.; Olla, S. Thermal Conductivity in Harmonic Lattices with Random Collisions. *Thermal Transport in Low Dimensions: From Statistical Physics to Nanoscale Heat Transfer*; Springer International Publishing: Cham, 2016; pp. 215-237.

3. Bansal, S.; Grenfell, B. T.; Meyers, L. A. When individual behaviour matters: homogeneous and network models in epidemiology. *J. Royal Soc. Interface* **2007**, *4*, 879-891.

4. Nelson, B. *Foundations and Methods of Stochastic Simulation: A First Course*; Springer US: New York, 2013.

5. Gillespie, D. T. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **2001**, *115*, 1716-1733.

6. Haseltine,   E. L.; Rawlings, J. B. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys*. **2002**, *117*, 6959-6969.

7. Oksendal, B. *Stochastic differential equations: an introduction with applications*; Springer-Verlag Berlin Heidelberg, 2003.

8. Erban, R.; Chapman, J. S.; Maini, P. A practical guide to stochastic simulations of reaction-diffusion processes, 2007, arXiv:0704.1908. arXiv preprint. https://arxiv.org/pdf/0704.1908.pdf

9. Andrews, S. S.; Bray, D. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys. Biol*. **2004**, *1*, 137.

10. Erban, R.; Chapman, J. S. Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Phys. Biol.* **2009**, *6*, 046001.

11. Erban, R.; Chapman, J. S. Reactive boundary conditions for stochastic simulations of reaction-diffusion processes. *Phys. Biol.* **2007**, *4*, 16.

12. Erban, R.; Chapman, J. S. Time scale of random sequential adsorption. *Phys Rev E.* **2007**, *75* (4), 041116.

13. Ulam, S.; Richtmyer, R. D.; Neumann, J. von Statistical methods in neutron diffusion. *Los Alamos Sci. Lab., [Rep.]* **1947**, *LAMS-551*.

14. Doucet, A.; De Freitas, N.; Gordon, N. *An Introduction to Sequential Monte Carlo Methods*; Springer: New York, 2001.

15. Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **1976**, *22*, 403-434.

16. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340-2361.

17. Huang, C.; Piccione, P. M.; Cattani, F.; Galvanin, F. Travelling traders' exchange problem: stochastic simulation and sensitivity analysis. *Proceedings of the AIChE Annual Meeting*, San Francisco, CA, November 13–18, 2016.

18. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **1974**, *19*, 716-723.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1-D Position-fixed Linear Model

Traders

| 1 | 2 | 3 | 4 | ... | n | n+1 | ... | N-2 | N-1 | N |

Adjacent Money Exchange Event

$$E^A \big|_t = f\left(k^A, M^A_{t-\Delta t}\right)$$

| n | $\longleftrightarrow$ | n+1 |

$$E^P \big|_t = f\left(k^P, M^P_{t-\Delta t}\right)$$

Stochastic Simulation and System Evolution

$t = t_0$ | M_0 | = | M_{01} | M_{02} | M_{03} | ... | M_{0N}

Δt

$t = t_1$ | M_1 | = | M_{11} | M_{12} | M_{13} | ... | M_{1N}

Δt

$t = t_2$ | M_2 | = | M_{21} | M_{22} | M_{23} | ... | M_{2N}

... $(T_{\text{total}} - 2)\Delta t$

$t = t_E$ | M_E | = | M_{E1} | M_{E2} | M_{E3} | ... | M_{EN}

Time evolution of money distribution in the 1-D Model

Regression model identification on population size of the 1-D Model