

Uniwersytet Warszawski
Wydział Fizyki

Rozprawa doktorska pt.

Modele błędzenia losowego
w czasie ciągłym z pamięcią.
Zastosowanie do opisu dynamiki
rynków finansowych

Tomasz Gubiec

Promotor:

prof. dr hab. Ryszard Kutner

Zakład Fizyki Biomedycznej

Instytut Fizyki Doświadczalnej UW

Warszawa, sierpień 2011 r.

Serdecznie dziękuję mojej żonie Joannie za wsparcie oraz Profesorowi Ryszardowi Kutnerowi za uwagi, wskazówki i cierpliwe sprawdzanie moich obliczeń.

Spis treści

1	Wprowadzenie	5
1.1	Cel pracy	8
1.2	Układ pracy	9
2	Dane empiryczne i metoda ich analizy	11
2.1	Opis danych empirycznych	11
2.2	Algorytm obliczania funkcji autokorelacji prędkości skokowego szeregu czasowego	13
2.3	Kwantyle autokorelacyjne	16
3	Kanoniczne błędzenie losowe w czasie ciągłym	19
3.1	Wyprowadzenie jawnej postaci propagatorów	22
3.2	Stacjonaryzacja formalizmu CTRW	23
4	Stacjonaryzacja modelu Montero i Masolivera	27
4.1	Geneza modelu	27
4.2	Pamięć jednokrokowa w ramach CTRW	28
4.3	Rozwiązanie modelu stacjonarnego	30
5	Model z pamięcią jednokrokową	35
5.1	Motywacja empiryczna modelu	35
5.2	Motywacja teoretyczna: zjawisko bid-ask bounce i jego mechanizm	38
5.3	Definicja modelu	41
5.4	Wyprowadzenie jawnej postaci propagatorów	44
5.5	Wariancja i funkcja autokorelacji prędkości	46
5.6	Porównanie z danymi empirycznymi	47
5.6.1	Estymacja parametru ϵ	48
5.6.2	Rozkład czasów wyczekiwania	50
5.6.3	Funkcja autokorelacji prędkości	53
6	Model z pamięcią dwukrokową	59
6.1	Motywacja modelu	59
6.1.1	Analiza co drugiej zmiany ceny	60
6.1.2	Zjawisko bid-ask bounce z opóźnieniem	62
6.2	Definicja modelu	67
6.3	Pamięć dwukrokowa w ramach CTRW	70
6.4	Wariancja i funkcja autokorelacji prędkości	71

6.5	Porównanie z danymi empirycznymi	73
7	Zakończenie	77
7.1	Podsumowanie uzyskanych wyników	77
7.2	Ocena wyników	79
7.3	Konkluzje	81
A	Wyprowadzenie propagatora dla modelu z pamięcią dwukrokową	83
B	Wpływ trendu dziennego na autokorelację prędkości	97
C	Autokorelacja znaku prędkości ceny	103

Rozdział 1

Wprowadzenie

Niniejsza praca leży w nurcie badań nad procesami o charakterze probabilistycznym, w których czas jest traktowany w sposób ciągły. Rozważane w pracy modele należą do tej kategorii tego nurtu, dla której czasy pomiędzy kolejnymi zdarzeniami mogą ulegać zmianie. Kanonicznym przedstawicielem tej kategorii jest Model Błądzenia Losowego w Czasie Ciągłym (ang. *Continuous-Time Random Walk*, CTRW).

Model CTRW został wprowadzony przez fizyków Montrolla i Weissa w roku 1965 [1], jako uogólnienie klasycznego błądzenia losowego, w którym czas pomiędzy kolejnymi skokami procesu stochastycznego jest zmienną losową. Model ten (zwany dalej kanonicznym) został po raz pierwszy zastosowany przez Schera i Laxa w 1973 roku [2] oraz niezależnie, rok później, przez Moore'a [3] do opisu anomalnej relaksacji fotoprądu w ciałach amorficznych; prąd ten przejawiał relaksację potęgową, a nie wykładniczą. Uzyskane wyniki zapoczątkowały rozwój wspomnianej kategorii procesów stochastycznych. Przykładowo, prawie dekadę później Tunaley [4, 5] rozszerzył kanoniczny formalizm¹ CTRW na stacjonarny model CTRW, wprowadzając odrębny rozkład czasów wyczekiwania dla pierwszego skoku. Pozwoliło to, m.in. na uniezależnienie charakterystyki procesu od warunków początkowych.

Kanoniczny CTRW znalazł szereg zastosowań takich jak np. opis starzenia się szkła [6, 7], transport w nieuporządkowanych przewodnikach jonowych [8], w modelowaniu zagadnień hydrologicznych [9, 10] czy też trzęsień ziemi [11].

Co więcej, formalizm CTRW można rozpatrywać jako przykład stochastycznego procesu odnowy [12, 13], w którym odstępy czasu pomiędzy skokami są zmienną losową, charakteryzowaną przez dowolny rozkład prawdopodobieństwa (a nie tylko rozkład Poissona). Teoria procesów odnowy (ang. *renewal theory*), zwanych też procesami półmarkowa (ang. *semi-markov model*), pojawiła się w okresie późniejszym (w stosunku do prac [1, 2]) w środowisku matematycznym. Ponadto, w środowisku tym rozwijano tzw. ukryte modele półmarkowa (ang. *hidden semi-markov model*). Znalazły one zastosowanie w wielu dziedzinach od biologii poprzez telekomunikację po ubezpieczenia i finanse [14–27], włączając w to ekonomię [28], a zwłaszcza ekonometrię [29]. W pierwszej dekadzie obecnego stulecia obserwuje się wzrost za-

¹Terminy 'formalizm' i 'model' używane są tutaj jako synonimy

interesowania zastosowaniami CTRW w szeroko rozumianych finansach [20–22].

Aby wskazać kierunek rozważań prowadzonych w niniejszej pracy zauważmy, że w ramach skokowego² formalizmu CTRW pojedynczą trajektorię, będącą realizacją błędzenia losowego, można przedstawić jako ciąg skoków kolejno o $r_1, r_2, \dots, r_n, \dots$, poprzedzonych okresami wyczekiwania w tym samym stanie, reprezentowanymi odpowiednio przez czasy $t_1, t_2, \dots, t_n, \dots$. Prawdopodobieństwo wystąpienia konkretnej trajektorii można sprowadzić do gęstości prawdopodobieństwa [12] wystąpienia pary (r_n, t_n) , pod warunkiem całej historii $(r_{n-1}, t_{n-1}; r_{n-2}, t_{n-2}; \dots; r_2, t_2; r_1, t_1)$. W kanonicznym modelu Montrolla i Weissa przyjmuje się następujące uproszczenie

$$\rho(r_n, t_n | r_{n-1}, t_{n-1}; r_{n-2}, t_{n-2}; \dots; r_2, t_2; r_1, t_1) \approx \rho(r_n, t_n) \approx h(r_n) \psi(t_n). \quad (1.1)$$

Oznacza to, że przyjmuje się poniższe założenia:

- (1) długość skoku r_n oraz czas wyczekiwania t_n poprzedzający ten skok są niezależne, a ponadto
- (2) nie zależą od wcześniejszych długości skoków i czasów wyczekiwania, co więcej
- (3) rozkłady wszystkich skoków są jednakowe, podobnie jak rozkłady wszystkich czasów wyczekiwania.

Założenia te wyniknęły z analizy danych doświadczalnych otrzymanych w ramach wspomnianego wcześniej eksperymentu, w którym mierzono relaksację fotoprądu. Analiza ta pozwoliła stwierdzić, że migrujący nośnik prądu przebywa w pojedynczym minimum potencjału przez czas znacznie dłuższy od okresu termalizacji, przez co wszelka pamięć o poprzednich skokach jest gubiona.

Modele błędzenia losowego w czasie ciągłym z pamięcią, wprowadzone i używane w mojej rozprawie (rozdziały 5 i 6), różnią się od kanonicznego modelu tym, że opierają się na ogólniejszym założeniu. Jest to podyktowane innym charakterem danych empirycznych (patrz poniżej). W możliwie najprostszym przypadku (rozdział 5) modyfikuję założenie (2) przyjmując, że kolejny skok może zależeć od skoku poprzedzającego, czyli:

$$\begin{aligned} \rho(r_n, t_n | r_{n-1}, t_{n-1}; r_{n-2}, t_{n-2}; \dots; r_2, t_2; r_1, t_1) &\approx \\ &\approx \rho(r_n, t_n | r_{n-1}) \approx h(r_n | r_{n-1}) \psi(t_n). \end{aligned} \quad (1.2)$$

W większości układów spotykanych dotychczas w przyrodzie jak i w układach badanych w ramach ekono- i socjofizyki wspomniana powyżej zależność, o ile występuje, ma charakter ujemnego sprzężenia zwrotnego. Sprzężenie to jest efektem działania układu, głównie przeciwko zewnętrznym czynnikom. Efekt ten stara się przywrócić pierwotny stan układu. Tego typu zachowanie układów fizycznych, znajdujących się w stanie równowagi lub w pobliżu stanu równowagi, jest opisane przez

²W dalszej części pracy opuszczać będę termin 'skokowy'.

regułą Le Chateliera – Brauna zwaną również regułą przekory [30]. Znanym przykładem rozszerzenia tej reguły na sferę finansów jest zasada eliminacji arbitrażu przez wolny rynek [31, 32].

Prawie trzy dekady temu ujemne sprzężenie zwrotne, w postaci korelacji powrotnych, było rozpatrywane w pracach [33–36] (i referencje tamże) w kontekście błędzeń losowych w gazach sieciowych. Korelacje te prowadziły nawet do znacznego zmniejszenia współczynnika samodyfuzji. Było to szczególnie widoczne dla dyfuzji przy skrajnie małej koncentracji wakancji [37], np. w przewodnictwie stałych elektrolitów [38], a zwłaszcza w dyfuzji wodoru w metalach przejściowych, gdzie koncentracja może być nawet “dowolna” [34]. Ponadto, tunelowanie wspomagane fononowo tworzące polaron, może być kolejnym interesującym przykładem korelacji powrotnej [37, 39–41].

W przypadku dyfuzji wodoru w metalach przejściowych korelacje powrotne pojawiają się w analizie dwóch kolejnych przeskoków atomu znaczonego pomiędzy położeniami międzywęzłowymi. Atom wykonując skok zostawia za sobą wolne miejsce, co skutkuje zwiększonym prawdopodobieństwem jego powrotu do wyjściowego położenia. Tendencja ta słabnie z upływem czasu pomiędzy skokami. W celu opisanie tego typu dynamiki stochastycznej, w ramach modelu gazu sieciowego, w pracy [34] rozszerzono kanoniczny formalizm CTRW, tak aby uwzględniał korelacje powrotne.

Ostatnio, problem powrotnych, czyli ujemnych korelacji pojawił się w innym, finansowym kontekście [42–45] (i referencje tamże). W finansowych szeregach czasowych wysokiej częstotliwości, jak np. cena pojedynczej akcji w funkcji czasu, powszechnie obserwowane były ujemne korelacje dwóch kolejnych skoków [26], natomiast korelacje poprzez trzy lub więcej skoków były praktycznie nieobecne. Ponadto, silna zależność pomiędzy dwoma kolejnymi skokami ceny, obserwowanymi na rynku finansowym, może dominować nad obserwowaną, słabszą, statystyczną zależnością kolejnych czasów oczekiwania pomiędzy transakcjami.

Wspomniana powyżej silna zależność pomiędzy kolejnymi skokami ceny ma swoje źródło w mikrostrukturze rynku i może być lepiej zrozumiana dopiero po przeprowadzeniu analizy tzw. księgi zleceń, która jest podstawą mechanizmu działania giełdy.

Aby usystematyzować rynek oparty na podwójnej aukcji [46–48] jakim jest giełda, wprowadza się pojęcie księgi zleceń, czyli zbioru wszystkich zleceń kupna i sprzedaży uporządkowanego względem ceny i rodzaju. Szczególnie interesującą własnością podwójnej aukcji jest występowanie tzw. zjawiska ‘bid-ask spread’³ [26]. Jest to przedział pomiędzy ceną oferty kupna o najwyższym dostępnym limicie (ang. bid) a ceną oferty sprzedaży o najniższym dostępnym limicie (ang. ask). Bid-ask spread stojący za obserwowaną dynamiką cen transakcyjnych prowadzi do zjawiska znanego jako ‘bid-ask bounce’ [26, 49]. Zjawisko to polega na “odbijaniu się” ceny transakcyjnej pomiędzy górną i dolną granicą bid-ask spreadu, co skutkuje wyraźną antykorelacją

³Termin ‘spread’ w kontekście finansowym nie ma swojego odpowiednika w języku polskim. Czasami tłumaczy się go jako “rozwarcie” lub “rozpiętość”, lecz tłumaczenie to nie jest powszechnie stosowane.

kolejnych zmian cen.

Mechanizm powstawania skoków powrotnych w dynamice cen akcji jest analogiczny do mechanizmu generującego korelacje powrotne we wspomnianej już samodyfuzji atomu wodoru w metalach przejściowych. Podobnie jak we wspomnianym przykładzie, zmieniająca się cena pozostawia za sobą wolne miejsce w księdze zleceń. Skutkuje to tendencją powrotną ceny do jej wartości wyjściowej. **Zatem, mamy tutaj do czynienia nie tylko z formalnym podobieństwem wzorów opisujących dynamikę, ale też z faktycznym podobieństwem mechanizmów.**

W kolejnej części mojej pracy (rozdział 6) w znacznie bardziej ogólny sposób zmieniam założenie (2). Dopuszczam możliwość istnienia zależności (nie mylić z korelacją) pomiędzy kolejnym skokiem a dwoma wcześniejszymi skokami bezpośrednio go poprzedzającymi, co można zapisać jako

$$\begin{aligned} \rho(r_n, t_n | r_{n-1}, t_{n-1}; r_{n-2}, t_{n-2}; \dots; r_2, t_2; r_1, t_1) &\approx \\ &\approx \rho(r_n, t_n | r_{n-1}, r_{n-2}) \approx h(r_n | r_{n-1}, r_{n-2}) \psi(t_n). \end{aligned} \quad (1.3)$$

Innymi słowy, rozszerzam założenie o istnieniu pamięci jednokrokowej na pamięć dwukrokową. Należy podkreślić, że modele tego typu nie były dotychczas rozwijane. Takie założenie znajduje swoje uzasadnienie w omawianych dalej faktach empirycznych oraz w efekcie bid-ask bounce działającym z opóźnieniem. W takiej sytuacji efekt skoku powrotnego pojawia się dopiero po wykonaniu przez cenę jednego skoku niezależnego, pomiędzy skokami o ujemnej korelacji.

W pracy rozważam systematycznie trzy modele oparte na formalizmie błędzenia losowego w czasie ciągłym. Dwa z nich zawierają pamięć jednokrokową: jeden (rozdział 4) będący własnym uogólnieniem istniejącego modelu opisanego w pracy [45] oraz drugi (rozdział 5) w pełni autor-ski [50]. Trzeci z przedstawionych modeli (rozdział 6) jest nowatorskim modelem zawierającym pamięć dwukrokową [51].

1.1 Cel pracy

Ogólnym celem mojej rozprawy doktorskiej jest ścisłe i analityczne modelowanie, na poziomie cen transakcyjnych, mechanizmu rządzącego dynamiką cen w skali mikroskopowej, czyli pojedynczych zleceń. Opisywane w pracy modele błędzeń losowych w czasie ciągłym z pamięcią mają na celu jak najdokładniejszy opis dynamiki rynku finansowego w najkrótszej możliwej skali czasowej, bazując na danych empirycznych wysokiej częstości (typu tick-by-tick). Modele opisywane w mojej rozprawie doktorskiej opierają się przede wszystkim na pracach [45, 50] i pracy przygotowywanej do druku [51].

W pracy [45] dokonane zostało uogólnienie modelu zaproponowanego przez Montero i Masolivera [42, 43]. Wprowadzono, podobnie jak w pracy Tunaley'a [4, 5],

nowy rozkład czasów oczekiwania dla pierwszego skoku. Pozwoliło to na stacjonaryzację używanych w modelu wielkości. Zmiana ta umożliwiła również porównanie teoretycznej i empirycznej funkcji autokorelacji prędkości procesu.

Model zaproponowany w pracach [42, 43] okazał się jednak zbyt przybliżony, aby jego uogólnienie w pracy [45] mogło dać satysfakcjonującą zgodność z danymi empirycznymi. Zatem, w pracy [50] rozwinięty został bardziej złożony model z pamięcią jednokrokową zainspirowany statystyką empiryczną. W tym modelu zgodność wyników teoretycznych z empirycznymi była już zadowalająca, jednak nie wszystkie dostępne fakty empiryczne były przez model odtwarzane.

Aby poprawić zgodność z danymi empirycznymi i uwzględnić np. obserwowany empirycznie brak korelacji między co drugą zmianą ceny, rozszerzyłem w niniejszej pracy powyższy model do pamięci dwukrokowej.

Ponadto, w pracy analizuję wpływ obserwowanego dziennego trendu [52–54] (i referencje tamże) na kształt empirycznej funkcji autokorelacji, a pośrednio na stacjonarność analizowanych danych.

1.2 Układ pracy

Niniejsza rozprawa składa się z pięciu uzupełniających się części:

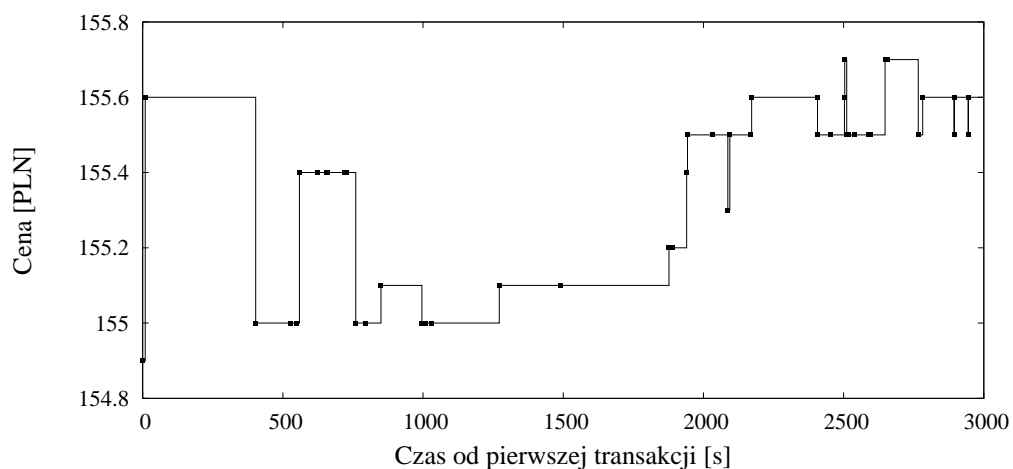
- w pierwszej części opisuję analizowane dane empiryczne (rozdział 2), wprowadzam stosowane metody (rozdziały 2 i 3) oraz stosowany formalizm (rozdział 3),
- w drugiej uwzględniam wpływ pamięci jednokrokowej na dynamikę ceny, stacjonaryzując model Montero i Masolivera (rozdział 4) i budując w pełni autorski model (rozdział 5),
- w trzeciej części rozszerzam podejście do pamięci dwukrokowej, wprowadzając kolejny oryginalny model (rozdział 6),
- w czwartej podsumowuję uzyskane w rozprawie wyniki, dokonuję ich oceny i przedstawiam konkluzje (rozdział 7),
- natomiast w dodatkach do niniejszej rozprawy (stanowiących piątą część) przedstawiam ściśle wyprowadzenie propagatora dla modelu z pamięcią dwukrokową (dodatek A), analizuję wpływ obserwowanego dziennego trendu na dynamikę ceny waloru giełdowego (dodatek B) oraz badam autokorelację znaku prędkości ceny (dodatek C).

Rozdział 2

Dane empiryczne i metoda ich analizy

2.1 Opis danych empirycznych

Część empiryczna niniejszej rozprawy opiera się na publicznie dostępnych, transakcyjnych danych wysokiej częstotliwości pochodzących z warszawskiej Giełdy Papierów Wartościowych (patrz rysunek 2.1).



Rysunek 2.1: Przykładowy fragment szeregu czasowego wysokiej częstotliwości dla spółki PEKAO z 29 lipca 2011 r. Punkty oznaczają chwilowe ceny transakcyjne. Cienkie linie łączą te punkty na sposób skokowy. Na osi pionowej wykreślona jest cena, natomiast na poziomej czas liczony od chwili pierwszej transakcji w dniu.

Zostały one udostępnione przez Bank Ochrony Środowiska SA [55]. Dane zarchiwizowane są w formacie plików programu komputerowego Metastock [56] i obejmują okres od rozpoczęcia działania systemu komputerowego WARSET [57], czyli 17 listopada 2000 r. do 29 lipca 2011 r. Dokładność cenowa powyższych danych jest maksymalna, tzn. odpowiada zawsze minimalnej możliwej zmianie ceny (ang. tick). Niestety, dostępne dane empiryczne są rejestrowane z dokładnością do jednej se-

kundy. Dlatego kilku transakcjom zrealizowanym w ciągu jednej sekundy zegarowej odpowiada jeden rekord danych.

Od dnia 1 stycznia 2011 harmonogram sesji na warszawskiej Giełdzie Papierów Wartościowych wygląda następująco:

- 8.00-9.00 przyjmowanie zleceń na otwarcie,
- 9.00 otwarcie (określenie kursu na otwarciu),
- 9.00-17.20 notowania ciągłe,
- 17.20-17.30 przyjmowanie zleceń na zamknięcie,
- 17.30 zamknięcie (określenie kursu na zamknięciu),
- 17.30-17.35 dogrywka.

W historii GPW harmonogram sesji był wielokrotnie zmieniany, a okres notowań ciągłych ulegał wydłużeniu.

W niniejszej pracy skupiam się na analizie wewnątrzdziennej dynamiki stochastycznej zmian cen, tzn. dynamiki odpowiadającej notowaniom ciągłym. Ponieważ okresy notowań ciągłych trwające obecnie 8 godzin 20 minut oddzielone są prawie 16 godzinną przerwą, błędem byłoby 'sklejanie' końca sesji z dnia poprzedniego z początkiem następnego. Nie można bowiem zakładać, iż struktura wewnątrzdzienna szeregu czasowego jest kontynuowana bez żadnej zmiany od początku sesji dnia następnego. Aby uniknąć popełnienia tego błędu, posiadane dla każdego analizowanego waloru dane traktuję jak zespół statystyczny dni. Badane charakterystyki będą wyliczał dla szeregu czasowego wewnątrz jednego dnia i uśredniał po takim zespole statystycznym (wagi elementów omówię w dalszej części tego rozdziału).

Analizując dane transakcyjne z jednego dnia należy zauważyć, iż posiadają one elementy (transakcje), które nie są interesujące z punktu widzenia dynamiki wewnątrzdziennej, m.in. należy do nich zapis pierwszej transakcji w danym dniu, zawierający realizację wszystkich ofert złożonych w czasie przed otwarciem. Sposób realizacji tych transakcji, zapisanych jako jeden rekord danych, jest zupełnie odmienny od sposobu realizacji następujących po nich transakcji w czasie notowań ciągłych. Transakcje otwarcia charakteryzują się znacznie większą zmiennością w porównaniu ze zmianami ceny wewnątrz dnia. Aby nie zaburzyć szeregu czasowego i nie zniekształcić końcowego wyniku, pomijam w dalszej analizie wszystkie transakcje otwarcia.

Drugą grupą transakcji o innym charakterze niż notowania ciągłe są transakcje na zamknięcie sesji i w czasie dogrywki, które odbywają się po stałej cenie zwanej ceną zamknięcia. W związku z tym, również te transakcje pomijam w dalszej analizie.

Kolejną cechą wymagającą ujednoczenia jest dokładność cenowa danych. Minimalna możliwa zmiana ceny η ulega zmianie w zależności od ceny waloru, jednak

w analizowanych przypadkach nigdy nie była większa od 0.1 PLN i nie mniejsza niż 0.01 PLN. W związku z tym, dla każdej spółki ograniczę się do badania danych z dokładnością równą maksymalnej wartości η występującej w trakcie jej notowań. Wszystkie ceny transakcyjne zostają zaokrąglone do tak znalezionej wartości η .

Dane uzyskane za pomocą powyższego podejścia, mogą być przedstawione, dla każdego analizowanego waloru, za pomocą zespołu statystycznego składającego się z d_{max} dni. W każdym dniu $d = 1, 2, \dots, d_{max}$, ma miejsce $n_d + 1$ transakcji, przy czym ich liczba jest na ogół różna dla różnych dni. Przez $r_{d,i}$ oznaczamy zmianę ceny pomiędzy transakcją $i + 1$ a transakcją i w dniu d , natomiast przez $t_{d,i}$ - okres czasu pomiędzy tymi transakcjami. Zdefiniujmy jeszcze czas pomiędzy pierwszą a $i + 1$ transakcją w dniu d ¹

$$\mathbf{t}_{d,k} = \sum_{i=1}^k t_{d,i}, \quad k = 1, 2, \dots, n_d, \quad (2.1)$$

oraz długość dnia transakcyjnego (sesji) $T_d = \mathbf{t}_{d,n_d}$. Jak widać, długość dnia transakcyjnego może fluktuować. Zdefiniowane powyżej wielkości są wykorzystane poniżej, w podrozdziale 2.2.

2.2 Algorytm obliczania funkcji autokorelacji prędkości skokowego szeregu czasowego

Przyjmijmy, że pomiędzy dwiema kolejnymi transakcjami cena waloru się nie zmienia, natomiast w chwili następnej transakcji zmienia się nieskończenie szybko. Zatem, analizowany szereg czasowy jest krzywą schodkową ceny w funkcji czasu - mówi się wtedy o skokowym szeregu czasowym (patrz rysunek 2.1). Chwilowa prędkość takiego procesu w funkcji czasu, liczonego od pierwszej transakcji w dniu (czyli od początku dnia transakcyjnego), przyjmuje postać:

$$v_d(\tau) = \sum_{i=1}^{n_d} r_{d,i} \delta(\mathbf{t}_{d,i} - \tau), \quad (2.2)$$

gdzie $v_d(\tau)$ jest właśnie chwilową prędkością² w dniu d w chwili τ , a δ oznacza deltę Diraca. Przy założeniu stacjonarności szeregu czasowego, składniki estymatora funkcji autokorelacji prędkości³ $C_d(t)$, liczonego jako średnia ruchoma (czyli skan oknem czasowym o szerokości t po czasie τ) w dniu d , można przedstawić w następującej

¹Czas pomiędzy transakcją i a $i + 1$ oznaczamy literą t (pisaną zwykłą czcionką w trybie matematycznym), natomiast czas od pierwszej do $i + 1$ transakcji oznaczamy przez \mathbf{t} , czyli czcionką pogrubioną.

²Dokładnie rzecz biorąc jest to pseudoprędkość, gdyż ma charakter dystrybucyjny.

³Używana tutaj nieunormowana funkcja autokorelacji nosi w ekonometrii nazwę autokowariancji.

postaci

$$\begin{aligned}
\langle v_d(\tau) v_d(\tau + t) \rangle_\tau &= \frac{1}{T_d - t} \int_0^{T_d - t} v_d(\tau) v_d(\tau + t) d\tau = \\
&= \frac{1}{T_d - t} \int_0^{T_d - t} \sum_{j=1}^{n_d} r_{d,j} \delta(\mathbf{t}_{d,j} - \tau) \sum_{i=1}^{n_d} r_{d,i} \delta(\mathbf{t}_{d,i} - \tau - t) d\tau = \\
&= \frac{1}{T_d - t} \sum_{j=1}^{n_d} \sum_{i=1}^{n_d} r_{d,i} r_{d,j} \int_0^{T_d - t} \delta(\mathbf{t}_{d,j} - \tau) \delta(\mathbf{t}_{d,i} - \tau - t) d\tau = \\
&= \frac{1}{T_d - t} \sum_{j=1}^{n_d} \sum_{i=1}^{n_d} r_{d,i} r_{d,j} \delta(\mathbf{t}_{d,i} - \mathbf{t}_{d,j} - t), \tag{2.3}
\end{aligned}$$

przy czym $n_d \geq 2$, $t \leq T_d$, natomiast $\langle \dots \rangle_\tau$ oznacza estymator średniej po czasie τ , a ponadto

$$\begin{aligned}
\langle v_d(\tau + t) \rangle_\tau &= \frac{1}{T_d - t} \int_0^{T_d - t} v_d(\tau + t) d\tau = \frac{1}{T_d - t} \int_0^{T_d - t} \sum_{i=1}^{n_d} r_{d,i} \delta(\mathbf{t}_{d,i} - \tau - t) d\tau = \\
&= \frac{1}{T_d - t} \sum_{i=1}^{n_d} r_{d,i} \int_0^{T_d - t} \delta(\mathbf{t}_{d,i} - \tau - t) d\tau = \frac{1}{T_d - t} \sum_{i=1}^{n_d} r_{d,i} \theta(\mathbf{t}_{d,i} - t) \tag{2.4}
\end{aligned}$$

oraz

$$\begin{aligned}
\langle v_d(\tau) \rangle_\tau &= \frac{1}{T_d - t} \int_0^{T_d - t} v_d(\tau) d\tau = \frac{1}{T_d - t} \int_0^{T_d - t} \sum_{i=1}^{n_d} r_{d,i} \delta(\mathbf{t}_{d,i} - \tau) d\tau = \\
&= \frac{1}{T_d - t} \sum_{i=1}^{n_d} r_{d,i} \int_0^{T_d - t} \delta(\mathbf{t}_{d,i} - \tau) d\tau = \frac{1}{T_d - t} \sum_{i=1}^{n_d} r_{d,i} \theta(T_d - t - \mathbf{t}_{d,i}), \tag{2.5}
\end{aligned}$$

gdzie θ jest funkcją skokową Heaviside'a. Można zauważyć, iż zdefiniowany za pomocą powyższych składników estymator funkcji autokorelacji prędkości⁴ wciąż zawiera składniki proporcjonalne do delty Diraca. Niedogodność tę daje się usunąć poprzez uśrednienie tego estymatora po krótkim przedziale czasu o długości Δt wokół wybranej chwili, czyli poprzez wprowadzenie definicji

$$C_d(t) = \frac{1}{\Delta t} \int_{t - \Delta t/2}^{t + \Delta t/2} \langle v_d(\tau) v_d(\tau + t') \rangle_\tau - \langle v_d(\tau) \rangle_\tau \langle v_d(\tau + t') \rangle_\tau dt'. \tag{2.6}$$

⁴Mowa tutaj o estymatorze będącym różnicą pomiędzy wyrażeniem (2.3) a iloczynem wyrażeń (2.4) i (2.5).

Tak zdefiniowany estymator dziennej funkcji autokorelacji prędkości można już zapisać w postaci umożliwiającej jej bezpośrednie wyznaczenie na drodze numerycznej

$$C_d(t) = \frac{1}{T_d - t} \sum_{j=1}^{n_d} \sum_{i=1}^{n_d} r_{d,i} r_{d,j} \left(\delta_{\Delta t}(\mathbf{t}_{d,i} - \mathbf{t}_{d,j} - t) - \frac{\theta_{\Delta t}(\mathbf{t}_{d,i} - t) \theta_{\Delta t}(T_d - t - \mathbf{t}_{d,j})}{T_d - t} \right), \quad (2.7)$$

gdzie

$$\delta_{\Delta t}(x) = \begin{cases} 0, & , x < -\frac{\Delta t}{2}, \\ \frac{1}{\Delta t}, & x \in [-\frac{\Delta t}{2}, \frac{\Delta t}{2}[, \\ 0, & x \geq \frac{\Delta t}{2} \end{cases} \quad \text{oraz} \quad \theta_{\Delta t}(x) = \begin{cases} 0, & , x < -\frac{\Delta t}{2}, \\ \frac{x}{\Delta t} + \frac{1}{2}, & x \in [-\frac{\Delta t}{2}, \frac{\Delta t}{2}[, \\ 1, & x \geq \frac{\Delta t}{2}. \end{cases}$$

Wyrażenie (2.7) dobrze nadaje się do implementacji algorytmicznej o ile prawidłowo dobrać się przedział czasu Δt . W pracy przyjmuję $\Delta t = 1s$.

Alternatywną metodą obliczenia funkcji autokorelacji prędkości byłoby stworzenie z posiadanych danych szeregu czasowego o stałym kroku czasowym i o zerowych skokach w okresach wyczekiwania. Warto zwrócić uwagę, iż algorytm oparty na wzorze (2.7) ma kilka charakterystycznych, dogodnych cech w porównaniu do algorytmów ze stałym krokiem czasowym:

- wymaga mniej pamięci na przechowywanie szeregu czasowego,
- może działać z dowolną dokładnością czasową, gdyż złożoność obliczeniowa i wymagana pamięć na przechowywanie szeregu od niej nie zależą,
- dopuszcza możliwość dwóch zmian ceny w tej samej chwili czasu, gdyż sumowanie następuje po transakcjach, a nie po czasie,
- jest wolny od artefaktów powodowanych przez stały krok czasowy, wyróżniająco odpowiadającą mu częstotliwość,
- nawet przy najprostszej implementacji wykonującej $\frac{n_d^2}{2}$ sumowań jest szybszy, o ile spełniona jest nierówność $\frac{n_d^2}{2} < \frac{T_d}{\Delta t}$.

Przypisując dniom transakcyjnym wagi proporcjonalne do długości danego dnia transakcyjnego T_d , czyli

$$w_d = \frac{T_d}{\sum_{i=1}^{d_{max}} T_i}, \quad (2.8)$$

definiuję estymator funkcji autokorelacji prędkości oparty na całym szeregu, obliczony jako średnia po zespole statystycznym dni

$$C(t) = \sum_{d=1}^{d_{max}} w_d C_d(t) = \frac{\sum_{d=1}^{d_{max}} T_d C_d(t)}{\sum_{d=1}^{d_{max}} T_d}. \quad (2.9)$$

Wzory (2.7) - (2.9) tworzą praktyczną podstawę algorytmu umożliwiającego numeryczne wyznaczenie funkcji autokorelacji prędkości.

2.3 Kwantyle autokorelacyjne

Posiadanie zespołu statystycznego pozwala również na szacowanie istotności uzyskanego wyniku bez zakładania postaci jakiegokolwiek rozkładu. W tym celu mogą wyznaczyć odpowiedniki kwantyli⁵ definiując

$$C^\alpha(t) = x(t) \Leftrightarrow \alpha \approx \frac{|\{d : C_d(t) \leq x(t)\}|}{d_{max}}, \quad (2.10)$$

gdzie $|A|$ oznacza tutaj moc zbioru A , natomiast $0 \leq \alpha \leq 1$ jest rzędem kwantyla autokowariancyjnego. Jak widać, rząd ten (w dalszym ciągu podawany w procentach) jest zdefiniowany jako względna liczba dni transakcyjnych, dla których dzienna funkcja autokowariancji prędkości $C_d(t)$, jest równa lub mniejsza od x . Zatem, ustalając rząd α , możemy dla każdej chwili czasu odczytać odpowiadającą jej wartość $x = x(t)$. Tak otrzymany przebieg $x(t)$ oznaczamy $C^\alpha(t)$ i nazywamy właśnie kwantylem autokowariancyjnym.

Mając na uwadze różne rozkłady skoków cen walorów poszczególnych spółek, a przez to różną wartość $C(0)$, wielkością dającą możliwość porównywania wyników dla różnych walorów jest unormowana funkcja autokorelacji prędkości⁶, czyli

$$C^n(t) = \frac{C(t)}{C(0)}, \quad (2.11)$$

a stąd także

$$C^{\alpha,n}(t) = \frac{C^\alpha(t)}{C(0)}. \quad (2.12)$$

Wielkość daną wyrażeniem (2.12) nazywamy kwantylem autokorelacyjnym rzędu α . Należy zaznaczyć, że wyrażenie (2.12) dostarcza oryginalnego narzędzia badawczego, nie wymagającego znajomości postaci jakiegokolwiek rozkładu.

Analizę narzędziami (2.11) i (2.12) przeprowadziłem dla ośmiu spółek o najwyższej liczbie transakcji wchodzących w skład indeksu WIG20. Lista spółek wraz z liczbą dni i liczbą analizowanych transakcji została podana w tabeli 2.1. Obliczone powyższą metodą funkcje autokorelacji prędkości wraz z kwantylami autokorelacyjnymi rzędu $\alpha = 10\%$ i $\alpha = 90\%$ zostały przedstawione na rysunku 2.2. Używając jako miary powyżej zdefiniowanych kwantyli autokorelacyjnych możemy powiedzieć, że na poziomie istotności 10% występowanie ujemnych autokorelacji jest zjawiskiem istotnym.

Warto zwrócić uwagę, że powyższa miara ufności nie zawiera żadnych założeń co do rozkładu zmian cen, z których uzyskiwano autokorelacje. Ponadto, przyjmuje ona jeden dzień transakcyjny jako naturalną jednostkę podziału szeregu czasowego.

⁵Te odpowiedniki kwantyli można nazwać autokowariancyjnymi.

⁶W ekonometrii dopiero tą unormowaną wielkość nazywa się autokorelacją, natomiast przed unormowaniem mówi się o autokowariancji.

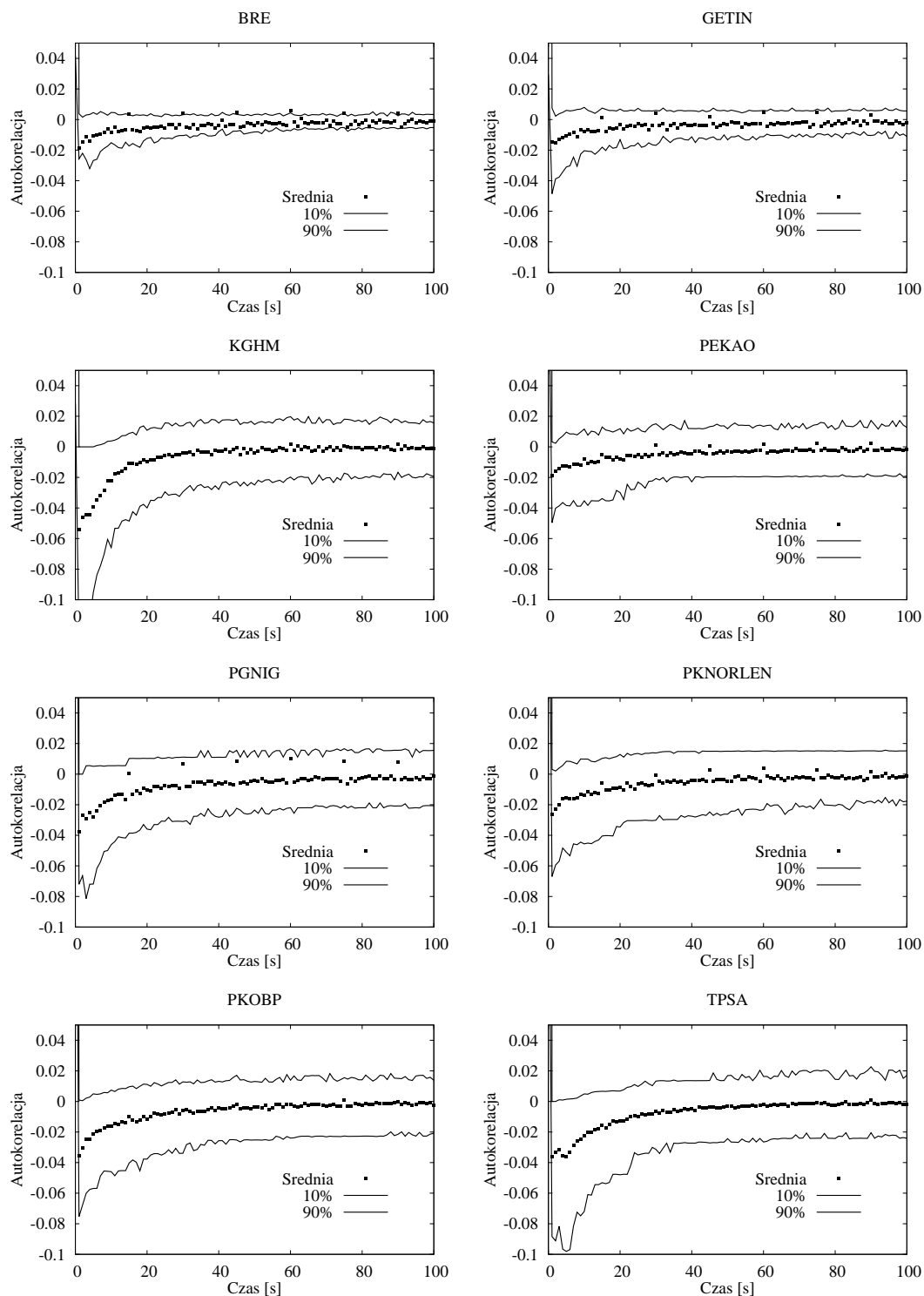
Tablica 2.1: Lista ośmiu badanych spółek wchodzących w skład indeksu WIG20 wraz z liczbą dni i transakcji.

Skrót nazwy	Pełna nazwa	Dni	Transakcje
BRE	BRE Bank	2684	550722
GETIN	Getin Holding	2559	653627
KGHM	KGHM Polska Miedź	2686	2163814
PEKAO	Bank Pekao	2686	1167810
PGNIG	Polskie Górnictwo Naftowe i Gazownictwo	1450	772398
PKNORLEN	PKN Orlen	2686	1539947
PKOBP	PKO Bank Polski	1688	1651724
TPSA	Telekomunikacja Polska	2686	1874336

Podział szeregu na dłuższe odcinki prowadziłby zapewne do zawężenia przedziałów ufności, czyli krzywe górne i dolne (oznaczone liniami ciągłymi na rysunku 2.2) zbliżyłyby się do siebie. Jednakże, zmniejszyłoby to liczebność zespołu statystycznego, pogarszając tym samym statystykę (linie ciągłe na rysunku 2.2 byłyby bardziej “poszarpane”). Zatem, wynik ten należy traktować jako ograniczenie górne, tzn. ujemność autokorelacji można również wykazać na niższym poziomie istotności. Analizując przedstawione na rysunku 2.2 średnie autokorelacje (małe czarne kwadraty), widzimy że są one wyraźnie ujemne przez blisko jedną minutę, dążąc od dołu do zera.

Zaskakującą obserwacją, obecną jedynie dla pięciu spółek (BRE, GETIN, PEKAO, PGNIG, PKNORLEN), jest występowanie dla pewnych pojedynczych wartości czasu, wyraźnie zwiększonej autokorelacji, dobrze widocznej jako czarne kwadraty znajdujące się powyżej trendu. Dostrzec można, że czasy dla których występuje to zjawisko to wielokrotności 15 sekund. Źródeł takiego zachowania można się dopatrywać w fakcie interakcji z indeksem WIG20, którego wartość publikowana jest dokładnie co 15 sekund przy wykorzystaniu algorytmicznego handlu (prowadzonego przez komputery) lub też w działaniu tzw. ‘market makerów’ - instytucji, które na zlecenie giełdy równoważą rynek.

W dalszej części pracy wracam do omawianych wyników przedstawionych na rysunku 2.2.



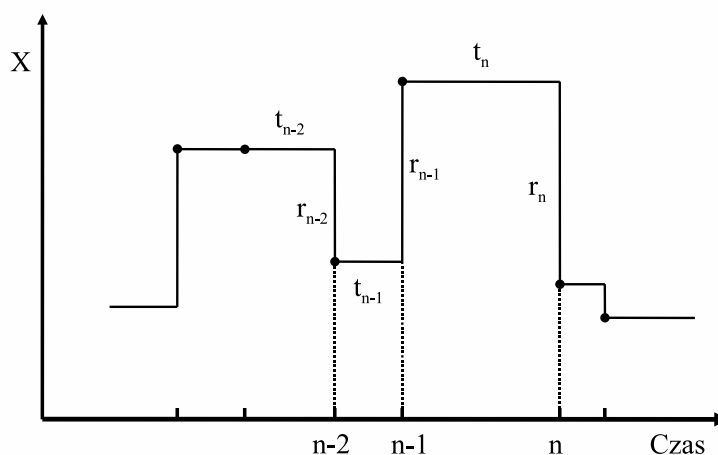
Rysunek 2.2: Unormowane funkcje autokorelacji prędkości $C^n(t)$ (małe czarne kwadraty) wraz z kwantylami autokorelacyjnymi $C^{0.1,n}(t)$ (krzywa ciągła poniżej kwadratów) i $C^{0.9,n}(t)$ (krzywa ciągła powyżej kwadratów) dla ośmiu badanych spółek. Dla wszystkich badanych spółek w około 90% dni występowała ujemna autokorelacja.

Rozdział 3

Kanoniczne błądzenie losowe w czasie ciągłym

Jak już wspomniałem w rozdziale 1, błądzenie losowe w czasie ciągłym zostało wprowadzone w pracy [1] przez fizyków Montrolla i Weissa. Obszerną część monograficzną wraz z odnośnikami literaturowymi i odniesieniami do inspirującego ten model eksperymentu, można znaleźć np. w rozprawie doktorskiej [53]. W poniższym rozdziale przedstawiam założenia i notację używaną we wspomnianym formalizmie. Stanowić to będzie bazę do oryginalnych rozszerzeń modelu zawartych w kolejnych rozdziałach niniejszej pracy.

Rozważmy, rozwijając elementy modelu przedstawione we Wprowadzeniu, proces stochastyczny, którego pojedyncza realizacja składa się z okresów wyczekiwania w tym samym stanie i natychmiastowych przeskoków do innego stanu (patrz rysunek 3.1). Dopuszczamy, że okres wyczekiwania między kolejnymi zmianami stanu jest



Rysunek 3.1: Przykładowy fragment trajektorii skokowego błądzenia losowego w czasie ciągłym wraz ze stosowanymi oznaczeniami. Trzy zaznaczone zmiany wartości procesu r_{n-2} , r_{n-1} i r_n poprzedzone są okresami wyczekiwania w tym samym stanie (oznaczone odpowiednio przez czasy t_{n-2} , t_{n-1} i t_n).

zmienną losową. Możemy teraz zdefiniować gęstość prawdopodobieństwa wystąpienia konkretnej trajektorii, wprowadzając następującą, zasadniczą wielkość¹

$\rho(r_n, t_n | r_{n-1}, t_{n-1}; r_{n-2}, t_{n-2}; \dots; r_2, t_2; r_1, t_1)$ - gęstość prawdopodobieństwa wystąpienia kolejnego przeskoku o r_n , pojawiającego się po okresie wyczekiwania t_n licząc od poprzedniego skoku, pod warunkiem całej historii, czyli wszystkich wcześniejszych skoków i czasów wyczekiwania, jakie je poprzedzały $(t_1, r_1; t_2, r_2; \dots; t_{n-1}, r_{n-1})$.

Kanoniczny model Montrolla i Weissa jest oparty o założenie wyrażone wzorem (1.1), czyli przyjmuje wymienione tam (tuż pod wspomnianym wzorem) założenia (1)-(3). Założenia te generują trajektorie, które są statystycznie niezmiennicze ze względu na przesunięcia w czasie² (stacjonarne) oraz w przestrzeni stanów, czyli w pełni jednorodne.

Kolejne, przedstawione w pracy rozszerzenia kanonicznego CTRW opierają się na zmianie wzoru (1.1), poprzez wprowadzenie innego rodzaju przybliżenia.

Zauważmy, że z aplikacyjnego punktu widzenia najbardziej przydatnym opisem procesu stochastycznego są funkcje rozkładu [58]. W przypadku procesów Markowa, wielkością wystarczającą do pełnego opisu procesu (poprzez funkcje rozkładu) jest jego propagator³ $P(x_2, t_2 | x_1, t_1)$. Jest to gęstość prawdopodobieństwa znalezienia procesu w stanie x_2 w chwili t_2 , pod warunkiem że w chwili t_1 był w stanie x_1 . Dla wspomnianych procesów właśnie z tego propagatora jesteśmy w stanie uzyskać wszystkie funkcje rozkładu [58]. Dodatkowo dla procesów stacjonarnych i translacyjnie niezmienniczych, czyli w pełni jednorodnych, zachodzi równość

$$P(x_2 + x, t_2 + \tau | x_1 + x, t_1 + \tau) = P(x_2, t_2 | x_1, t_1), \quad (3.1)$$

przez co bez straty ogólności, proces zdefiniowany jest poprzez propagator przesunięty w czasie i przestrzeni $P(x_2 - x_1, t_2 - t_1 | 0, 0)$. Dla uproszczenia zapisu w niniejszej pracy będę oznaczał ten propagator przez $P(X, t)$, gdzie $X = x_2 - x_1$ oraz $t = t_2 - t_1$.

W przypadku błędzenia losowego w czasie ciągłym, proces z jakim mamy do czynienia nie jest w ogólności procesem Markowa. Z tego powodu pełen opis procesu można uzyskać dopiero przez podanie wszystkich funkcji rozkładu. Na szczęście większość interesujących i badanych charakterystyk procesów stochastycznych, jakie są porównywane z wynikami eksperymentów stanowią charakterystyki drugorzędowe, do uzyskania których wystarczy sam propagator.

Uwzględniając specyficzny charakter opisywanego procesu skokowego i dla ułatwienia zapisu występujących zależności, używamy różnych propagatorów zdefiniowanych poniżej. Przyjmuję następujące oznaczenia:

¹Jest to precyzyjniejsza, w stosunku do wprowadzonej we Wprowadzeniu, definicja gęstości prawdopodobieństwa warunkowego ρ .

²Przy dodatkowym założeniu o istnieniu średniego czasu wyczekiwania.

³W szeroko cytowanej książce [58] wielkość ta jest oznaczana jako $P_{1|1}(x_2, t_2 | x_1, t_1)$.

- $P(X, t)$ - jak wcześniej, jest to gęstość prawdopodobieństwa tego, że w chwili t proces będzie miał wartość X , pod warunkiem, że w chwili początkowej ($t = 0$) był w zerze ($X = 0$). Tę wielkość nazywam 'miękkim' propagatorem.
- $Q(X, t)$ - gęstość prawdopodobieństwa tego, że **dokładnie** w chwili t proces **przeskoczy** do wartości X , pod warunkiem, że w chwili początkowej ($t = 0$) był w zerze ($X = 0$). Tę wielkość nazywam 'ostrym' propagatorem i rozważam tylko dla $t > 0$.

Różnica pomiędzy ostrym i miękkim propagatorem jest taka, że w przypadku miękkiego propagatora, proces może osiągnąć końcową wartość X wcześniej, przed chwilą t i pozostać w końcowym stanie (czyli bez przeskoku) aż do chwili t . Natomiast, dla ostrego propagatora wartość X osiągana jest dokładnie w chwili t .

Wprowadźmy tzw. prawdopodobieństwo przetrwania (ang. sojourn probability)

$$\Psi(t) = \int_t^{\infty} \psi(\tau) d\tau, \quad (3.2)$$

dzięki czemu możemy wyrazić zależność pomiędzy powyższymi dwoma propagatorami (miękkim i ostrym) równaniem

$$P(X, t) = \int_0^t [\delta(X) + Q(X, \tau)] \Psi(t - \tau) d\tau. \quad (3.3)$$

W powyższej relacji składnik w postaci delty Diraca pojawia się, aby uwzględnić prawdopodobieństwo niewykonania żadnego skoku przez cały okres t .

Następnie, dokonujemy podziału 'ostrego' propagatora ze względu na liczbę wykonanych skoków definiując

- $Q_n(X, t)$ - gęstość prawdopodobieństwa tego, że **dokładnie w chwili t proces przeskoży** do wartości X **wykonując dokładnie n -ty skok**, pod warunkiem, że w chwili początkowej ($t = 0$) był w zerze ($X = 0$). Tę wielkość również będę rozważał tylko dla $t > 0$.

Związek między $Q_n(X, t)$, a $Q(X, t)$ wyraża się wzorem

$$Q(X, t) = \sum_{n=1}^{\infty} Q_n(X, t). \quad (3.4)$$

Ze względu na występujące w formalizmie CTRW równania o charakterze konwolucji (splotu), uzasadnionym jest stosowanie transformaty Fouriera–Laplace'a, definiowanej (patrz poniżej) np. dla miękkiego propagatora

$$\tilde{P}(k, s) = \int_0^{\infty} dt e^{-st} \int_{-\infty}^{\infty} dx e^{ikx} P(x, t); \quad (3.5)$$

analogiczne transformacje mają miejsce dla pozostałych propagatorów. Zwróćmy uwagę, iż symbol \tilde{A} , raz będzie oznaczał transformatę Fouriera–Laplace’a funkcji A (w przypadku funkcji dwóch zmiennych), raz transformatę Fouriera (w przypadku funkcji jednej zmiennej od argumentu rzeczywistego), a raz transformatę Laplace’a (w przypadku funkcji jednej zmiennej od argumentu nieujemnego) przy czym, stosowany jest w taki sposób, aby nie prowadził do konfliktu oznaczeń. Po wprowadzeniu tych transformat zależność (3.3) przyjmuje postać

$$\tilde{P}(k, s) = \tilde{\Psi}(s) + \tilde{\Psi}(s)\tilde{Q}(k, s), \quad (3.6)$$

gdzie

$$\tilde{\Psi}(s) = \frac{1 - \tilde{\psi}(s)}{s}. \quad (3.7)$$

Natomiast z uwagi na liniowość transformaty Fouriera-Laplace’a, zależność (3.4) łatwo przekłada się na

$$\tilde{Q}(k, s) = \sum_{n=1}^{\infty} \tilde{Q}_n(k, s). \quad (3.8)$$

Przechodzimy teraz do wyprowadzenia zależności, które pozwolą nam na znaleźć propagatory w postaci jawnej.

3.1 Wyprowadzenie jawnej postaci propagatorów

Zauważmy, że pomiędzy dwoma propagatorami ostrymi typu $Q_n(X, t)$ dla kolejnych n zachodzi (dla dowolnego $h(r_n)$) zależność rekurencyjna postaci

$$Q_n(X, t) = \int_0^t dt' \psi(t') \int_{-\infty}^{\infty} dr_n h(r_n) Q_{n-1}(X - r_n, t - t'), \quad n \geq 2. \quad (3.9)$$

Jeżeli proces dokładnie w chwili t przeskakuje do wartości X wykonując n -ty skok, to musiał w $n - 1$ skokach w pewnej chwili $t - t'$ dotrzeć do innego miejsca $X - r_n$, po czym po czasie wyczekiwania t' , czyli w chwili t , wykonać skok r_n . Równanie (3.9) jest uśrednieniem po wszystkich możliwych wartościach t' oraz r_n wraz z wagami $\psi(t')$ oraz $h(r_n)$.

Równanie (3.9) po przeprowadzeniu transformaty Fouriera-Laplace’a przyjmuje prostszą postać

$$Q_n(k, s) = \tilde{\psi}(s)\tilde{h}(k)Q_{n-1}(k, s), \quad n \geq 2. \quad (3.10)$$

Wiedząc, że początkowo

$$Q_1(X, t) = \psi(t)h(X), \quad (3.11)$$

czyli

$$\tilde{Q}_1(k, s) = \tilde{\psi}(s)\tilde{h}(k), \quad (3.12)$$

łatwo można wyprowadzić z równania (3.10) ogólne rozwiązania dla dowolnego $n \geq 1$

$$\tilde{Q}_n(k, s | \xi) = \tilde{\psi}^n(s)\tilde{h}^n(k). \quad (3.13)$$

Po dokonaniu sumowania (szeregu geometrycznego) według wzoru (3.8) i wykorzystaniu po drodze wyrażenia (3.13) otrzymujemy

$$\tilde{Q}(k, s) = \frac{\tilde{\psi}(s)\tilde{h}(k)}{1 - \tilde{h}(k)\tilde{\psi}(s)}. \quad (3.14)$$

Stąd, na podstawie wyrażeń (3.6) i (3.7) otrzymujemy ostatecznie

$$\tilde{P}(k, s) = \tilde{\Psi}(s) + \tilde{\Psi}(s) \frac{\tilde{\psi}(s)\tilde{h}(k)}{1 - \tilde{h}(k)\tilde{\psi}(s)} = \frac{\tilde{\Psi}(s)}{1 - \tilde{h}(k)\tilde{\psi}(s)} = \frac{1}{s} \frac{1 - \tilde{\psi}(s)}{1 - \tilde{h}(k)\tilde{\psi}(s)}. \quad (3.15)$$

Jak widać, wyznaczenie propagatora P sprowadza się do znajomości dwóch gęstości, przestrzennej h i czasowej ψ , opisujących pojedynczy krok w procesie stochastycznym. **Jest to kluczowy wynik otwierający możliwość dokonania koniecznych uogólnień formalizmu CTRW.**

3.2 Stacjonaryzacja formalizmu CTRW

Zauważmy, że przy wyprowadzeniu wyrażenia (3.15) przyjęliśmy założenie, iż rozkład prawdopodobieństwa dla pierwszego okresu wyczekiwania, rozpoczynającego się od $t = 0$, jest identyczny jak w pozostałych przypadkach. Jest to równoważne założeniu, że dokładnie w chwili początkowej $t = 0$ miał miejsce skok do stanu $X = 0$. Uzyskane w ten sposób propagatory są związane warunkiem opisującym proces, w którym dokładnie w chwili początkowej nastąpił skok procesu do miejsca początkowego. Tak wyprowadzony propagator nie jest stacjonarny. Sposobem na stacjonaryzację propagatora zaproponowanym przez Tunaley'a [4, 5] jest założenie osobnego rozkładu czasów wyczekiwania dla pierwszego skoku. Oczywiście, metoda ta jest bardzo ogólna i może być zastosowana nie tylko do stacjonaryzacji, ale również do innych zagadnień [36, 44, 59, 60], gdyż mamy pełną dowolność zakładanego rozkładu.

W niniejszej pracy zajmuję się głównie przypadkiem stacjonarnym, w którym rozkład pierwszego skoku jest uśrednionym rozkładem po wszystkich możliwych chwilach wystąpienia przeskoku przed chwilą zero. Za równaniami (3.3) i (3.4) w pracy [36] przyjmuję, że

$$\psi_1(t) = \frac{\int_0^\infty dt' \psi(t+t')}{\int_0^\infty dt'' \int_0^\infty dt' \psi(t'+t'')} \quad (3.16)$$

oraz

$$\Psi_1(t) = \int_t^{\infty} \psi_1(\tau) d\tau. \quad (3.17)$$

Po dokonaniu transformacji Laplace'a otrzymuje się

$$\tilde{\psi}_1(s) = \frac{1}{\langle t \rangle} \frac{1 - \tilde{\psi}(s)}{s} \quad (3.18)$$

oraz

$$\tilde{\Psi}_1(s) = \frac{1 - \tilde{\psi}_1(s)}{s}. \quad (3.19)$$

Zatem, aby można było wprowadzić tak zdefiniowany rozkład czasów wyczekiwania dla pierwszego skoku $\psi_1(t)$, musi istnieć średni czas wyczekiwania pozostałych skoków, czyli

$$\langle t \rangle = \int_0^{\infty} t \psi(t) dt < \infty. \quad (3.20)$$

Warto zwrócić uwagę, że jedynym ciągłym przypadkiem rozkładu dla którego $\psi_1(t) = \psi(t)$, jest eksponencjalny rozkład czasów oczekiwania procesu Poissona [12].

Według [61] jest to również warunek słabej ergodyczności procesu. Argument umożliwiający zrozumienie tego warunku jest następujący. Aby średnia po zespole mogła być równa średniej po czasie, czas uśredniania musi być odpowiednio długi. Jeśli układ nie ma charakterystycznego, tzn. średniego czasu między skokami, to dowolnie długi czas uśredniania nigdy nie będzie wystarczający.

Wyprowadzenie przedstawione w poprzednim paragrafie daje się łatwo uogólnić na powyżej opisany przypadek innego traktowania sytuacji początkowej. Zmianie ulega wówczas zależność między propagatorami wyrażona wzorem (3.3); przyjmuje ona postać

$$P(X, t) = \int_0^t [\delta(X) \Psi_1(t - \tau) + Q(X, \tau) \Psi(t - \tau)] d\tau \quad (3.21)$$

oraz (3.6)

$$\tilde{P}(k, s) = \tilde{\Psi}_1(s) + \tilde{\Psi}(s) \tilde{Q}(k, s), \quad (3.22)$$

przy czym propagator $Q_1(X, t)$, obecny we wzorze (3.11), przybiera postać:

$$Q_1(X, t) = \psi_1(t) h(X) \quad (3.23)$$

oraz

$$\tilde{Q}_1(k, s) = \tilde{\psi}_1(s) \tilde{h}(k). \quad (3.24)$$

Natomiast zależność rekurencyjna (3.10) nie ulega zmianie, zatem

$$\tilde{Q}_n(k, s | \xi) = \tilde{\psi}_1(s) \tilde{\psi}^{n-1}(s) \tilde{h}^n(k). \quad (3.25)$$

Po dokonaniu sumowania według wzoru (3.8) oraz wykorzystaniu wyrażenia (3.25) otrzymujemy

$$\tilde{Q}(k; s) = \frac{\tilde{\psi}_1(s) \tilde{h}(k)}{1 - \tilde{h}(k) \tilde{\psi}(s)}. \quad (3.26)$$

Stąd, oraz na podstawie (3.22)

$$\tilde{P}(k, s) = \tilde{\Psi}_1(s) + \tilde{\Psi}(s) \frac{\tilde{\psi}_1(s) \tilde{h}(k)}{1 - \tilde{h}(k) \tilde{\psi}(s)}. \quad (3.27)$$

Jeżeli wariancja

$$\mu_2 = \int_{-\infty}^{\infty} dx x^2 h(x) \quad (3.28)$$

używanego rozkładu skoków jest skończona, to uzyskany w ten sposób w pełni stacjonarny propagator dany wzorem (3.27) umożliwi nam wyprowadzenie drugorzędowych charakterystyk procesu w szczególnie dogodnej postaci.

Zdefiniujmy teraz zależną od czasu wariancję (średnie przesunięcie kwadratowe) procesu

$$m_2(t) = \langle X^2(t) \rangle. \quad (3.29)$$

Możemy otrzymać jej transformatę Laplace'a poprzez dwukrotne różniczkowanie miękkiego, stacjonarnego propagatora (3.27) wyrażonego w zmiennych Fouriera-Laplace'a

$$\tilde{m}_2(s) = - \left. \frac{\partial^2 \tilde{P}(k, s)}{\partial k^2} \right|_{k=0} = \frac{\mu_2}{s^2 \langle t \rangle}, \quad (3.30)$$

gdzie po drodze skorzystaliśmy z wyrażen (3.7), (3.18), (3.19) oraz (3.28). Odwracając tę transformatę Laplace'a, otrzymujemy

$$m_2(t) = 2D|t|, \quad (3.31)$$

czyli liniową zależność średniego przesunięcia kwadratowego od czasu, charakterystyczną dla procesu dyfuzji (procesu Wienera [58]), gdzie współczynnik dyfuzji $D = \frac{\mu_2}{2\langle t \rangle} > 0$. Zauważmy, iż wynik ten nie zależy od konkretnej postaci rozkładów $h(x)$ oraz $\psi(t)$, a jedynie od skończonej wartości średniego czasu wyczekiwania (3.20) oraz skończonej wariancji rozkładu skoków (3.28).

Kolejną charakterystyką drugorzędową, stanowiącą jeden z kluczowych elementów analizy przeprowadzonej w niniejszej pracy, jest funkcja autokorelacji prędkości. Zdefiniujmy chwilową prędkość procesu, przez

$$v(t) = \frac{dX(t)}{dt}. \quad (3.32)$$

Funkcja autokorelacji prędkości dla procesu stacjonarnego (bez dryfu) jest funkcją jednej zmiennej (tzn. różnicy czasów), zdefiniowaną wzorem

$$C(t) = \langle v(0)v(t) \rangle, \quad (3.33)$$

gdzie $\langle \dots \rangle$ oznacza uśrednianie po zespole statystycznym złożonym z wielu realizacji procesu. Średnie przesunięcie kwadratowe $m_2(t)$ łączy z funkcją autokorelacji prędkości $C(t)$ zależność [62]

$$C(t) = \frac{1}{2} \frac{d^2}{dt^2} m_2(t), \quad (3.34)$$

co w przestrzeni Fouriera-Laplace'a daje związek

$$\tilde{C}(s) = \frac{s^2}{2} \tilde{m}_2(s). \quad (3.35)$$

Zatem, podstawiając (3.30) do (3.35) otrzymujemy

$$\tilde{C}(s) = \frac{\mu_2}{2 \langle t \rangle}, \quad (3.36)$$

co po odwróceniu transformaty prowadzi do

$$C(t) = \frac{\mu_2}{2 \langle t \rangle} \delta(t). \quad (3.37)$$

Jak należało oczekiwać otrzymaliśmy wynik charakterystyczny dla procesu dyfuzji, czyli autokorelację prędkości w postaci delty Diraca, w której stała proporcjonalności jest wariancją prędkości procesu. Wynik ten można też uzyskać podstawiając bezpośrednio wyrażenie (3.31) do (3.34) i wykorzystując fakt, że $\frac{d}{dt} m_2(t)$ ma w chwili $t = 0$ skok z $-\frac{\mu_2}{\langle t \rangle}$ do $\frac{\mu_2}{\langle t \rangle}$.

Porównując wzór (3.37) z przebiegami empirycznych funkcji autokorelacji prędkości przedstawionymi na rysunku 2.2 (małe czarne kwadraty) widzimy, że **kanoniczny formalizm CTRW nie jest w stanie odtworzyć ujemnej części przebiegu pokazanego na wykresach.**

Wyniki przedstawione w powyższym rozdziale stanowią jedynie punkt odniesienia do uogólnień modelu zaproponowanych w dalszej części pracy.

Rozdział 4

Stacjonaryzacja modelu Montero i Masolivera

4.1 Geneza modelu

Poprzedni rozdział można podsumować stwierdzeniem, że kanoniczny formalizm CTRW oparty o wyrażenie (1.1) nie jest wystarczający do opisanie stochastycznej dynamiki cen waloru giełdowego. Stwierdzenie to bazuje na wynikach analizy empirycznej przedstawionych na rysunku 2.2, gdzie dobrze widoczne są ujemne autokorelacje (patrz małe czarne kwadraty). Ponadto, jak wspomniałem już we Wprowadzeniu do niniejszej rozprawy, zjawisko występowania silnej korelacji dwóch kolejnych zmian cen było szeroko omawiane w literaturze ekonometrycznej [26, 31]. Dodatkowo, pojawiały się podejścia uwzględniające zależność dwóch kolejnych okresów wyczekiwania, oparte głównie na modelu Engle’a i Russella zwanym *Autoregressive conditional duration*¹ [63].

W kontekście finansowym, rezygnacja z założeń tkwiących u podstaw wyrażenia (1.1), będącego fundamentem kanonicznego CTRW, pojawiła się w pracach [42, 43]. Według autorów korelacje kolejnych przeskoków są wyraźnie silniejsze od korelacji czasów wyczekiwania i głównie w tych pierwszych upatruje się źródła niezgodności przewidywań kanonicznego formalizmu CTRW z danymi empirycznymi. Dlatego wyrażenie (1.1) zostało zastąpione przez wyrażenie (1.2). Należy podkreślić, iż w swoich rozważaniach [64] autorzy przyjmują, że istotne w dynamice są jedynie zmiany cen o ich najmniejszą możliwą wartość, czyli tzw. tick. Zatem, w pracy [42] przyjmują, że rozkład zmian cen przybiera postać: dwustanową

$$h(x) = \frac{\delta(x - c) + \delta(x + c)}{2}, \quad (4.1)$$

oznaczającą, że dyskretna wartość ticku $\pm c$ jest osiągnięta z jednakowym prawdopodobieństwem $1/2$ albo trzystanową

$$h(x) = \frac{1 - Q}{2} \delta(x - c) + Q \delta(x) + \frac{1 - Q}{2} \delta(x + c). \quad (4.2)$$

¹Brak jest powszechnie akceptowanego tłumaczenia. W wolnym tłumaczeniu oznaczałoby to autoregresyjny model warunkowych czasów międzyzdarzeniowych.

W przypadku trzystanowym, zmienna losowa X z prawdopodobieństwem Q , które jest kolejnym parametrem rozkładu, przyjmuje wartość 0, natomiast dyskretna wartość ticku $\pm c$ jest przyjmowana z równym prawdopodobieństwem wynoszącym $(1 - Q)/2$.

Dla takiego dwu- lub trzystanowego rozkładu skoków, rozkład warunkowy $h(r_n | r_{n-1})$ skoku r_n , pod warunkiem ustalonej wartości skoku go poprzedzającego r_{n-1} , wyraża się odpowiednio, przez macierz o rozmiarze 2x2 albo 3x3. Dla tych macierzy zaproponowano ich jawną postać (patrz wzór (31) w pracy [42]). W ramach takiego ujęcia zbadano jedynie średni czas ucieczki ceny z zadanego przedziału.

W pracy [43] rozwinięto powyższe podejście w ramach założenia o istotności jedynie najmniejszych zmian ceny. Autorzy tej pracy postulują ogólniejszą postać zależności dwóch kolejnych zmian cen dla dowolnego rozkładu pojedynczego skoku²

$$h(r_n | r_{n-1}) = h(r_n)(1 + \epsilon \operatorname{sgn}(r_n) \operatorname{sgn}(r_{n-1})), \quad (4.3)$$

gdzie $h(r_n)$ jest rozkładem pojedynczego skoku, będącym funkcją parzystą $h(r_n) = h(-r_n)$. Można zauważyć, że we wzorze (4.3) zależność między kolejnymi skokami występuje jedynie poprzez ich znak, gdzie parametr $\epsilon \in [-1, 1]$ jest miarą tej zależności. Dla $\epsilon > 0$ mamy do czynienia z ujemnym sprzężeniem zwrotnym, dla $\epsilon < 0$ z dodatnim sprzężeniem zwrotnym, natomiast dla $\epsilon = 0$ kolejne skoki ceny są niezależne.

Autorom udało się uzyskać zamkniętą postać propagatora procesu opartego o wyrażenie (1.2) i (4.3) w wersji niestacjonarnej oraz średnie przesunięcie kwadratowe w funkcji czasu. Jednakże, ze względu na niestacjonarność uzyskanego procesu, nie można było skorzystać ze wzoru (3.34) i wyznaczyć funkcji autokorelacji prędkości.

W poniższym rozdziale przedstawiam inne od stosowanego w pracy [43] (przedstawione w pracy [45]) wyprowadzenie zamkniętej postaci propagatora, które daje możliwość zastosowania procedury stacjonaryzacji modelu opisanej w podrozdziale 3.2. **Dzięki uzyskaniu propagatora w wersji stacjonarnej mogłem wyprowadzić jawną postać funkcji autokorelacji prędkości. Stanowi to jeden z kluczowych elementów strategii postępowania przyjętej w niniejszej pracy.**

4.2 Pamięć jednokrokowa w ramach CTRW

Rozważam teraz model CTRW z pamięcią jednokrokową zdefiniowaną poprzez wyrażenie (1.2). Szukając ostrego propagatora, należy zwrócić uwagę na dwa aspekty. Po pierwsze, propagatora $Q_n(X, t)$ nie można już połączyć z propagatorem $Q_{n-1}(X, t)$ zależnością (3.9), gdyż propagator $Q_{n-1}(X, t)$ nie niesie ze sobą informacji o ostatnim skoku, od czego zależy przecież rozkład kolejnego skoku. Po drugie, istotne jest

²Zmieniony został tutaj, w stosunku do oryginału, znak przed parametrem ϵ , co podyktowane jest możliwością łatwiejszego wykazania zgodności modelu przedstawionego w tym rozdziale z modelem opisanym w kolejnym rozdziale.

jaki skok wykonał proces przybywając do $X = 0$ w chwili $t = 0$ (w wersji stacjonarnej - przed chwilą $t = 0$), gdyż od tego zależy rozkład pierwszego skoku. Aby uwzględnić oba powyższe fakty, definiuję nowy propagator. Mianowicie:

- $Q_n(X, r_n; t | \xi)$ - gęstość prawdopodobieństwa tego, że **dokładnie w chwili t proces wykona n -ty przeskok** z wartości $X - r_n$ do wartości X (wykonując skok o r_n), pod warunkiem, że w chwili początkowej ($t = 0$) był w zerze ($X = 0$), do którego przybył wykonując (przedpoczątkowy) skok ξ . Jak poprzednio, interesujemy się tutaj czasem $t > 0$.

Ten nowo zdefiniowany propagator można połączyć z wcześniej wprowadzonym ostrym propagatorem poprzez relację

$$Q_n(X, t) = \int_{-\infty}^{\infty} dr_n \int_{-\infty}^{\infty} d\xi Q_n(X, r_n; t | \xi) h(\xi), \quad n = 1, 2, \dots, \quad (4.4)$$

będącą uśrednieniem po skoku ξ . Używam też uboższych propagatorów zależnych tylko od jednej z wymienionych zmiennych r_n albo ξ

$$Q_n(X, t | \xi) = \int_{-\infty}^{\infty} dr_n Q_n(X, r_n; t | \xi) \quad (4.5)$$

oraz

$$Q_n(X, r_n; t) = \int_{-\infty}^{\infty} d\xi Q_n(X, r_n; t | \xi) h(\xi), \quad (4.6)$$

gdzie $n = 1, 2, \dots$. Dla dowolnej postaci $h(r_n | r_{n-1})$, przy użyciu nowozdefiniowanego propagatora, mogę zapisać równanie rekurencyjne, analogiczne do równania (3.9), w postaci

$$Q_n(X, r_n; t | \xi) = \int_0^t dt' \psi(t') \int_{-\infty}^{\infty} dr_{n-1} h(r_n | r_{n-1}) Q_{n-1}(X - r_n, r_{n-1}; t - t' | \xi), \quad n = 2, 3, \dots \quad (4.7)$$

Tak jak poprzednio, jeżeli proces dociera do punktu X w chwili t , wykonując n -ty skok o r_n , to znaczy, że do punktu $X - r_n$ dotarł w chwili $t - t'$ w $n - 1$ skokach. Powyższe równanie uwzględnia jednak więcej informacji, a mianowicie wszystkie możliwe długości skoku r_{n-1} , a przez to odpowiadające im wszystkie możliwe wartości $h(r_n | r_{n-1})$. Ponieważ zmienne w powyższym równaniu mogą być ciągłe, zależność ta jest ogólniejsza od używanych w pracach [33–36] w przypadku błędzeń losowych na sieciach.

Przeprowadzamy teraz transformatę Laplace'a powyższego równania

$$\frac{1}{\tilde{\psi}(s)} \tilde{Q}_n(X, r_n; s | \xi) = \int_{-\infty}^{\infty} dr_{n-1} h(r_n | r_{n-1}) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi). \quad (4.8)$$

Powyższe równania (4.7) i (4.8) mają charakter ogólny, tzn. są prawdziwe dla dowolnej postaci $h(r_n | r_{n-1})$, w związku z tym, wyników tego podrozdziału używam również w kolejnym rozdziale 5, w którym przedstawiam autorski model o pamięci jednokrokowej.

4.3 Rozwiązanie modelu stacjonarnego

Ograniczam się teraz do postaci $h(r_n | r_{n-1})$ danej równaniem (4.3). W zależności tej, wystarczającą informacją potrzebną do znajomości rozkładu kolejnego skoku jest zwrot skoku poprzedzającego. Zgodnie z tym, dokonajmy następującego podziału wyrażenia (4.8)

$$\begin{aligned} \frac{1}{\tilde{\psi}(s)} \int_0^{\infty} dr_n \tilde{Q}_n(X, r_n; s | \xi) &= \\ &= \int_0^{\infty} dr_n \int_0^{\infty} dr_{n-1} h(r_n | r_{n-1}) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi) + \\ &\quad + \int_0^{\infty} dr_n \int_{-\infty}^0 dr_{n-1} h(r_n | r_{n-1}) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi) \end{aligned} \quad (4.9)$$

oraz

$$\begin{aligned} \frac{1}{\tilde{\psi}(s)} \int_{-\infty}^0 dr_n \tilde{Q}_n(X, r_n; s | \xi) &= \\ &= \int_{-\infty}^0 dr_n \int_0^{\infty} dr_{n-1} h(r_n | r_{n-1}) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi) + \\ &\quad + \int_{-\infty}^0 dr_n \int_{-\infty}^0 dr_{n-1} h(r_n | r_{n-1}) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi). \end{aligned} \quad (4.10)$$

Dla konkretnej postaci $h(r_n | r_{n-1})$ danej przez wzrażenie (4.3), powyższe wzory przyjmują postać, odpowiednio

$$\begin{aligned} \frac{1}{\tilde{\psi}(s)} \int_0^{\infty} dr_n \tilde{Q}_n(X, r_n; s | \xi) &= \\ &= (1 + \epsilon) \int_0^{\infty} dr_n \int_0^{\infty} dr_{n-1} h(r_n) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi) + \\ &+ (1 - \epsilon) \int_0^{\infty} dr_n \int_{-\infty}^0 dr_{n-1} h(r_n) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi) \end{aligned} \quad (4.11)$$

oraz

$$\begin{aligned} \frac{1}{\tilde{\psi}(s)} \int_{-\infty}^0 dr_n \tilde{Q}_n(X, r_n; s | \xi) &= \\ &= (1 - \epsilon) \int_{-\infty}^0 dr_n \int_0^{\infty} dr_{n-1} h(r_n) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi) + \\ &+ (1 + \epsilon) \int_{-\infty}^0 dr_n \int_{-\infty}^0 dr_{n-1} h(r_n) \tilde{Q}_{n-1}(X - r_n, r_{n-1}; s | \xi). \end{aligned} \quad (4.12)$$

Jak widać z postaci zależności (4.11) i (4.12) propagator $\tilde{Q}_n(X, r_n; s | \xi)$ niesie nadmierną informację, gdyż po obu stronach równania występuje on całkowany po zmiennej r_n wzdłuż dodatniej albo ujemnej półosi rzeczywistej. Wynika to wprost z faktu, że w zależności $h(r_n | r_{n-1})$ danej równaniem (4.3), istotny jest jedynie znak poprzedzającego skoku r_{n-1} .

Aby uprościć zapis powyższych zależności wprowadzam połówkową transformatę Fouriera rozkładu h

$$\tilde{H}(k) = \int_0^{\infty} dr e^{ikr} h(r), \quad \tilde{h}(k) = \tilde{H}(k) + \tilde{H}(-k), \quad (4.13)$$

a zamiast propagatora $Q_n(X, r_n; t | \xi)$ wprowadzam propagatory $Q_n^{\pm}(X, t | \xi)$ niosące jedynie informację o zwrocie ostatniego wykonanego skoku, zdefiniowane równaniami

$$Q_n^+(X, t | \xi) = \int_0^{\infty} dr_n Q_n(X, r_n; t | \xi), \quad (4.14)$$

$$Q_n^-(X, t | \xi) = \int_{-\infty}^0 dr_n Q_n(X, r_n; t | \xi). \quad (4.15)$$

Przechodząc do zmiennych Fouriera-Laplace'a, możemy równania (4.11) oraz (4.12) zapisać w postaci macierzowej dla $n \geq 2$

$$\begin{pmatrix} \tilde{Q}_n^+(k, s | \xi) \\ \tilde{Q}_n^-(k, s | \xi) \end{pmatrix} = \tilde{\psi}(s) \begin{pmatrix} \tilde{H}(k)(1 + \epsilon) & \tilde{H}(k)(1 - \epsilon) \\ \tilde{H}(-k)(1 - \epsilon) & \tilde{H}(-k)(1 + \epsilon) \end{pmatrix} \begin{pmatrix} \tilde{Q}_{n-1}^+(k, s | \xi) \\ \tilde{Q}_{n-1}^-(k, s | \xi) \end{pmatrix}. \quad (4.16)$$

Zauważmy, że ostre propagatory dla pierwszego kroku można wyrazić w postaci

$$Q_1(X, t | \xi) = h(X | \xi)\psi_1(t) \quad (4.17)$$

oraz

$$Q_1(X, r_1; t | \xi) = \delta(X - r_1)h(X | \xi)\psi_1(t). \quad (4.18)$$

Uśredniając równanie (4.18) po ξ z wagą $h(\xi)$ (według wzoru (4.6)), a następnie korzystając z (4.3) otrzymujemy

$$Q_1(X, r_1; t) = \delta(X - r_1)h(X)\psi_1(t). \quad (4.19)$$

Stąd

$$\tilde{Q}_1^+(k, s) = \tilde{\psi}_1(s)\tilde{H}(k), \quad (4.20)$$

$$\tilde{Q}_1^-(k, s) = \tilde{\psi}_1(s)\tilde{H}(-k). \quad (4.21)$$

Wprowadźmy teraz sumacyjną reprezentację

$$\tilde{Q}^\pm(k, s) = \sum_{n=1}^{\infty} \tilde{Q}_n^\pm(k, s). \quad (4.22)$$

Łatwo dostrzec, iż spełniona jest równość

$$\tilde{Q}(k, s) = \tilde{Q}^+(k, s) + \tilde{Q}^-(k, s). \quad (4.23)$$

Dokonajmy teraz sumowania stronami równania (4.16) po n wraz z uśrednieniem po ξ według wzoru (4.6)

$$\begin{aligned} \begin{pmatrix} \tilde{Q}^+(k, s) \\ \tilde{Q}^-(k, s) \end{pmatrix} - \tilde{\psi}_1(s) \begin{pmatrix} \tilde{H}(k) \\ \tilde{H}(-k) \end{pmatrix} &= \\ &= \tilde{\psi}(s) \begin{pmatrix} \tilde{H}(k)(1 + \epsilon) & \tilde{H}(k)(1 - \epsilon) \\ \tilde{H}(-k)(1 - \epsilon) & \tilde{H}(-k)(1 + \epsilon) \end{pmatrix} \begin{pmatrix} \tilde{Q}^+(k, s) \\ \tilde{Q}^-(k, s) \end{pmatrix}. \end{aligned} \quad (4.24)$$

Dokonując elementarnych przekształceń algebraicznych uzyskujemy

$$\begin{aligned} \begin{pmatrix} \tilde{Q}^+(k, s) \\ \tilde{Q}^-(k, s) \end{pmatrix} &= \\ &= \frac{1}{\tilde{\psi}_1(s)} \begin{pmatrix} 1 - \tilde{\psi}(s)\tilde{H}(k)(1 + \epsilon) & -\tilde{\psi}(s)\tilde{H}(k)(1 - \epsilon) \\ -\tilde{\psi}(s)\tilde{H}(-k)(1 - \epsilon) & 1 - \tilde{\psi}(s)\tilde{H}(-k)(1 + \epsilon) \end{pmatrix}^{-1} \begin{pmatrix} \tilde{H}(k) \\ \tilde{H}(-k) \end{pmatrix}, \end{aligned} \quad (4.25)$$

gdzie $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$ oznacza macierz odwrotną do $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Z tej postaci uzyskujemy miękki propagator

$$\begin{aligned} \tilde{P}(k, s) &= \tilde{\Psi}(s) \left(\tilde{Q}^+(k, s) + \tilde{Q}^-(k, s) \right) + \tilde{\Psi}_1(s) = \\ &= \frac{1 - \epsilon \tilde{h}(k) \tilde{\psi}(s) + \left(\tilde{\psi}_1(s) - \tilde{\psi}(s) \right) \left(\tilde{h}(k) - 4\epsilon \tilde{H}(k) \tilde{H}(-k) \tilde{\psi}(s) \right)}{1 - \tilde{h}(k)(1 + \epsilon) \tilde{\psi}(s) + 4\epsilon \tilde{H}(k) \tilde{H}(-k) \tilde{\psi}(s)^2} \tilde{\Psi}(s) + \\ &\quad + \left(\tilde{\Psi}_1(s) - \tilde{\Psi}(s) \right). \end{aligned} \quad (4.26)$$

Uzyskany w ten sposób propagator opiera się na założeniu, że rozkład czasów oczekiwania przed pierwszym skokiem dany jest przez $\psi_1(t)$. Zatem, z technicznego punktu widzenia, aby odtworzyć wynik uzyskany w [43] dla przypadku niestacjonarnego, w którym pierwszy skok również podlega rozkładowi $\psi(t)$, wystarczy we wzorze (4.26) zamienić po prostu $\psi_1(t)$ na $\psi(t)$.

Posiadając stacjonarny, miękki propagator dany równaniem (4.26) oraz korzystając z (3.30) obliczam transformatę Laplace'a zależnego od czasu, średniego przesunięcia kwadratowego. Zatem,

$$\tilde{m}_2(s) = - \left. \frac{\partial^2 \tilde{P}(k, s)}{\partial k^2} \right|_{k=0} = \frac{\mu_2 \tilde{\psi}_1(s)}{s(1 - \tilde{\psi}(s))} + \frac{2\epsilon \mu_1^2 \tilde{\psi}(s) \tilde{\psi}_1(s)}{s(1 - \tilde{\psi}(s))(1 - \epsilon \tilde{\psi}(s))}, \quad (4.27)$$

gdzie μ_2 dane jest wzorem (3.28), natomiast

$$\mu_1 = \int_{-\infty}^{\infty} dx |x| h(x). \quad (4.28)$$

Wracając w równaniu (4.27) do przestrzeni czasu uzyskujemy

$$m_2(t) = \frac{\mu_2}{\langle t \rangle} t + \frac{2\epsilon \mu_1^2}{\langle t \rangle} \mathcal{L}^{-1} \left\{ \frac{\tilde{\psi}(s)}{s^2(1 - \epsilon \tilde{\psi}(s))} \right\} (t), \quad (4.29)$$

gdzie \mathcal{L}^{-1} oznacza odwrotną transformatę Laplace'a. Zauważmy, że jeżeli średni czas pomiędzy skokami $\langle t \rangle$ oraz drugi moment rozkładu skoków μ_2 istnieją (czyli są skończone), to niezależnie od postaci $\psi(t)$ i wartości ϵ , asymptotyczne zachowanie średniego przesunięcia kwadratowego jest dyfuzyjne, ze stałą dyfuzji

$$D = \frac{\mu_2}{2 \langle t \rangle}, \quad (4.30)$$

taką samą jak dla kanonicznego modelu CTRW.

Kształt odstępstwa od zachowania dyfuzyjnego, zawartego w drugim składniku w wyrażeniu (4.29), jest rządzony rozkładem czasów oczekiwania $\psi(t)$ oraz parametrem ϵ , natomiast skalowane jest ono parametrem $\frac{2\epsilon \mu_1^2}{\langle t \rangle}$.

Możemy teraz wyznaczyć, dla badanego procesu stochastycznego, funkcję autokorelacji prędkości korzystając ze wzoru (3.34)

$$C(t) = \frac{\mu_2}{2\langle t \rangle} \delta(t) + \frac{\epsilon \mu_1^2}{\langle t \rangle} \mathcal{L}^{-1} \left\{ \frac{\tilde{\psi}(s)}{1 - \epsilon \tilde{\psi}(s)} \right\} (t) \quad (4.31)$$

lub w postaci unormowanej

$$C^n(t) = \delta(t) + \frac{\mu_1^2}{\mu_2} 2\epsilon \mathcal{L}^{-1} \left\{ \frac{\tilde{\psi}(s)}{1 - \epsilon \tilde{\psi}(s)} \right\} (t). \quad (4.32)$$

Podobnie jak w przypadku równania (4.29), także w powyższym równaniu jest obecny drugi składnik, odróżniający wyrażenie (4.32) od postaci czysto dyfuzyjnej. Można przypuszczać, że mógłby on odtwarzać zachowanie funkcji autokorelacji prędkości dla $t > 0$, pokazane na rysunku 2.2. Porównanie wyników empirycznych z przewidywaniami wzoru (4.32), jak również z przewidywaniami wprowadzonego kolejnego modelu mojego autorstwa, przedstawione jest w następnym rozdziale niniejszej pracy.

Rozdział 5

Model z pamięcią jednokrokową

W niniejszym rozdziale przedstawiam już własny model oparty na formalizmie CTRW. Model ten rozwinąłem w oparciu o założenia, które legły u podstaw równania (1.2) i używane były w pracach [42, 43]. Należy podkreślić, iż model ten również posiada pamięć jednokrokową. Sformułowanie takiego modelu, dla dowolnego rozkładu pojedynczego skoku $h(r_n)$ i dowolnego rozkładu czasów oczekiwania $\psi(t)$, sprowadza się, w istocie rzeczy, do zapostulowania rozkładu warunkowego $h(r_n | r_{n-1})$, czyli zależności występującej między dwoma kolejnymi zmianami ceny. Motywacją do stworzenia tego modelu są zarówno dane empiryczne, jak też przesłanki teoretyczne - omawiam je systematycznie poniżej.

5.1 Motywacja empiryczna modelu

Podstawą empiryczną rozważań zawartych w tym rozdziale są publicznie dostępne dane transakcyjne wysokiej częstotliwości dla spółek notowanych na warszawskiej Giełdzie Papierów Wartościowych. Zostały one opisane w rozdziale 2.1. Dysponując danymi tego typu, możemy sporządzić m.in. dwuwymiarowy histogram empiryczny dwóch następujących bezpośrednio po sobie zmian cen, co przedstawiono na rysunku 5.1. Odstęp czasu pomiędzy tymi zmianami oczywiście fluktuuje i nie jest obrazowany na tym histogramie. Dzięki założeniom tkwiącym w wyrażeniu (1.2), histogram ten powinien być statystyczną realizacją rozkładu $h(r_n, r_{n-1})$. Stanowi on empiryczną inspirację pozwalającą zaproponować funkcyjną postać rozkładu warunkowego $h(r_n | r_{n-1})$.

Rysunek 5.1 pokazuje, że empiryczna gęstość prawdopodobieństwa dwóch kolejnych zmian cen $h(r_n, r_{n-1})$ zbudowana jest z dwóch głównych składników:

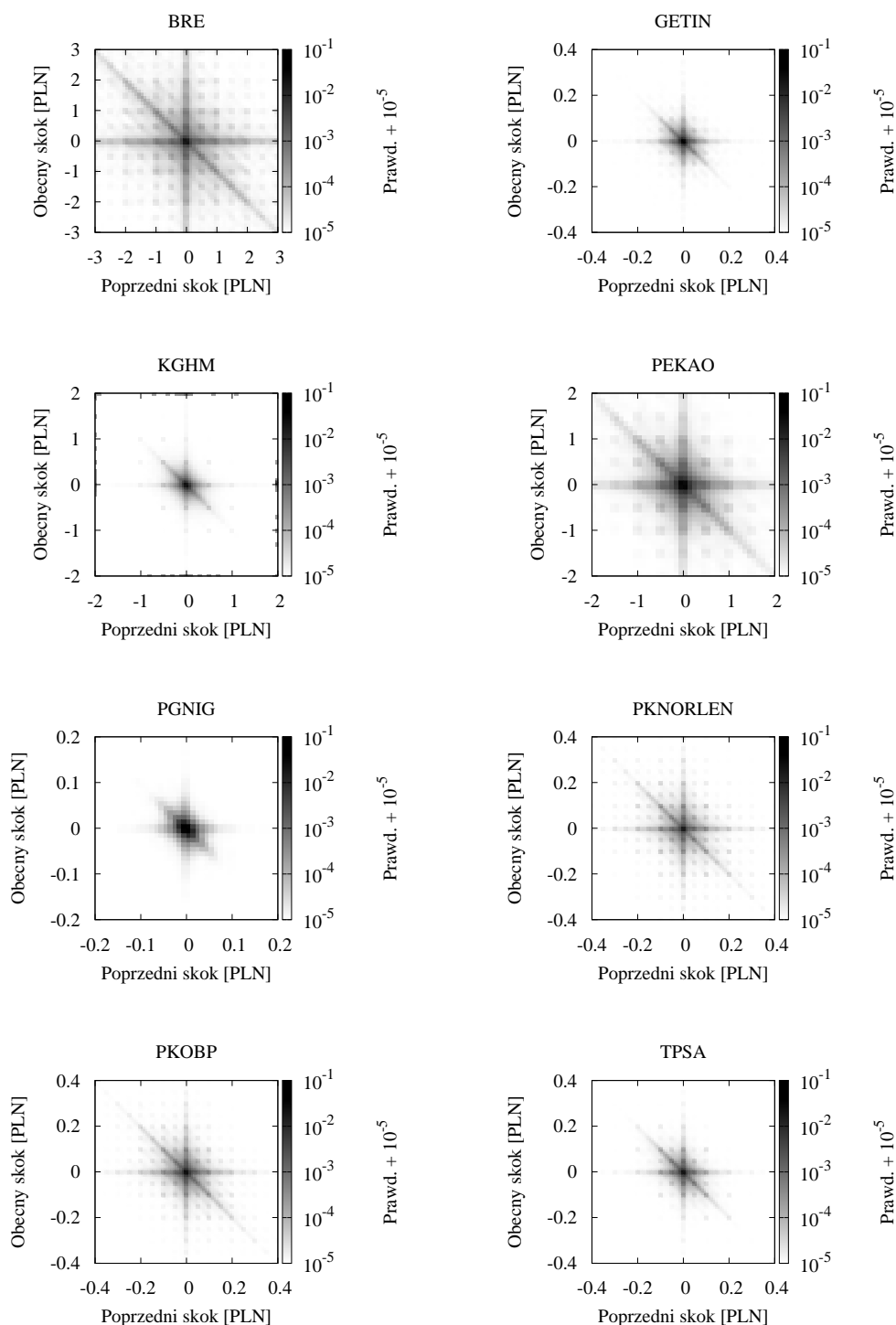
- (i) z centralnego krzyża zdefiniowanego przez punkty $(r_{n-1}, 0)$ i $(0, r_n)$, co potwierdza tylko powszechną obserwację, że co najmniej jedna z dwóch kolejnych transakcji może pojawić się bez zmiany ceny,
- (ii) z punktów $(r_{n-1}, r_n = -r_{n-1} \pm \eta)$ należących do antydiagonali, których zbiór definiuje w przybliżeniu składnik proporcjonalny do delty Diraca $\delta(r_{n-1} + r_n)$.

Delta ta oznacza, że dwa kolejne skoki mają przeciwne zwroty lecz w przybliżeniu tą samą długość, gdyż η jest małe¹. Taką zależność pomiędzy dwoma kolejnymi skokami nazywamy ostrą korelacją powrotną. W szczególności ma to miejsce dla wartości skoków znacznie większych od wartości pojedynczego ticku. Warto zwrócić uwagę, że analogiczny histogram wykonany dla stóp zwrotu lub logarytmicznych stóp zwrotu jest bardziej rozmyty, co uniemożliwia użycie (wygodnej w dalszych obliczeniach) delty Diraca.

Pierwsza obserwacja (i) jest bezpośrednią konsekwencją pojawiania się dużej liczby transakcji bez zmiany ceny. Obserwacja ta nie świadczy o żadnej zależności pomiędzy dwoma kolejnymi skokami, gdyż występowałaby ona również w przypadku gdyby dwie kolejne zmiany ceny były całkowicie niezależne.

Druga obserwacja (ii) jest reminiscencją na poziomie mikroskopowym analogonu reguły Le Chateliera-Brauna, jakim jest tutaj zjawisko bid-ask bounce. Zjawisko to oznacza, mówiąc w uproszczeniu, odbijanie się ceny transakcyjnej od dolnej do górnej granicy bid-ask spreadu i z powrotem. Powtarza się to wielokrotnie w czasie, gdy sam bid-ask spread fluktuuje. W stosunku do przypadku niezależnych skoków, zjawisko to skutkuje zwiększonym łącznym prawdopodobieństwem dwóch kolejnych skoków o przeciwnych zwrotach, ale równej lub prawie równej długości. Mechanizm tego zjawiska nie jest bezpośrednio widoczny na poziomie cen transakcyjnych, gdzie obserwować możemy jedynie jego skutki, lecz dopiero na poziomie księgi zleceń. Niestety, dane z księgi zleceń umożliwiające bezpośrednią obserwację zjawiska bid-ask bounce nie są dla warszawskiej Giełdy Papierów Wartościowych gromadzone. W kolejnym podrozdziale opisuję mechanizm powstawania tego zjawiska.

¹Wielkość η definiuje rozmiar kwadratów widocznych w tle rysunku 5.1.

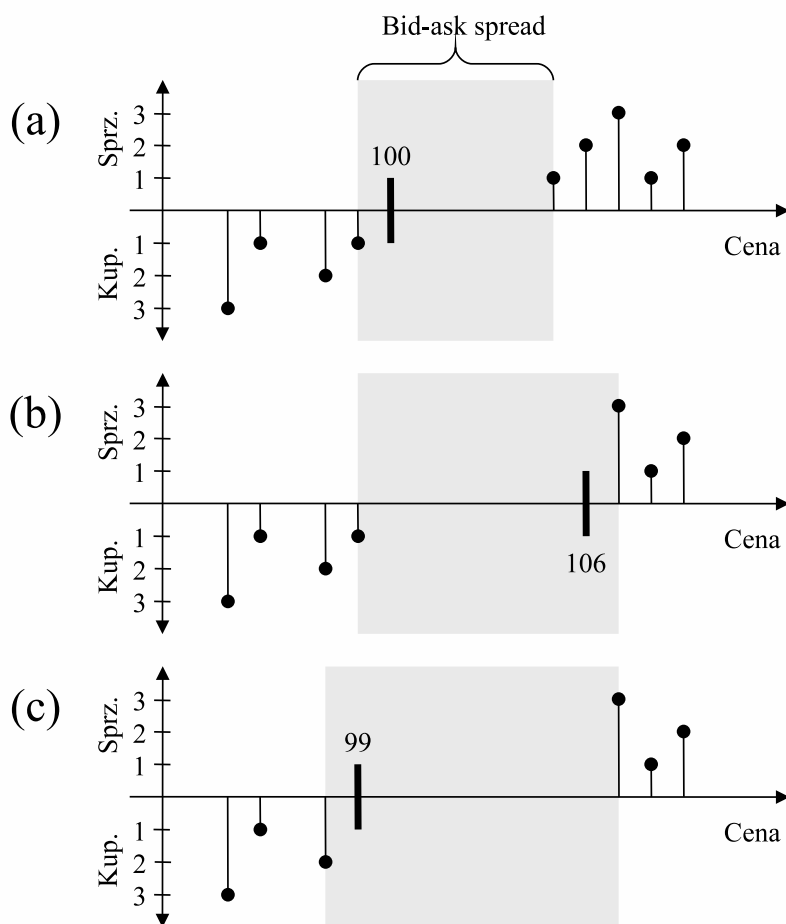


Rysunek 5.1: Histogramy empiryczne dwóch kolejnych zmian ceny dla ośmiu badanych spółek. Wielkości skoków oznaczone są na osiach, natomiast odcieniem szarości oznaczona jest częstotliwość wystąpienia danej pary (ciemniejszy odcień szarości odpowiada wyższej częstotliwości). Skala szarości na rysunku jest logarytmiczna, zatem aby usunąć problem osobliwości w zerze wszystkie częstotliwości zwiększono o stałą wartość 10^{-5} . Dobrze widoczne są dwa charakterystyczne elementy: centralny krzyż i antydiagonała, różniące się jednak natężeniem (częstotliwością występowania danej pary skoków).

5.2 Motywacja teoretyczna: zjawisko bid-ask bounce i jego mechanizm

Rozważamy teraz przykład, który ilustruje sposób działa mechanizmu bid-ask bounce. Jest to przykład kluczowy dla niniejszej pracy. Pozwoli on na wskazanie, jaka sekwencja zleceń prowadzi do korelacji powrotnych kolejnych skoków ceny danego waloru.

Założmy, że początkowo mamy do czynienia z księgą zleceń zawierającą różnego typu zlecenia, przedstawioną na rysunku 5.2a. Założmy dalej, że ostatnia cena trans-



Rysunek 5.2: Schematyczna ilustracja księgi zleceń w trzech kolejnych fazach (a), (b), (c), przedstawiająca mechanizm zjawiska bid-ask bounce. Na każdym z trzech schematów zlecenia kupna (pionowe kreski z kółkami zwrócone w dół znajdujące się po lewej stronie) oddzielone są od zleceń sprzedaży (analogiczne pionowe kreski z kółkami po prawej stronie) bid-ask spreadem (szary obszar). Cena poprzedzającej transakcji oznaczona jest pogrubioną kreską podpisaną ceną. Punkty powyżej poziomej osi oznaczającej limit ceny, mówią o całkowitym wolumenie ofert sprzedaży z danym limitem (oś Sprz.). Analogicznie, punkty poniżej osi ceny oznaczają całkowity wolumen ofert kupna z danym limitem (oś Kup.). Minimalne odległości pomiędzy najbliższymi pionowymi kreskami wynoszą 1 w jednostkach umownych.

akcyjna, zaznaczona na rysunku pogrubioną kreską, miała miejsce przy cenie równej 100 umownych jednostek. Przypuśćmy, dla przykładu, że pojawiła się rynkowa oferta kupna 3 akcji². Oferta zostanie natychmiastowo zrealizowana, a cena wzrośnie do 106 umownych jednostek (patrz rysunek 5.2a i 5.2b). Z tego powodu prawa krawędź bid-ask spreadu przesunęła się do wartości 107 jednostek podczas gdy jego lewa krawędź pozostała przy wartości 99 (tzn. cenie najwyższej oferty kupna). Kolejna oferta pojawiająca się na rynku może należeć tylko do jednego z poniższych typów:

- (a) **wykonana oferta kupna**, czyli albo rynkowa oferta kupna albo oferta kupna z limitem ceny, który jest równy lub wyższy od prawej krawędzi bid-ask spreadu. Zauważmy, że oferty sprzedaży z limitem wyższym niż ostatnia cena transakcyjna, tzn. 106 jednostek, były już obecne w księdze zleceń. Możemy zatem z dobrym przybliżeniem założyć, że zmiana ceny wywołana takim zleceniem (zawsze dodatnia) jest niezależna od zmiany ją poprzedzającej.
- (b) **Wewnętrzna oferta kupna**, czyli oferta kupna z limitem ceny znajdującym się wewnątrz bid-ask spreadu. Oferta ta nie może być natychmiastowo zrealizowana - jej możliwa realizacja jest opóźniona w czasie. Zmniejsza ona bid-ask spread przesuwając jego lewą krawędź w prawo, pozostawiając jednocześnie ostatnią cenę transakcyjną, tj. 106 jednostek, wewnątrz nowego, zmniejszonego bid-ask spreadu. Oznacza to, że następna transakcja może być efektem realizacji albo wspomnianej wewnętrznej oferty kupna albo ofert sprzedaży czekających na realizację a znajdujących się po prawej stronie ostatniej ceny transakcyjnej. Stąd, następna zmiana ceny transakcyjnej może być zarówno dodatnia jak i ujemna z jednakowym (w dobrym przybliżeniu) prawdopodobieństwem. Możemy więc przyjąć, że dla tego typu transakcji nie ma zależności pomiędzy poprzedzającą a obecną zmianą ceny.
- (c) **Wykonana oferta sprzedaży**, czyli albo rynkowa oferta sprzedaży albo oferta sprzedaży z limitem równym bądź niższym od lewej krawędzi bid-ask spreadu. Nawet jeśli taka oferta sprzedaży jest bardzo mała, mająca np. wolumen równy 1 akcji, to transakcja dojdzie do skutku po cenie znajdującej się w otoczeniu lewego brzegu bid-ask spreadu, który tutaj wynosi 99 jednostek (patrz rysunek 5.2b i 5.2c). Dłuższe skoki ceny, tj. odpowiadające znacznie niższym cenom transakcyjnym są znacznie mniej prawdopodobne, gdyż oferty o małym wolumenie zdarzają się znacznie częściej. W tym przypadku możemy założyć, że poprzedzająca i obecna zmiana ceny mają w przybliżeniu tą samą długość lecz przeciwne zwroty tj. są ostro ujemnie skorelowane. Tego typu transakcje odgrywają zasadniczą rolę w moich rozważaniach.

²Przez ofertę rynkową rozumiemy ofertę bez limitu ceny, która musi być natychmiastowo zrealizowana, tj. ofertę sprzedaży z limitem 0 lub ofertę kupna z nieskończonym limitem ceny. Na GPW takie zlecenia nazywane są zleceniami 'Po Każdej Cenie' (PKC).

- (d) **Wewnętrzna oferta sprzedaży**, czyli oferta sprzedaży z limitem ceny znajdującym się wewnątrz bid-ask spreadu. Również ta transakcja nie może być natychmiastowo wykonana prowadząc tylko do zawężenia bid-ask spreadu przez przesunięcie jego prawej krawędzi w lewą stronę. W tej sytuacji ostatnia cena transakcyjna, tj. 106 jednostek, znajduje się na prawo od nowego bid-ask spreadu. Zatem, najbardziej prawdopodobną następną ceną transakcyjną jest cena znajdująca się w pobliżu tego nowego bid-ask spreadu. Stąd możemy założyć, że kolejna zmiana ceny będzie ujemnie skorelowana z obecną. Ponadto, wprowadzamy uproszczenie polegające na tym, iż zakładamy, że poprzednia i następna zmiana ceny mają tę samą długość pomimo, że średnio rzecz biorąc cena następnej transakcji znajduje się w przybliżeniu w środku nowego bid-ask spreadu. Przybliżenie to jest dyskutowane poniżej.
- (e) **Pozostałe rodzaje ofert** (np. oferty kupna z limitem ceny poniżej lewej krawędzi bid-ask spreadu, jak też oferty sprzedaży z limitem ceny powyżej prawej krawędzi bid-ask spreadu) w tej analizie mogą być zignorowane, ponieważ w chwili bieżącej nie mają bezpośredniego wpływu na dynamikę ceny. Jeżeli taka nieistotna oferta pojawi się, czekamy na kolejną należącą do jednego z czterech wymienionych wcześniej typów.

Analogiczny mechanizm ma oczywiście miejsce przy założeniu odwrotnego kierunku początkowej zmiany ceny, czyli gdy zaczynamy naszą analizę od realizowanej oferty sprzedaży i poprzedzającej ceny przy prawej krawędzi bid-ask spreadu. Powyższa sekwencja kroków jest ciągle powtarzana i stanowi mechanizm powstawania efektu bid-ask bounce, czyli odbijania się ceny transakcyjnej pomiędzy dwiema krawędziami bid-ask spreadu. W moim modelu zakładam, że pierwsze dwa przypadki (a) i (b) pojawiają się ze stałym prawdopodobieństwem $1 - \epsilon$, gdzie $0 \leq \epsilon \leq 1$, podczas gdy sytuacje (c) i (d) pojawiają się z prawdopodobieństwem ϵ .

Bardziej wyrafinowane podejście wymagałoby oddzielnego traktowania sytuacji z punktu (d), gdyż mamy wówczas do czynienia z szerszym rozkładem możliwych skoków o ujemnej korelacji. Rozkład ten jest bardziej rozmyty niż w sytuacji opisanej w punkcie (c). Jednakże z rozkładu empirycznego przedstawionego na rysunku 5.1, który pokazuje ostro zarysowaną antydiagonalę możemy wnioskować, że przypadek (d) może być przybliżony opisanym powyżej założeniem lub też ma znacznie mniejsze znaczenie.

Warto nadmienić, że mechanizm przedstawiony powyżej działa niezależnie od tego jak duży był początkowy skok ceny transakcyjnej, która w naszym przykładzie wzrosła ze 100 do 106 jednostek. Wystarczy, że skok jest większy od najmniejszej możliwej zmiany ceny. **Interesującą, nadzwyczaj ważną obserwacją jest to, że nawet znacząca, początkowa zmiana ceny transakcyjnej może zostać odwrócona lub prawie odwrócona nawet przez transakcję o znikomym wolumenie. Taka transakcja prowadzi do powrotnego skoku ceny o tej samej lub prawie takiej samej długości co skok poprzedzający.**

Transakcje o dużym wolumenie mogą być rozpatrywane jako duże fluktuacje lub jako działania zewnętrznej siły starającej się wyprowadzić układ z równowagi. Jeśli rozpatrzmy oferty o małym wolumenie jako wewnętrzny szum zawarty w układzie, to z powyższej analizy wynika, iż przywracają one cenę do stanu początkowego, co może być rozpatrywane jako rozszerzenie na ten układ reguły przekory.

Powyższa analiza daje nam podstawy do opisu zależności dwóch kolejnych skoków. Zależność tę, można wyrazić warunkową gęstością prawdopodobieństwa kolejnego skoku r_n , pod warunkiem skoku go poprzedzającego r_{n-1} , czyli $h(r_n | r_{n-1})$. W następnym podrozdziale 5.3 proponuję oryginalną formułę uwzględniającą powyżej opisane zachowania.

5.3 Definicja modelu

Analiza danych empirycznych zawarta w podrozdziale 5.1, a uwidoczniiona na rysunku 5.1 oraz rozważania teoretyczne przedstawione w podrozdziale 5.2, pozwoliły mi zaproponować oryginalną formę warunkowej gęstości prawdopodobieństwa $h(r_n | r_{n-1})$. Mianowicie,

$$h(r_n | r_{n-1}) \stackrel{\text{def.}}{=} (1 - \epsilon) h(r_n) + \epsilon \delta(r_n + r_{n-1}). \quad (5.1)$$

Występujący tutaj rozkład $h(r_n)$ jest niezależną od parametru ϵ funkcją parzystą swojego argumentu co oznacza, że w tym modelu nie uwzględniamy dryfu. Parametr $0 \leq \epsilon \leq 1$ jest stałą wagą, rządzącą stopniem korelacji dwóch kolejnych skoków. W przeciwieństwie do modelu Montero i Masolivera, w którym gęstość $h(r_n | r_{n-1})$ dana jest wzorem (4.3), tutaj rozważamy jedynie ujemne korelacje sterowane dodatnim parametrem ϵ .

Analizując równanie (5.1) można zauważyć, iż składa się ono z dwóch ważonych składników. Pierwszy, posiadający wagę $1 - \epsilon$ odpowiada sytuacji, w której kolejny skok jest niezależny od poprzedniego i losowany z rozkładu $h(r_n)$. Drugi składnik, pojawiający się z prawdopodobieństwem ϵ , odpowiada powrotowi ceny do poprzedniej wartości, czyli ostrej korelacji powrotnej prowadzącej do powtórzenia poprzedniego skoku, ale ze zmienionym zwrotem.

Mnożąc równanie (5.1) stronami przez $h(r_{n-1})$, łatwo uzyskujemy $h(r_n, r_{n-1})$, czyli łączny rozkład prawdopodobieństwa dwóch kolejnych skoków. Teraz, dysponując empirycznym rozkładem (unormowanym histogramem) pojedynczego skoku $h(r_n)$ uzyskanym z danych empirycznych opisanych w podrozdziale 2.1 oraz parametrem ϵ , którego estymacja jest opisana w dalszej części pracy, mogę zbudować teoretyczny odpowiednik wykresów przedstawionych na rysunku 5.1; zamieściłem go na rysunku 5.3.

Warunkowa gęstość prawdopodobieństwa dana wzorem (5.1), lepiej opisuje dostępne dane empiryczne niż ta, zaproponowana w pracy [43] przez Montero i Masolivera, dana w niniejszej pracy równaniem (4.3). Fakt ten wynika z istnienia we

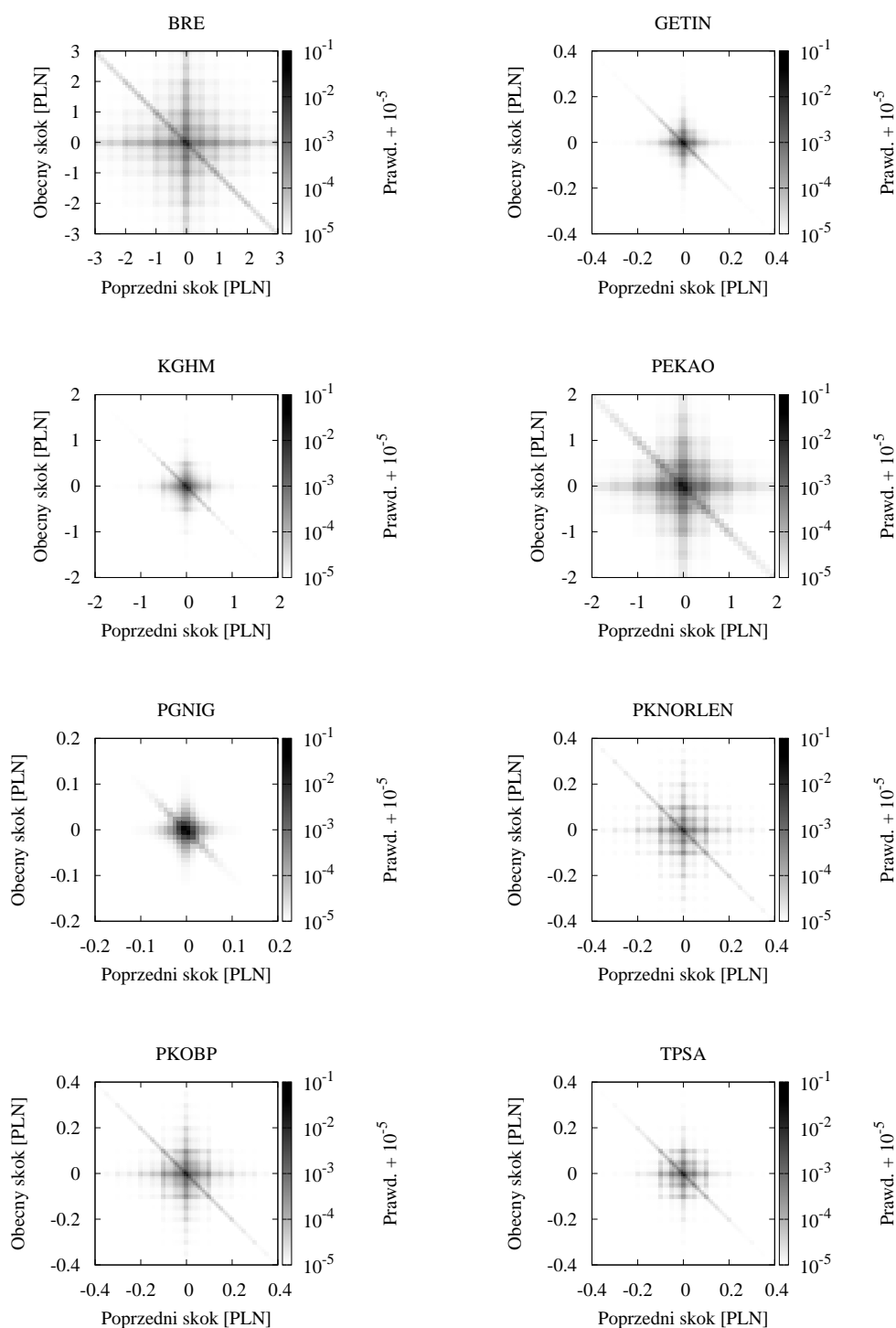
wzorze (5.1) składnika zawierającego deltę Diraca, która faworyzuje skoki powrotne obrazowane przez antydiagonalę wyraźnie widoczną na obu rysunkach 5.1 i 5.3.

Zauważmy, że dla $\epsilon \neq 0$ zaproponowany przeze mnie model oraz model Montero i Masolivera pokrywają się tylko dla szczególnego przypadku. Mianowicie, ma to miejsce tylko wtedy, gdy gęstość rozkładu pojedynczego skoku dana jest przez rozkład (4.1). Opisuje on nierealistyczną sytuację dwustanową dla zmiany ceny transakcyjnej. Oczywiście w przypadku $\epsilon = 0$ obie warunkowe gęstości rozkładu skoków dane równaniami (5.1) i (4.3) stają się identycznymi, bezwarunkowymi rozkładami pojedynczego skoku ceny.

Na zakończenie tego podrozdziału sprawdzimy, że warunkowa gęstość prawdopodobieństwa (5.1) spełnia warunki samozgodności

$$\begin{aligned} h(r_n) &= \int_{-\infty}^{\infty} h(r_n | r_{n-1}) h(r_{n-1}) dr_{n-1}, \\ h(r_{n-1}) &= \int_{-\infty}^{\infty} h(r_n | r_{n-1}) h(r_{n-1}) dr_n. \end{aligned} \quad (5.2)$$

Opierając się teraz na wyrażeniu (1.2), możemy (używając formalizmu CTRW) przystąpić do wyprowadzenia jawnej postaci propagatora procesu z pamięcią jednokrokwą.



Rysunek 5.3: Rozkład teoretyczny dwóch kolejnych zmian ceny generowany z rozkładu (5.1). Przyjęte na rysunku oznaczenia są identyczne jak te na rysunku 5.1. Przyjąłem tutaj empiryczny rozkład pojedynczego skoku (bezwarunkowy). Metoda estymacji parametru ϵ opisana została w podrozdziale 5.6. Widać, że rozkład (5.1) odtwarza zarówno centralny krzyż, jak też antydiagonalę dla wszystkich analizowanych spółek, co (przynajmniej jakościowo) zgadza się z analogicznymi rozkładami empirycznymi przedstawionymi na rysunku 5.1.

5.4 Wyprowadzenie jawnej postaci propagatorów

Rozważany w bieżącym rozdziale model, podobnie jak model Montero i Masolivera, jest modelem o pamięci jednokrokowej. Zatem, zastosowanie mają oznaczenia wprowadzone przeze mnie w podrozdziale 4.2, a przede wszystkim zasadnicza zależność rekurencyjna pomiędzy propagatorami dana równaniem (4.7). Wstawiając postać $h(r_n | r_{n-1})$, daną wyrażeniem (5.1), do równania (4.7) i wykonując transformatę Fouriera-Laplace'a otrzymujemy

$$\begin{aligned} \tilde{Q}_n(k, r_n; s | \xi) &= \tilde{\psi}(s) e^{ikr_n} \times \\ &\times \int_{-\infty}^{\infty} dr_{n-1} [(1 - \epsilon) h(r_n) + \epsilon \delta(r_n + r_{n-1})] \tilde{Q}_{n-1}(k, r_{n-1}; s | \xi). \end{aligned} \quad (5.3)$$

Powyższe równanie jest w moim modelu równaniem bazowym, służącym wyprowadzeniu jawnej postaci propagatora procesu.

Całkując teraz równanie (5.3) stronami po r_n otrzymujemy

$$\frac{\tilde{Q}_n(k; s | \xi)}{\tilde{\psi}(s)} = (1 - \epsilon) \tilde{h}(k) \tilde{Q}_{n-1}(k; s | \xi) + \epsilon \int_{-\infty}^{\infty} dr_{n-1} e^{ik(-r_{n-1})} \tilde{Q}_{n-1}(k, r_{n-1}; s | \xi). \quad (5.4)$$

Całka występująca w drugim składniku po prawej stronie powyższego równania, może być ponownie obliczona z równania bazowego (5.3) (po zamianie indeksu n na $n - 1$ i pomnożeniu obu stron równania przez $e^{ik(-r_{n-1})}$)

$$\begin{aligned} &\int_{-\infty}^{\infty} dr_{n-1} e^{ik(-r_{n-1})} \tilde{Q}_{n-1}(k, r_{n-1}; s | \xi) = \\ &= \tilde{\psi}(s) \int_{-\infty}^{\infty} dr_{n-1} \int_{-\infty}^{\infty} dr_{n-2} [(1 - \epsilon) h(r_{n-1}) + \epsilon \delta(r_{n-1} + r_{n-2})] \tilde{Q}_{n-2}(k, r_{n-2}; s | \xi) = \\ &= \tilde{\psi}(s) \int_{-\infty}^{\infty} dr_{n-2} \tilde{Q}_{n-2}(k, r_{n-2}; s | \xi) = \tilde{\psi}(s) \tilde{Q}_{n-2}(k; s | \xi). \end{aligned} \quad (5.5)$$

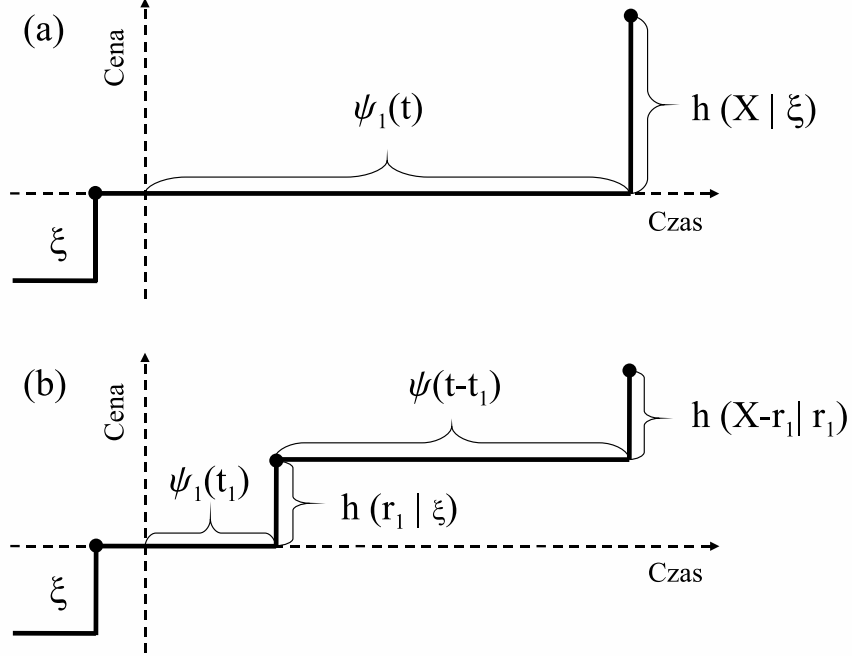
Ostatecznie otrzymujemy zależność rekurencyjną w zmiennych Fouriera-Laplace'a dla $n \geq 3$ w postaci

$$\tilde{Q}_n(k; s | \xi) = (1 - \epsilon) \tilde{h}(k) \tilde{\psi}(s) \tilde{Q}_{n-1}(k; s | \xi) + \epsilon \tilde{\psi}^2(s) \tilde{Q}_{n-2}(k; s | \xi). \quad (5.6)$$

Dokonując sumowania powyższego równania po n od 3 do nieskończoności oraz uzupełniając je odpowiednio o wielkości $\tilde{Q}_1(k, s | \xi)$ i $\tilde{Q}_2(k, s | \xi)$ otrzymujemy, po elementarnych przekształceniach algebraicznych

$$\tilde{Q}(k; s | \xi) = \frac{\tilde{Q}_2(k; s | \xi) + \tilde{Q}_1(k; s | \xi) (1 - \tilde{h}(k)(1 - \epsilon)\tilde{\psi}(s))}{1 - (1 - \epsilon)\tilde{h}(k)\tilde{\psi}(s) - \epsilon\tilde{\psi}^2(s)}. \quad (5.7)$$

Aby uzyskać jawną postać propagatora potrzebujemy jeszcze pierwszych dwóch ostrych propagatorów, czyli $\tilde{Q}_1(k, s | \xi)$ i $\tilde{Q}_2(k, s | \xi)$. Wielkości te możemy wyznaczyć bezpośrednio z definicji. Schemat ich wyprowadzenia, dla dowolnego rozkładu $h(r_n, r_{n-1})$, przedstawia rysunek 5.4.



Rysunek 5.4: Schematyczna ilustracja wyprowadzenia (a) pierwszego ostrego propagatora $Q_1(X, t | \xi)$ i (b) drugiego ostrego propagatora $Q_2(X, t | \xi)$ dla przypadku pamięci jednokrokowej. W obu przypadkach charakterystyczne sekwencje skoków i czasów wyczekiwania oznaczone są nawiasami klamrowymi, przy których podana jest odpowiednia gęstość rozkładu, z którego dana wielkość jest losowana. Ponadto, zaznaczony jest skok ξ mający miejsce przed chwilą początkową $t = 0$.

Zatem, wspomniane propagatory przyjmują następujące postacie

$$Q_1(X, t | \xi) = \psi_1(t)h(X | \xi), \quad (5.8)$$

$$Q_2(X, t | \xi) = \int_0^t dt_1 \psi_1(t_1) \psi(t - t_1) \int_{-\infty}^{\infty} dr_1 h(X - r_1 | r_1) h(r_1 | \xi). \quad (5.9)$$

Przejdźmy teraz do zmiennych Fouriera-Laplace'a (wciąż dla dowolnego rozkładu warunkowego $h(r_n | r_{n-1})$)

$$\tilde{Q}_1(k; s | \xi) = \tilde{\psi}_1(s) \int_{-\infty}^{\infty} dX e^{ikX} h(X | \xi), \quad (5.10)$$

$$\tilde{Q}_2(k; s | \xi) = \tilde{\psi}_1(s)\tilde{\psi}(s) \int_{-\infty}^{\infty} dX e^{ikX} \int_{-\infty}^{\infty} dr_1 h(X - r_1 | r_1). \quad (5.11)$$

Dalej, podstawiając zależność daną równaniem (5.1) otrzymujemy

$$\frac{\tilde{Q}_1(k; s | \xi)}{\tilde{\psi}_1(s)} = \int_{-\infty}^{\infty} dX e^{ikX} [(1 - \epsilon) h(X) + \epsilon \delta(X + \xi)] = \left[(1 - \epsilon) \tilde{h}(k) + \epsilon e^{ik(-\xi)} \right] \quad (5.12)$$

oraz

$$\begin{aligned} \frac{\tilde{Q}_2(k; s | \xi)}{\tilde{\psi}_1(s) \tilde{\psi}(s)} &= \\ &= \int_{-\infty}^{\infty} dX e^{ikX} \int_{-\infty}^{\infty} dr_1 [(1 - \epsilon) h(r_1) + \epsilon \delta(r_1 + \xi)] [(1 - \epsilon) h(X - r_1) + \epsilon \delta(X)] \\ &= (1 - \epsilon)^2 \tilde{h}^2(k) + \epsilon (1 - \epsilon) + \epsilon (1 - \epsilon) \tilde{h}(k) e^{ik(-\xi)} + \epsilon^2. \end{aligned} \quad (5.13)$$

Wstawiając powyższe równania do (5.7) otrzymujemy już jawną postać ostrego propagatora, pod warunkiem skoku ξ poprzedzającego chwilę $t = 0$, w postaci

$$\tilde{Q}(k; s | \xi) = \tilde{\psi}_1(s) \frac{(1 - \epsilon) \tilde{h}(k) + \epsilon \left(e^{ik(-\xi)} + \tilde{\psi}(s) \right)}{1 - (1 - \epsilon) \tilde{h}(k) \tilde{\psi}(s) - \epsilon \tilde{\psi}^2(s)}. \quad (5.14)$$

Dokonujemy teraz uśrednienia powyższego równania po ξ

$$\begin{aligned} \tilde{Q}(k; s) &= \int_{-\infty}^{\infty} d\xi \tilde{Q}(k, s | \xi) h(\xi) = \tilde{\psi}_1(s) \frac{(1 - \epsilon) \tilde{h}(k) + \epsilon \left(\tilde{h}(k) + \tilde{\psi}(s) \right)}{1 - (1 - \epsilon) \tilde{h}(k) \tilde{\psi}(s) - \epsilon \tilde{\psi}^2(s)} = \\ &= \tilde{\psi}_1(s) \frac{\tilde{h}(k) + \epsilon \tilde{\psi}(s)}{1 - (1 - \epsilon) \tilde{h}(k) \tilde{\psi}(s) - \epsilon \tilde{\psi}^2(s)}. \end{aligned} \quad (5.15)$$

Jak widać, konkretna postać ostrego propagatora zależy od dwóch rozkładów czasów wyczekiwania $\psi(t)$ i $\psi_1(t)$, jednego rozkładu pojedynczego skoku $h(x)$ oraz parametru ϵ . Zauważmy, że podstawiając w równaniu (5.15) $\epsilon = 0$ otrzymujemy, tak jak to być powinno, równanie (3.26), czyli wynik dla przypadku niezależnych zmian ceny. **Należy podkreślić, że wzór (5.15) stanowi jeden z dwóch filarów niniejszej pracy.** Dzięki niemu możliwe będzie wyprowadzenie interesujących nas charakterystyk drugorzędowych.

5.5 Wariancja i funkcja autokorelacji prędkości

Aby znaleźć miękki propagator, posiadając propagator ostry dany równaniem (5.15), skorzystam z zależności (3.22). W dalszym ciągu interesować nas będzie jedynie przypadek stacjonarny. Zatem, możemy wykorzystać jawną postać $\tilde{\psi}_1(s)$ daną równaniem (3.18). Wówczas uzyskuje się

$$\tilde{P}(k, s) = \frac{1}{s} - \frac{1 - \tilde{\psi}(s)}{\langle t \rangle s^2} + \frac{1}{\langle t \rangle s^2} \frac{\left(1 - \tilde{\psi}(s) \right)^2 \left(\tilde{h}(k) + \epsilon \tilde{\psi}(s) \right)}{1 - (1 - \epsilon) \tilde{h}(k) \tilde{\psi}(s) - \epsilon \tilde{\psi}^2(s)}, \quad (5.16)$$

gdzie dwa pierwsze składniki sumy stanowią element $\tilde{\Psi}_1(s)$ obecny w równaniu (3.22) a uzyskany ze wzorów (3.18) i (3.19). Równanie (5.16) nie zależy już jawnie od rozkładu $\tilde{\psi}_1(s)$, gdyż został on wyrażony przez rozkład $\tilde{\psi}(s)$. Stąd, oraz na podstawie wzoru (3.30), otrzymujemy średnie przesunięcie kwadratowe

$$\tilde{m}_2(s) = - \left. \frac{\partial^2 \tilde{P}(k, s)}{\partial k^2} \right|_{k=0} = \frac{\mu_2}{s^2 \langle t \rangle} \frac{1 - \epsilon \tilde{\psi}(s)}{1 + \epsilon \tilde{\psi}(s)}, \quad (5.17)$$

a z (3.35) funkcję autokorelacji prędkości

$$\tilde{C}(s) = \frac{\mu_2}{2 \langle t \rangle} \frac{1 - \epsilon \tilde{\psi}(s)}{1 + \epsilon \tilde{\psi}(s)} \quad (5.18)$$

w zmiennej Lapalce'a. Dokonajmy teraz odwrotnej transformaty Laplace'a wyrażenia (5.18)

$$C(t) = \frac{\mu_2}{2 \langle t \rangle} \delta(t) - \epsilon \frac{\mu_2}{\langle t \rangle} \mathcal{L}_t^{-1} \left\{ \frac{\tilde{\psi}(s)}{1 + \epsilon \tilde{\psi}(s)} \right\}, \quad (5.19)$$

gdzie $\mathcal{L}_t^{-1}\{\dots\}$ jest odwrotną transformatą Laplace'a do zmiennej czasowej. Podobnie jak w poprzednim rozdziale, dokonując normalizacji funkcji autokorelacji prędkości do delty Diraca w zerze otrzymujemy

$$C^n(t) = \delta(t) - 2\epsilon \mathcal{L}_t^{-1} \left\{ \frac{\tilde{\psi}(s)}{1 + \epsilon \tilde{\psi}(s)} \right\}. \quad (5.20)$$

Ważną własnością równania (5.20) jest jego niezależność od rozkładu pojedynczego skoku (nie występują ani moment μ_1 ani μ_2) w przeciwieństwie do analogicznego równania (4.32) dla modelu Montero i Masolivera. W tym drugim przypadku autokorelacja (poza deltą Diraca w zerze) jest skalowana przez czynnik μ_1^2/μ_2 . Jak wspomniałem wcześniej, dla rozkładu pojedynczego skoku danego przez (4.1) oba modele są sobie równoważne, a wówczas czynnik skalujący $\mu_1^2/\mu_2 = 1$, jak należało oczekiwać.

W uzyskanym tutaj modelu kształt unormowanej funkcji autokorelacji prędkości (5.20) zależy jedynie od rozkładu czasów oczekiwania $\psi(t)$ oraz parametru ϵ . Należy podkreślić, że zależność od parametru ϵ nie jest jedynie prostym skalowaniem, gdyż występuje on również w mianowniku równania (5.20) przez co, jego wpływ na wyniki jest wyraźnie nieliniowy.

5.6 Porównanie z danymi empirycznymi

Zasadniczym celem niniejszego podrozdziału jest porównanie przewidywań zależności czasowej teoretycznej (unormowanej) funkcji autokorelacji prędkości z wynikami empirycznymi przedstawionymi na rysunku 2.2 (małe czarne kwadraty). Porównanie

takie przeprowadzam zarówno dla modelu Montero i Masolivera, w którym funkcja autokorelacji prędkości opisana jest równaniem (4.32), jak i dla modelu opisanego w niniejszym rozdziale, gdzie unormowana funkcja autokorelacji prędkości dana jest przez wyrażenie (5.20).

Jak podkreśliłem w poprzednim podrozdziale, kształt funkcji autokorelacji prędkości wyprowadzonej w ramach rozwiniętego przeze mnie modelu zależy jedynie od rozkładu czasów wyczekiwania $\psi(t)$ oraz parametru ϵ . W stacjonaryzowanym przeze mnie modelu Montero i Masolivera istotny jest również stosunek momentów μ_1^2/μ_2 . Są to jedyne parametry, które należy estymować na bazie posiadanych danych empirycznych. Zauważmy, że do uzyskania każdego z tych parametrów, nie potrzebujemy całości danych, składających się ze zmian ceny i czasów pomiędzy tymi zmianami. Do uzyskania parametru ϵ potrzebny jest jedynie szereg kolejnych zmian ceny, natomiast do uzyskania rozkładu czasów oczekiwania - szereg kolejnych czasów wyczekiwania. Ponadto, w drugim przypadku interesujemy się jedynie histogramem czasów wyczekiwania zatem, ich kolejność w szeregu czasowym nie ma znaczenia. Powyższe obserwacje są istotne, gdyż żaden z parametrów obecnych w równaniu (5.20) nie zawiera w sobie struktury czasowo-cenowej, która w zupełności określona jest założeniami stosowanego modelu. W pierwszej kolejności opisuję metodę estymacji parametru ϵ .

5.6.1 Estymacja parametru ϵ

Zauważmy, że przyjmując zależność pomiędzy dwoma kolejnymi skokami daną przez wyrażenie (5.1), otrzymuję korelację dwóch kolejnych skoków

$$\begin{aligned}
\langle r_n r_{n-1} \rangle &= \int_{-\infty}^{\infty} dr_n \int_{-\infty}^{\infty} dr_{n-1} r_n r_{n-1} h(r_n, r_{n-1}) = \\
&= \int_{-\infty}^{\infty} dr_n \int_{-\infty}^{\infty} dr_{n-1} r_n r_{n-1} [(1 - \epsilon)h(r_n)h(r_{n-1}) + \epsilon h(r_n)\delta(r_n + r_{n-1})] = \\
&= \epsilon \int_{-\infty}^{\infty} dr_n \int_{-\infty}^{\infty} dr_{n-1} r_n r_{n-1} h(r_n)\delta(r_n + r_{n-1}) = -\epsilon \int_{-\infty}^{\infty} dr_n r_n^2 h(r_n) = \\
&= -\epsilon \mu_2,
\end{aligned} \tag{5.21}$$

stąd

$$\epsilon = -\frac{\langle r_n r_{n-1} \rangle}{\mu_2}. \tag{5.22}$$

W powyższych wyrażeniach brak jest iloczynu średnich $\langle r_n \rangle \langle r_{n-1} \rangle$, gdyż z parzystości rozkładu skoków wynika znikanie ich średniej długości. Ta prosta estymacja

parametru ϵ przez korelację jest podobna do znanej momentowej metody estymacji rozkładów [65]. W analizowanej przeze mnie sytuacji metoda oparta na wzorze (5.22) ma jednak istotne wady, o czym mowa poniżej.

Jak wynika z obserwacji poczynionych w trakcie analizy rysunku 5.1 (w podrozdziale 5.1), rozkład prawdopodobieństwa pojedynczego skoku zawiera znaczącą liczbę skoków o długości zero (niemal 50%), odpowiadających transakcjom bez zmiany ceny. W empirycznym rozkładzie dwóch kolejnych skoków uwidacznia się to (na tym rysunku) w postaci centralnego krzyża. Są to sytuacje, gdy co najmniej jedna z dwóch kolejnych zmian cen jest równa zero. Zauważmy, że w takim przypadku nie dają one wkładu do estymatora korelacji. Wyznaczony estymator korelacji bazuje wyłącznie na przeskokach poza centralnym krzyżem (blisko 25% przypadków), co w badanej sytuacji nie daje wystarczająco dobrej statystyki.

W związku z powyższym, wykorzystuję metodę estymacji bazującą na macierzy 3x3 użytą zarówno przez Tsay'a [26] jak i przez Montero i Masolivera [42]. Zgodnie z tą metodą skoki dzielone są na ujemne (-1), zerowe (0) i dodatnie (1), a elementy macierzowe $h_{i,j}$, gdzie $i, j = -1, 0, 1$, reprezentują unormowany histogram dwóch kolejnych (tego typu) skoków. Zsymetryzowany histogram pojedynczego skoku oznaczam przez h_i , gdzie $i = -1, 0, 1$ ($h_{-1} = h_1$). Bazując na wyrażeniu (5.1), teoretyczny odpowiednik empirycznej macierzy otrzymuję w postaci

$$h_{i,j}^{teor}(\epsilon) = (1 - \epsilon) h_i h_j + \epsilon \delta_{i,-j} h_i, \quad (5.23)$$

gdzie δ oznacza teraz deltę Kroneckera. Maksymalizując funkcjonal

$$L[\epsilon] = \sum_{i=-1}^1 \sum_{j=-1}^1 h_{i,j} \log(h_{i,j}^{teor}), \quad (5.24)$$

ze względu na ϵ otrzymujemy estymator metodą największej wiarygodności [65]. Natomiast, minimalizując funkcjonal

$$S[\epsilon] = \sum_{i=-1}^1 \sum_{j=-1}^1 (h_{i,j}^{teor} - h_{i,j})^2 \quad (5.25)$$

otrzymujemy estymator metodą najmniejszych kwadratów [65]. Numeryczna optymalizacja, dla dwóch wspomnianych powyżej metod, została przeprowadzona metodą złotego podziału [66], a jej wyniki wraz z estymacją parametru ϵ metodą korelacyjną (dla porównania) przedstawiono w tabeli 5.1

Zauważmy (kolumna druga i trzecia tabeli 5.1), że dla większości badanych watorów otrzymana wartość parametru ϵ mieści się w przedziale 0.2 - 0.3. Natomiast różnica pomiędzy wynikami uzyskanymi dla każdej spółki z osobna metodą największej wiarygodności oraz najmniejszych kwadratów mieści się w akceptowalnych granicach 10%. Wartości parametru ϵ uzyskane powyższymi dwoma metodami są do siebie zbliżone, natomiast wartości parametru ϵ otrzymane za pomocą korelacji dwóch kolejnych skoków są wyraźnie większe. Warto zwrócić uwagę także na to, że

Tablica 5.1: Wyniki estymacji parametru ϵ dla badanych spółek z WIG20. W drugiej kolumnie przedstawiono wyniki uzyskane metodą największej wiarygodności (maksymalizacji funkcjonau (5.24)), w trzeciej kolumnie – wyniki uzyskane metodą najmniejszych kwadratów (minimalizacji funkcjonau (5.25)), natomiast w czwartej – metodą korelacyjną (opartą o równanie (5.22)).

Skrót nazwy spółki	ϵ : m.najw.wiaryg.	ϵ : m.najm.kwadratów	ϵ : korelacja
BRE	0.220	0.222	0.304
GETIN	0.312	0.317	0.283
KGHM	0.249	0.228	0.401
PEKAO	0.198	0.196	0.341
PGNIG	0.289	0.251	0.410
PKNORLEN	0.290	0.271	0.354
PKOBP	0.313	0.303	0.371
TPSA	0.381	0.356	0.460

rozzut wyników uzyskanych dla wszystkich spółek dwiema pierwszymi metodami jest mniejszy niż w przypadku trzeciej.

Można teraz postawić pytanie: do jakiego stopnia analizowany szereg czasowy zmian ceny jest stacjonarny z punktu widzenia estymacji badanego powyżej parametru? Kolejne, związane z tym pytanie dotyczy zależności estymowanej wartości parametru ϵ od czasu pomiędzy analizowanymi zmianami ceny. Na te pytania odpowiedziano w pracy licencjackiej [67]. Z przedstawionych w tej pracy badań wynika, iż po podzieleniu szeregów czasowych (dla każdej spółki) na poszczególne lata oraz przeprowadzeniu estymacji dla każdego roku oddzielnie, nie obserwuje się istotnych różnic pomiędzy otrzymanymi wynikami. W pracy tej dokonano również estymacji parametru ϵ w dwóch przypadkach, tzn. gdy odległość czasowa między zmianami ceny była większa i mniejsza od średniego czasu pomiędzy zmianami. Gdy czasy pomiędzy skokami były dłuższe wystąpiła nieznaczna tendencja prowadząca do zmniejszenia wartości parametru ϵ . Może to sugerować konieczność rozważenia w przyszłości bardziej wyrafinowanych modeli, w których skok ceny i czas między skokami są od siebie zależne.

5.6.2 Rozkład czasów wyczekiwania

Kolejnym elementem, który należy wyznaczyć z danych empirycznych jest rozkład czasów wyczekiwania. W takim przypadku mamy do czynienia z estymowaniem funkcji, której argumentem jest przedział czasu. Aby uzyskać analityczną, zamkniętą postać funkcji autokorelacji prędkości, należy modelować posiadany rozkład empiryczny przez odpowiednio dobraną funkcję analityczną.

Analiza struktury czasów międzytransakcyjnych była już szeroko badana [68], stanowiąc np. jeden z głównych celów badawczych rozprawy doktorskiej [53] (gdzie zamieszczono obszerną listę odnośników literaturowych na ten temat). Dotychczas udało się zaproponować kilka postaci funkcji, które zdaniem autorów dobrze opi-

sują dane odpowiadające rozkładowi czasów wyczekiwania. Szczególnie popularne są: funkcja rozciągniętego eksponensa, funkcja Write'a i funkcja Mittag-Lefflera [68]. Jednak w niniejszej pracy, wymagania stawiane postaci funkcyjnej są specyficzne:

- (1) rozkład prawdopodobieństwa czasów wyczekiwania powinien posiadać transformatę Laplace'a w postaci zamkniętej, a ponadto
- (2) funkcjonal wymierny gęstości rozkładu w przestrzeni Laplace'a (patrz wzór (5.20)) powinien umożliwiać dokonanie transformaty powrotnej do zmiennej czasowej w postaci analitycznej.

To drugie założenie sugeruje, iż sama transformata Laplace'a rozkładu powinna być funkcją wymierną parametru s . Poniżej przedstawiam tylko dwa szczególnie proste przypadki.

Eksponencjalny rozkład czasów wyczekiwania. Jest to najczęściej stosowany w fizyce rozkład czasów międzydziennej, charakterystyczny dla procesu Poissona, dany wzorem

$$\psi(t) = \frac{1}{\langle t \rangle} e^{-t/\langle t \rangle}, \quad (5.26)$$

gdzie $\langle t \rangle$ jest średnim czasem międzydziennej. W zmiennej Laplace'a rozkład ten przyjmuje postać

$$\tilde{\psi}(s) = \frac{1}{\langle t \rangle s + 1}. \quad (5.27)$$

Przypomnijmy, iż jest to jedyny przypadek rozkładu ciągłego, dla którego $\psi_1(t) = \psi(t)$, a kanoniczne błądzenie losowe w czasie ciągłym jest procesem Markowa. Ponieważ rozkład eksponencjalny scharakteryzowany jest jednym parametrem, zatem najprostszym sposobem jego estymacji jest metoda momentowa, czyli użycie estymatora średniej.

Podwójnie eksponencjalny rozkład czasów wyczekiwania, czyli w postaci ważonej sumy dwóch rozkładów eksponencjalnych. Bardziej realistyczne, a zarazem możliwe do analitycznego potraktowania jest modelowanie rozkładu czasów wyczekiwania poprzez rozkład będący ważoną sumą dwóch rozkładów danych wzorem (5.26). Rozkład taki ma postać

$$\psi(t) = \frac{w}{\tau_1} e^{-t/\tau_1} + \frac{1-w}{\tau_2} e^{-t/\tau_2}, \quad (5.28)$$

gdzie $0 \leq w \leq 1$ jest wagą, a τ_1 i τ_2 cząstkowymi czasami relaksacji. W zmiennej Laplace'a rozkład ten przyjmuje postać

$$\tilde{\psi}(s) = \frac{w}{1 + s\tau_1} + \frac{1-w}{1 + s\tau_2}. \quad (5.29)$$

Powyższy rozkład posiada trzy parametry, które należy wyznaczyć metodą dopasowania (fitowania) rozkładu do histogramu empirycznego. Aby uzgodnić to fitowanie z wynikiem otrzymanym dla pojedynczego rozkładu eksponencjalnego, a przede wszystkim zmniejszyć liczbę dopasowywanych parametrów z trzech (w, τ_1, τ_2) do dwóch (τ_1, τ_2) , korzystam z prostej zależności

$$\langle t \rangle = w\tau_1 + (1 - w)\tau_2. \quad (5.30)$$

Dzięki temu, znając już średni czas $\langle t \rangle$ (obliczony z danych empirycznych jako średnia arytmetyczna wszystkich czasów międzytransakcyjnych) możemy uzyskać np. wagę w jako funkcję τ_1 i τ_2 . Wybór dwóch z trzech dopasowywanych parametrów jest oczywiście dowolny, jednak dopasowywanie parametrów tego samego rodzaju i należących do tego samego przedziału (od zera do nieskończoności) ułatwia znalezienie najlepszego dopasowania, przez co wybranie pary τ_1, τ_2 wydaje się bardziej naturalne.

Do estymacji parametrów użyłem metody najmniejszych kwadratów wraz z metodą optymalizacji Nelder–Meada [69] (zwaną również metodą sympleksową), w implementacji podanej w książce [66]. Ze względu na braki w danych empirycznych dotyczące transakcji zawartych w ramach pojedynczej sekundy zegarowej musiałem uznać, iż wartość histogramu dla wartości czasu $t = 0$ nie jest miarodajna. Zauważyć można również, iż dla czasu $t = 1$ występuje systematyczne odchylenie od regularności widocznej dla kolejnych chwil czasu, które może być również spowodowane zaokrągleniem czasu transakcji do jednej sekundy w danych empirycznych. Ponieważ kształt rozkładu czasów wyczekiwania dla relatywnie dłuższych czasów nie jest kluczowy dla dynamiki w najkrótszej skali czasu, zakresem dopasowywania objęto przedział $t \in [2, 100]$. Uzyskane parametry dla obu proponowanych rozkładów przedstawia tabela 5.2. Natomiast histogramy wraz z estymowanymi rozkładami pokazano na rysunku 5.5.

Tablica 5.2: Wyniki estymacji parametrów rozkładów modelujących rozkład czasów wyczekiwania dla badanych spółek z WIG20. W drugiej kolumnie zamieszczono średni czas międzytransakcyjny używany w rozkładzie (5.26), w trzeciej i czwartej kolumnie znajdują się cząstkowe czasy relaksacji τ_1 i τ_2 rozkładu podwójnie eksponencjalnego (5.28), natomiast w kolumnie piątej waga w .

Skrót nazwy spółki	$\langle t \rangle$	τ_1	τ_2	w
BRE	109.871	11.711	179.438	0.415
GETIN	85.064	12.292	137.976	0.421
KGHM	30.312	6.819	72.182	0.641
PEKAO	55.888	9.221	102.304	0.499
PGNIG	49.078	8.829	75.524	0.397
PKNORLEN	42.596	9.022	76.431	0.502
PKOBP	26.266	6.720	50.681	0.555
TPSA	34.979	8.048	66.714	0.541

Warto zwrócić uwagę, że zbliżone do przedstawionych w tabeli 5.2 wartości średnie czasów międzytransakcyjnych otrzymano w pracy [42] dla kilku spółek notowanych na NYSE³ w latach 1995-1998.

Drugą obserwacją, którą możemy poczynić jest rozdzielenie skal czasowych, gdyż estymaty τ_1 i τ_2 różnią się o rząd wielkości. Pozwala to traktować je jako cząstkowe czasy relaksacji na dwóch różnych skalach.

Na rysunku 5.5 widoczne jest w skali pół-logarytmicznej, niewielkie systematyczne odstępstwo przewidywań teoretycznych, opisanych rozkładem podwójnie eksponencjalnym, od danych empirycznych. Mimo, iż na kształt funkcji autokorelacji prędkości (wyrażonej wzorem (5.20)) ma wpływ transformata Laplace'a rozkładu, zależna od wartości rozkładu dla wszystkich czasów międzytransakcyjnych, to można przypuszczać, iż rola wspomnianego odstępstwa w rozkładzie dla dłuższych czasów ma stosunkowo niewielki wpływ na różnice pomiędzy teoretyczną i empiryczną funkcją autokorelacji prędkości dla krótkich czasów.

W ten sposób zgromadziliśmy już wszystkie informacje niezbędne do porównania przewidywań teoretycznych rozważanych modeli z empiryczną funkcją autokorelacji prędkości ceny. Porównanie to stanowi jeden z najważniejszych elementów niniejszej pracy.

5.6.3 Funkcja autokorelacji prędkości

Wyprowadzamy teraz zamknięte formuły na unormowaną funkcję autokorelacji prędkości, przyjmując rozkłady czasów wyczekiwania w postaciach (5.26) i (5.28) zaproponowanych w poprzednim podrozdziale 5.6.2.

Wstawiając równanie (5.27) do (5.20), dokonujemy odwrotnej transformaty Laplace'a uzyskując

$$C^n(t) = \delta(t) - \frac{2\epsilon}{\langle t \rangle} e^{-(1+\epsilon)t/\langle t \rangle}. \quad (5.31)$$

Jak poprzednio, dla stacjonaryzowanego modelu Montero i Masolivera należy uzupełnić powyższe równanie o czynnik μ_1^2/μ_2 przed drugim składnikiem po prawej stronie równania.

Dla podwójnie eksponencjalnego rozkładu czasów wyczekiwania możemy podstawić (5.29) do (5.20), co prowadzi do

$$C^n(t) = \delta(t) - 2\epsilon (A_1 e^{-\nu_1 t} + A_2 e^{-\nu_2 t}), \quad (5.32)$$

³NYSE to akronim angielskiej nazwy New York Stock Exchange - giełdy na Wall Street w Nowym Jorku.

gdzie

$$\begin{aligned}
\omega_j &= 1/\tau_j, \quad j = 1, 2, \\
v &= w \omega_1 + (1 - w) \omega_2, \\
\nu_{1,2} &= \frac{1}{2} \left[\omega_1 + \omega_2 + \epsilon v \pm \sqrt{(\omega_1 + \omega_2 + \epsilon v)^2 - 4\omega_1\omega_2(1 + \epsilon)} \right], \\
A_j &= (-1)^j \omega_1 \omega_2 \frac{1 - \nu_j v}{\nu_1 - \nu_2}, \quad j = 1, 2.
\end{aligned} \tag{5.33}$$

Wartości współczynników $1/\nu_1, 1/\nu_2, A_1, A_2$ funkcji autokorelacji prędkości określonej równaniem (5.32) dla parametrów uzyskanych metodami opisanymi w podrozdziałach 5.6.1 i 5.6.2 wraz z wartością współczynnika μ_1^2/μ_2 , przedstawiono w tabeli 5.3.

Tablica 5.3: Wartości parametrów $1/\nu_1, 1/\nu_2, A_1, A_2$, teoretycznej funkcji autokorelacji prędkości opisanej równaniem (5.32) ośmiu badanych spółek wchodzących w skład indeksu WIG20. Parametry wyznaczone według wzorów (5.33), bazując na estymowanych parametrach, których wartości podano w tabelach 5.1 (metoda największej wiarygodności) i 5.2. W ostatniej kolumnie podano również wartość współczynnika μ_1^2/μ_2 występującego w stacjonaryzowanym modelu Montero i Masolivera (4.32).

Skrót nazwy spółki	$1/\nu_1$	$1/\nu_2$	A_1	A_2	μ_1^2/μ_2
BRE	10.725	160.631	0.03598	0.00270	0.201
GETIN	10.844	119.180	0.03526	0.00319	0.195
KGHM	5.876	67.090	0.09533	0.00359	0.238
PEKAO	8.384	93.886	0.05500	0.00398	0.215
PGNIG	7.905	65.447	0.04669	0.00622	0.310
PKNORLEN	7.860	68.024	0.05738	0.00477	0.239
PKOBP	5.711	45.434	0.08538	0.00604	0.225
TPSA	6.651	58.466	0.06963	0.00447	0.244

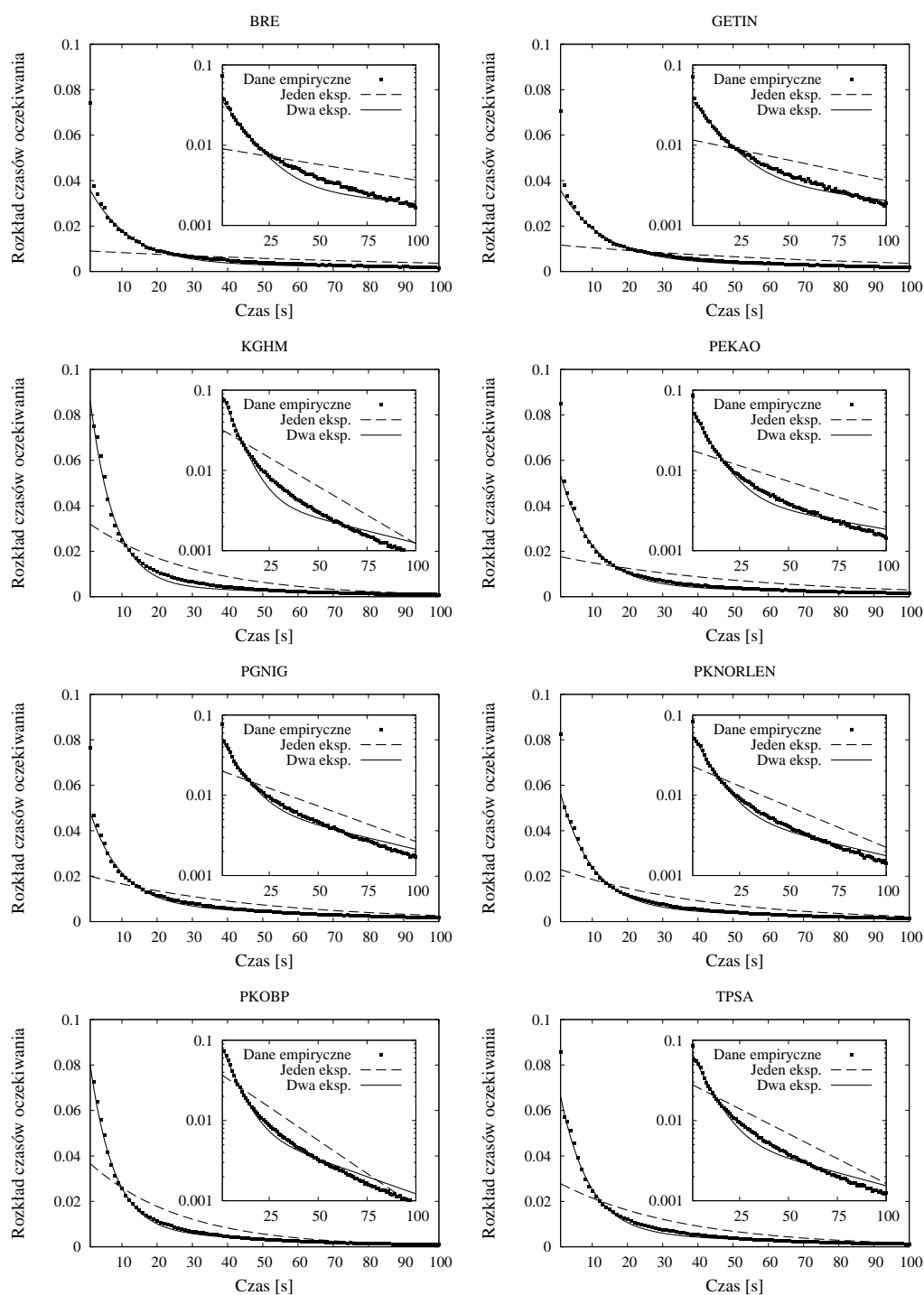
Na rysunku 5.6 porównałem przebieg empirycznej funkcji autokorelacji prędkości uzyskany metodą opisaną w rozdziale 2, a przedstawiony wcześniej na rysunku 2.2 (małe czarne kwadraty), z wynikami przewidywań wzorów (5.31) dla eksponencjalnego rozkładu czasów wyczekiwania (linia przerywana) i (5.32) dla podwójnie eksponencjalnego rozkładu czasów wyczekiwania (linia ciągła).

Pragnę w tym miejscu podkreślić szczególnie istotny aspekt mojego podejścia. Mianowicie krzywe teoretyczne na rysunku 5.6 nie są uzyskane metodą dopasowania ich parametrów do widocznych na tym samym rysunku danych empirycznych. Wszystkie parametry funkcji zostały bowiem wyznaczone już wcześniej niezależnymi metodami, na dwóch osobnych zbiorach danych, a mianowicie na szeregu kolejnych skoków oraz zbiorze czasów wyczekiwania.

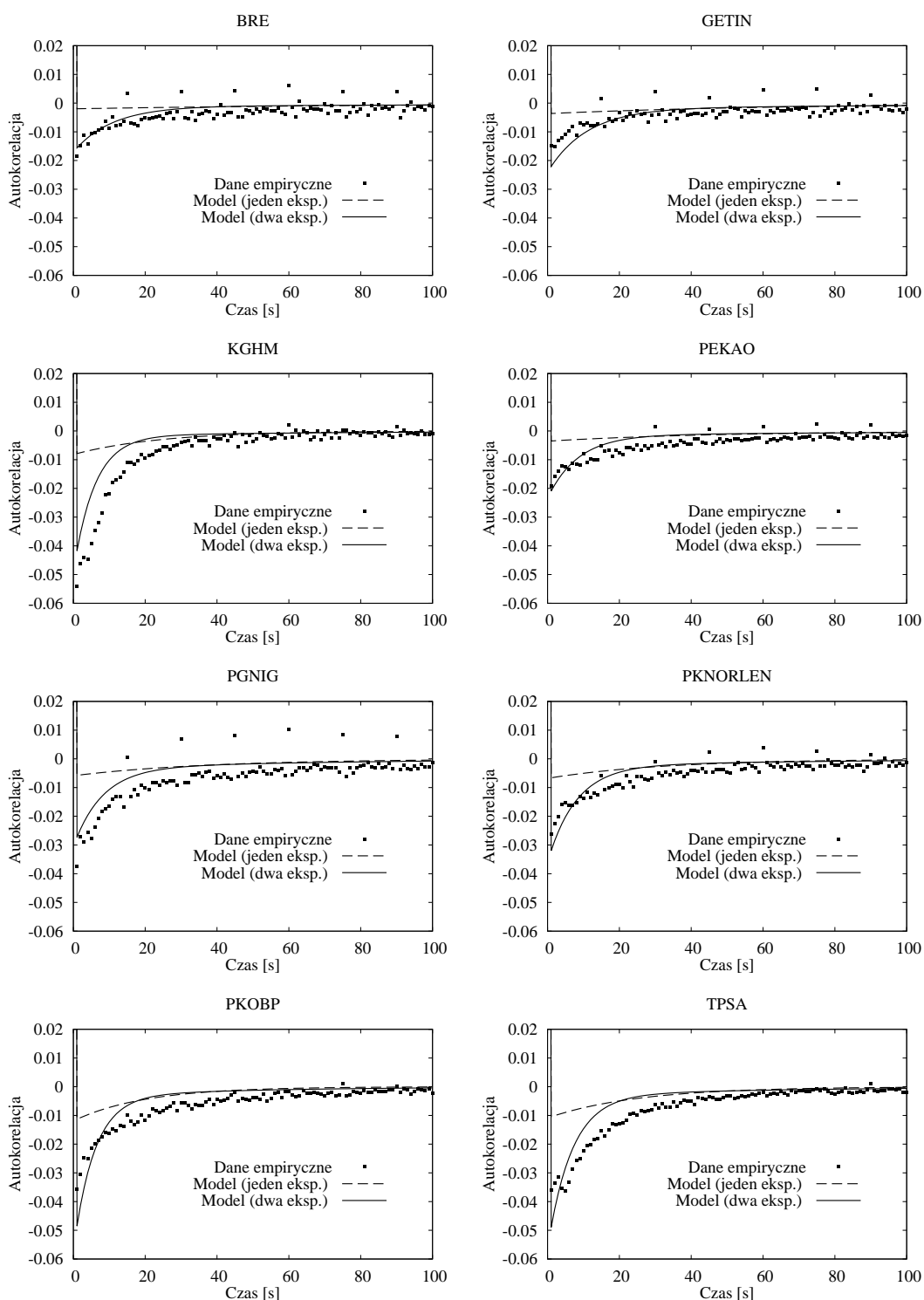
Pomimo, iż zgodność przewidywań teoretycznych wyrażonych formułą (5.32) z danymi empirycznymi jest satysfakcjonująca, wciąż widoczne jest systematyczne odstępstwo, w szczególności dla pośrednich czasów (patrz rysunek 5.6). Na rysunku

pominałem przewidywania stacjonaryzowanego modelu Montero i Masolivera, gdyż byłoby to jedynie przeskalowanie krzywej zaznaczonej linią ciągłą o parametr μ_1^2/μ_2 . Wartości tego parametru, jak pokazano w tabeli 5.3, są rzędu 0.2. Zatem, zgodność przewidywań tego modelu z danymi empirycznymi byłaby znacząco gorsza. Wynika stąd, iż zaproponowany w niniejszym rozdziale model znacznie lepiej opisuje dane empiryczne niż stacjonaryzowany model Montero i Masolivera.

Przedstawiona na rysunku 5.6 linią przerywaną funkcja autokorelacji prędkości dla eksponencjalnego rozkładu czasów wyczekiwania (dana wzorem (5.31)) wykazuje znaczne odstępstwo od danych empirycznych, czego można było się spodziewać analizując rysunek 5.5. Widoczna niezgodność na poziomie dopasowania do rozkładu czasów wyczekiwania przekłada się na niezgodność funkcji autokorelacji prędkości widoczną na rysunku 5.6. Jednak funkcja autokorelacji prędkości dana równaniem (5.20) opisuje z grubsza charakterystyczne zachowanie, tzn. znak i rząd wielkości ujemnej części funkcji autokorelacji. Zatem, zgodność przewidywań równania (5.32) z danymi empirycznymi jest bezpośrednią konsekwencją ulepszenia rozkładu czasów wyczekiwania. Możemy przypuszczać, że dla rozkładów czasów wyczekiwania, które lepiej opisywałyby rozkład empiryczny, zgodności przebiegu funkcji autokorelacji prędkości z danymi empirycznymi uległaby poprawie.



Rysunek 5.5: Porównanie empirycznego histogramu czasów wyczekiwania (małe czarne kwadraty) z dopasowanymi przewidywaniami modeli opisującymi odpowiadający histogramowi rozkład. Linia przerywaną przedstawiony jest rozkład eksponencjalny wyrażony wzorem (5.26), natomiast linią ciągłą rozkład podwójnie eksponencjalny (5.28). Wewnątrz wykresów wykonanych w skali liniowej znajdują się ich zmniejszone odpowiedniki w skali pół-logarytmicznej. Dopiero na zmniejszonych wykresach widać niewielkie, chociaż systematyczne odstępstwo rozkładu podwójnie eksponencjalnego od rozkładu empirycznego. Niezgodność ta ma jednak niewielki wpływ na uzyskane wyniki końcowe.



Rysunek 5.6: Unormowana funkcja autokorelacji prędkości ceny: empiryczna (czarne kwadraty), wyrażona wzorem (5.31) dla eksponencjalnego rozkładu czasów wyczekiwania (linia przerywana) i wyrażona wzorem (5.32) dla podwójnie eksponencjalnego rozkładu czasów wyczekiwania (linia ciągła). Parametry krzywych teoretycznych zostały otrzymane osobnymi metodami estymacji. Zatem, przedstawione krzywe uzyskano bez potrzeby dopasowywania (fitowania) jakiegokolwiek parametru do danych empirycznych.

Rozdział 6

Model z pamięcią dwukrokową

Kolejnym, bardziej realistycznym, autorskim modelem procesu stochastycznego rozważanego w niniejszej rozprawie, jest model zawierający pamięć obejmującą dwa skoki poprzedzające aktualny. Tego typu modele nie były dotychczas rozważane w ramach formalizmu błędzenia losowego w czasie ciągłym, ze względu na znaczne komplikacje rachunkowe. **Innymi słowy, model z pamięcią dwukrokową stanowi całkowicie oryginalną część teoretyczną niniejszej rozprawy.**

6.1 Motywacja modelu

Zasadniczą przyczyną rozwinięcia kolejnego modelu jest konieczność poprawienia zgodności przewidywań teoretycznych, uzyskanych w ramach modelu z pamięcią jednokrokową, z danymi empirycznymi (patrz rysunek 5.6). Bezpośrednią motywacją do ulepszenia modelu była analiza autokorelacji empirycznego szeregu samych zmian cen.

Jeżeli badany szereg danych empirycznych zubożymy usuwając z niego czasy wyczekiwania, otrzymamy dyskretny w czasie szereg czasowy, gdzie rolę (dyskretnego) czasu pełni numer kolejnego skoku. Możemy analizować autokorelację takiego szeregu, używając jako argumentu liczby naturalnej lub zera. Na przykład, dla argumentu 0 badana autokorelacja odpowiada wariancji rozkładu skoków, dla argumentu 1 - kowariancji dwóch następujących po sobie skoków, dla argumentu 2 - kowariancji co drugiego skoku, itd. Innymi słowy, liczba ta jest liczbą kroków rozdzielających dwa ustalone skoki - wyjściowy i aktualny. Przebieg tak opisanej funkcji, po unormowaniu do wartości 1 w zerze (czyli w nazewnictwie ekonometrycznym autokorelacji), przedstawia rysunek 6.1.

Dla argumentu 1 wartości wykreślonej funkcji były już analizowane w ramach równania (5.22). Odpowiada ona parametrowi ϵ , co przedstawiono w ostatniej kolumnie tabeli 5.1. Co szczególnie ważne, dla kolejnych wartości liczby rozdzielających kroków nie obserwuje się już istotnych korelacji. Obserwacja ta była już poczyniona we wcześniejszych pracach, m.in w [42]. Stanowi to zasadniczą różnicę w stosunku do autokorelacji prędkości w funkcji czasu, pokazanej na rysunku 5.6, gdzie znacząca

zależność dwóch kolejnych zmian cen ulega rozmyciu z powodu zmiennego czasu pomiędzy kolejnymi skokami.

W świetle powyższego, należy zadać kluczowe pytanie, **czy z obserwowanego braku korelacji wynika brak jakiejkolwiek zależności pomiędzy co drugą zmianą ceny?**

6.1.1 Analiza co drugiej zmiany ceny

Aby odpowiedzieć na postawione w poprzednim podrozdziale pytanie konstruując, na podstawie analizowanych w niniejszej pracy danych empirycznych (analogicznie do wykresów pokazanych na rysunku 5.1), dwuwymiarowy histogram co drugiej zmiany ceny. Stanowi on statystyczną realizację rozkładu prawdopodobieństwa dla wspomnianej zależności co drugiej zmiany ceny (patrz rysunek 6.2).

Wyniki przedstawione na rysunku 6.2 są charakterystycznym przykładem dwóch zmiennych losowych, które są zależne chociaż ich korelacja znika. Mówiąc ogólniej, ze znikania korelacji nie wynika bezwarunkowość rozkładów.

Przedstawione histogramy, podobnie jak w przypadku tych na rysunku 5.1, składają się z dwóch głównych części:

- z centralnego krzyża, zdefiniowanego przez punkty $(r_{n-2}, 0)$ i $(0, r_n)$, czyli sytuacji, gdy co najmniej jedna z dwóch analizowanych zmian ceny była zerowa. Fakt ten nie implikuje istnienia zależności pomiędzy tymi skokami ceny.
- Z diagonalami i antydiagonalami, zdefiniowanymi poprzez punkty $(r_{n-2}, r_n \approx r_{n-2})$ i $(r_{n-2}, r_n \approx -r_n)$, tworzących krzyż obrócony względem poprzedniego o 45° . Wynik ten odpowiada sytuacji, gdy analizowane dwie zmiany cen mają taką samą długość, przy czym mogą mieć taki sam lub przeciwny zwrot, w przybliżeniu z równym prawdopodobieństwem.

Rozważmy najpierw, jaką zależność co drugiej zmiany ceny uzyskujemy w ramach modelu opisywanego w poprzednim rozdziale (czyli modelu z pamięcią jednokrokową), opierającego się na założeniach leżących u podstaw wyrażenia (1.2).

Skonstruujmy rozkład trzech następujących po sobie zmian ceny. Przy założeniu zależności dwóch kolejnych skoków danej wyrażeniem (5.1), otrzymuję

$$\begin{aligned}
 h(r_n, r_{n-1}, r_{n-2}) &= h(r_n | r_{n-1}, r_{n-2}) h(r_{n-1}, r_{n-2}) = \\
 &= h(r_n | r_{n-1}) h(r_{n-1}, r_{n-2}) = h(r_n | r_{n-1}) h(r_{n-1} | r_{n-2}) h(r_{n-2}) = \\
 &= [(1 - \epsilon)h(r_n) + \epsilon \delta(r_n + r_{n-1})] [(1 - \epsilon)h(r_{n-1}) + \epsilon \delta(r_{n-1} + r_{n-2})] h(r_{n-2}) = \\
 &= [(1 - \epsilon)^2 h(r_n) h(r_{n-1}) + (1 - \epsilon) \epsilon h(r_n) \delta(r_{n-1} + r_{n-2}) + \\
 &+ (1 - \epsilon) \epsilon \delta(r_n + r_{n-1}) h(r_{n-1}) + \epsilon^2 \delta(r_n + r_{n-1}) \delta(r_{n-1} + r_{n-2})] h(r_{n-2}).
 \end{aligned} \tag{6.1}$$

Zauważmy, że istnienie pamięci jednokrokowej zostało wykorzystane w drugim wierszu, gdzie rozkład warunkowy $h(r_n | r_{n-1}, r_{n-2})$ został zastąpiony przez uproszczony rozkład warunkowy $h(r_n | r_{n-1})$.

Dysponując powyższym rozkładem trójelementowym, możemy już zdefiniować gęstość rozkładu co drugiego skoku¹ $h_2(r_n, r_{n-2})$

$$h_2(r_n, r_{n-2}) = \int_{-\infty}^{\infty} dr_{n-1} h(r_n, r_{n-1}, r_{n-2}), \quad (6.2)$$

Dla rozkładu $h(r_n, r_{n-1}, r_{n-2})$ danego wzorem (6.1) wyrażenie (6.2) sprowadza się do postaci

$$\begin{aligned} h_2(r_n, r_{n-2}) &= [(1 - \epsilon)^2 h(r_n) + (1 - \epsilon) \epsilon h(r_n) + \\ &\quad + (1 - \epsilon) \epsilon h(-r_n) + \epsilon^2 \delta(r_n - r_{n-2})] h(r_{n-2}) = \\ &= (1 - \epsilon^2) h(r_n) h(r_{n-2}) + \epsilon^2 \delta(r_n - r_{n-2}) h(r_{n-2}). \end{aligned} \quad (6.3)$$

Powyższą zależność przestawiam w postaci wykresu (patrz rysunek 6.3), analogicznie do wykresu na rysunku 5.3. Można zauważyć, że **zależność (6.3) odtwarza tylko jedną z dwóch obecnych na rysunku 6.2 przekątnych** (diagonali i antydiagonali). Ponadto, z równania (6.3) można łatwo dowieść (analogicznie do wyprowadzenia wzoru (5.22)), że w ramach modelu z pamięcią jednokrokową korelacja co drugiej zmiany ceny przybiera postać

$$\frac{\langle r_n r_{n-2} \rangle}{\mu_2} = \epsilon^2, \quad (6.4)$$

a więc nie znika, w przeciwieństwie do tego co wynika z danych empirycznych. Zatem, **przewidywania modelu z pamięcią jednokrokową są niewystarczające do opisu danych empirycznych.**

Fenomenologiczna analiza rysunku 6.2, analogiczna do przedstawionej w rozdziale 5.1, dopuszcza następujące, formalne uogólnienie wyrażenia (6.3), odtwarzające (jak zobaczymy) istnienie diagonali i antydiagonali

$$\begin{aligned} h_2(r_n, r_{n-2}) &= (1 - 2\zeta) h(r_n) h(r_{n-2}) + \zeta \delta(r_n - r_{n-2}) h(r_{n-2}) + \\ &\quad + \zeta \delta(r_n + r_{n-2}) h(r_{n-2}). \end{aligned} \quad (6.5)$$

W celu uzyskania formalnej zgodności parametru ϵ^2 z równania (6.3) z parametrem ζ z równania (6.5) wystarczy po prostu położyć $\zeta = \epsilon^2$. Dlatego, na razie pozostajemy przy ogólniejszym sformułowaniu, czyli wyrażeniu (6.5).

Sporządzamy teraz wykres analizowanej zależności będący odpowiednikiem wykresów na rysunkach 5.3 i 6.3. Postulowana przez wyrażenie (6.5) postać zależności co drugiej zmiany ceny przedstawiona jest na rysunku 6.4.

Widać, że równanie (6.5) odtwarza zarówno digonałę przewidywaną już wcześniej równaniem (6.3) jak również antydiagonałą, której brakowało na rysunku 6.3. Fakt

¹Dotychczas literą h oznaczałem gęstości i warunkowe gęstości prawdopodobieństwa, w których argumenty były kolejnymi, występującymi bezpośrednio po sobie zmianami ceny (w kolejności odwrotnej do chronologicznej, jeżeli było to istotne). Tutaj dwa argumenty funkcji nie są kolejnymi zmianami ceny, więc aby uniknąć konfliktu oznaczeń wprowadzam nową funkcję h_2 .

ten jest spowodowany pojawieniem się w wyrażeniu (6.5) składnika zawierającego $\delta(r_n + r_{n-2})$. Składnik ten odpowiada sytuacji, gdy aktualny skok r_n , ma zwrot przeciwny do przedostatniego r_{n-2} , ale identyczną długość.

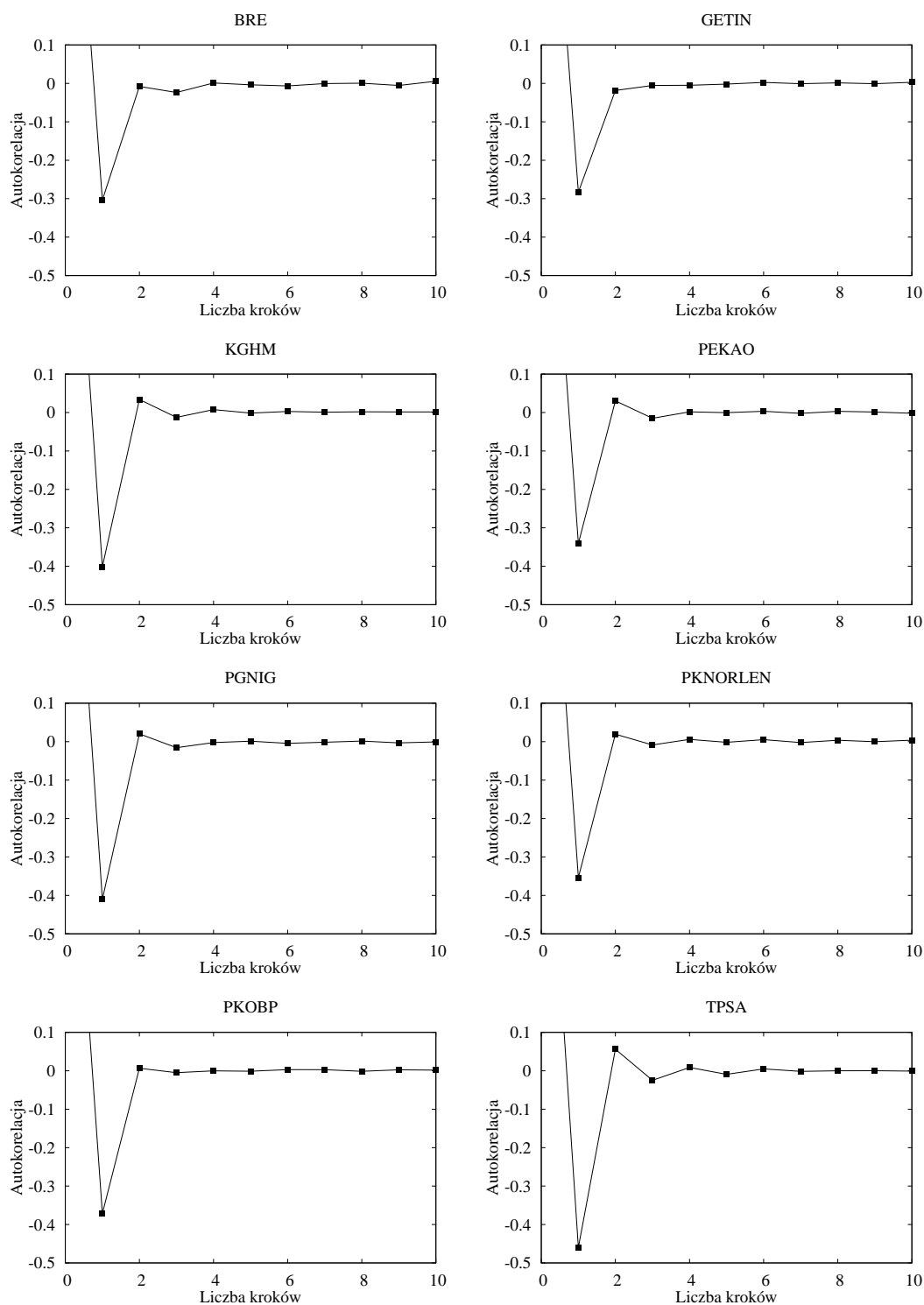
6.1.2 Zjawisko bid-ask bounce z opóźnieniem

Teoretyczne uzasadnienie wyrażenia (6.5) wynika z bardziej szczegółowej analizy mechanizmu bid-ask bounce w stosunku do przeprowadzonej w rozdziale 5.2. Powracając do wspomnianych rozważań, rozwijamy zamieszczony tam podpunkt (a).

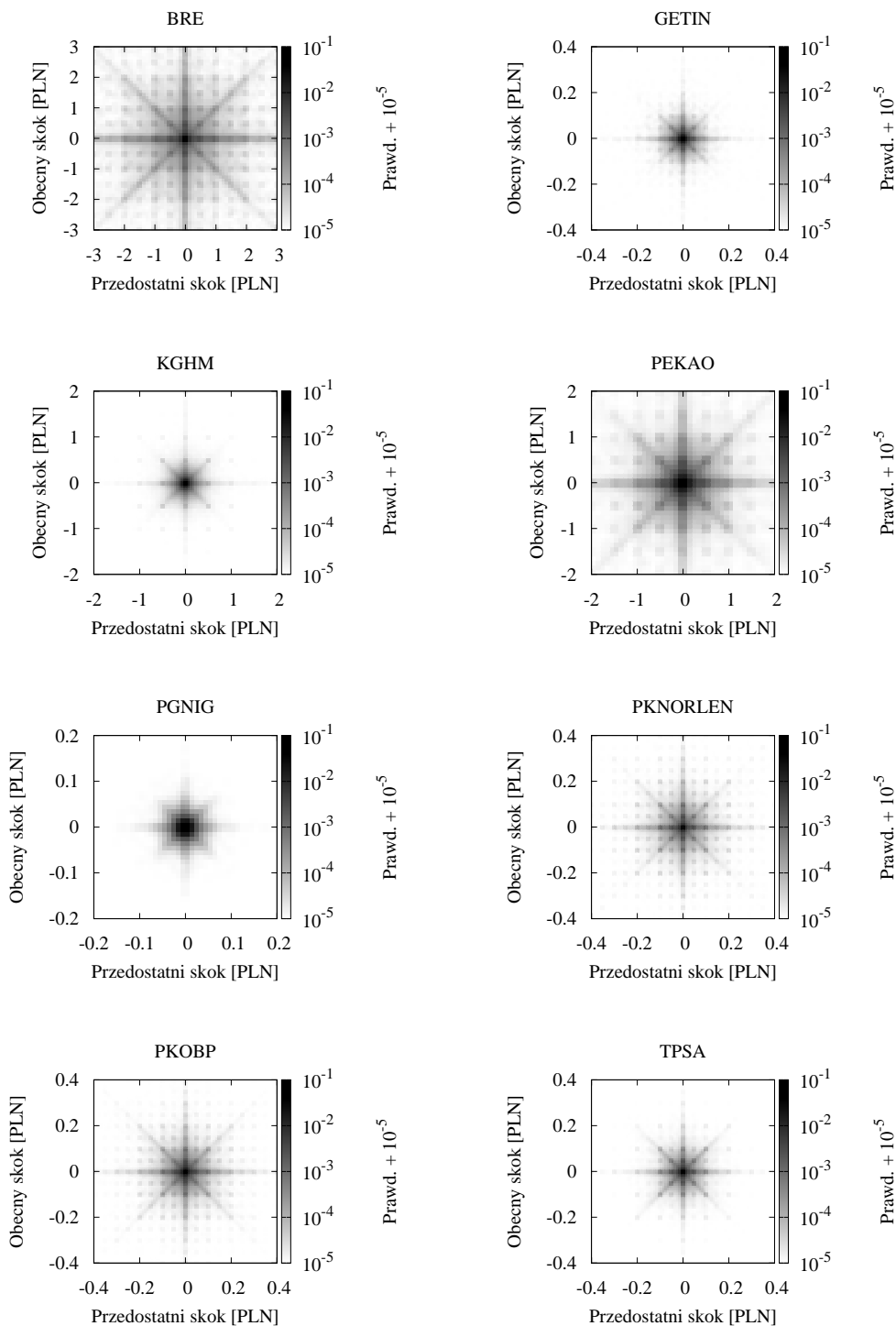
Przewidywaliśmy tam, że po pojawieniu się wykonanej oferty kupna, kolejny (drugi) skok ceny, wywołany tego typu ofertą, będzie dodatni i niezależny od poprzedniego. Nie interesowało nas jednak czy następujący po nim (trzeci) skok ceny zależy od któregośkolwiek z poprzedzających. Zauważmy, że w tej sytuacji lewa krawędź bid-ask spreadu pozostała niezmienną. Zatem, układ wciąż przechowuje pamięć o przedostatniej zmianie ceny. Analogicznie do rozważań przedstawionych już w podrozdziale 5.2, po wykonaniu pojedynczego, całkowicie niezależnego skoku, z nieznikającym prawdopodobieństwem, cena powróci do lewej krawędzi bid-ask spreadu, pokazanego na rysunku 5.2c.

Zauważmy, że wspomniany niezależny (drugi) skok będzie nieujemny. W dodatku, aby zależność co drugiej zmiany ceny można było przybliżyć przez $\delta(r_n + r_{n-2})$, skok ten powinien być niewielkiej długości. Jednak, aby zależności skoków w parach (r_n, r_{n-1}) i (r_{n-1}, r_{n-2}) były identyczne, musimy przyjąć, iż środkowy, niezależny skok r_{n-1} , następujący pomiędzy dwoma skorelowanymi ujemnie, jest losowany z rozkładu $h(r_{n-1})$.

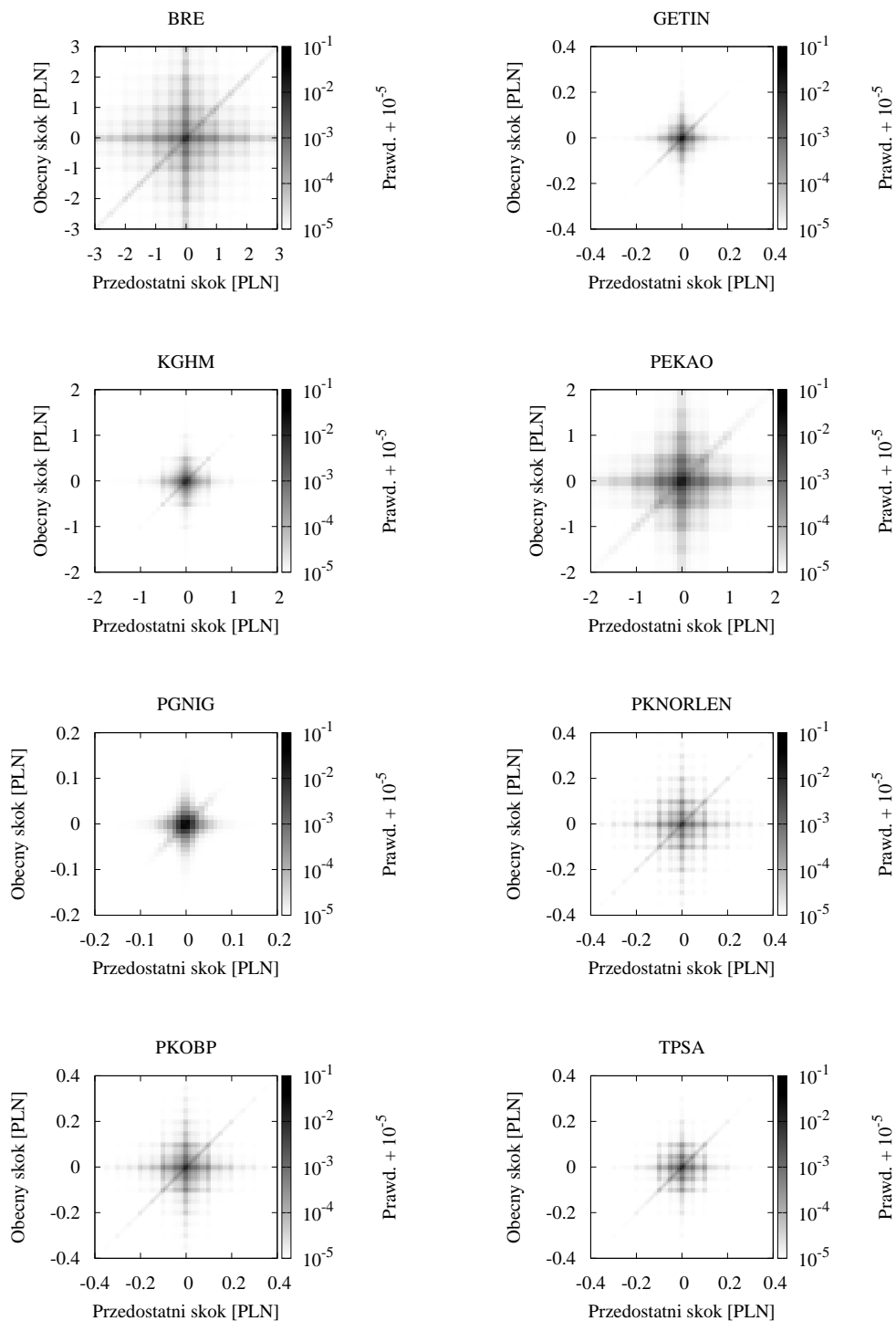
Rozważania przedstawione w niniejszym podrozdziale 6.1 stanowią zasadniczą motywację do skonstruowania przeze mnie oryginalnego modelu, w którym kolejny skok ceny jest zależny od dwóch poprzedzających go zmian ceny. Model taki konstruuję poprzez podanie gęstości prawdopodobieństwa trzech kolejnych skoków cen, przy założeniu niezależności czasów wyczekiwania. Postać tej gęstości przedstawiam w następnym podrozdziale.



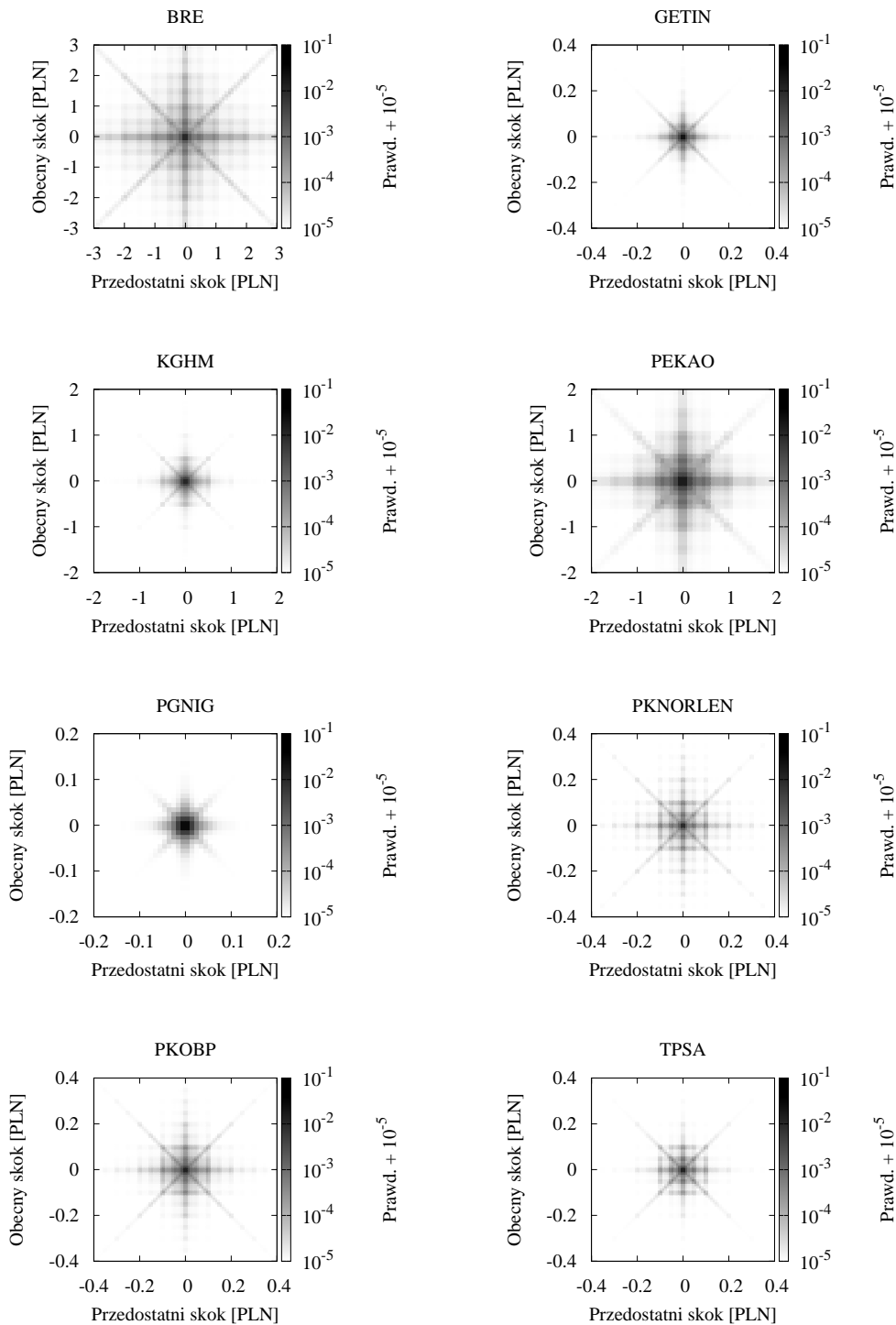
Rysunek 6.1: Unormowana funkcja autokorelacji szeregu kolejnych zmian ceny w funkcji przesunięcia (liczby rozdzielających je kroków). Wartość dla danego przesunięcia zaznaczona jest małym czarnym kwadratem, linie łączące wprowadziłem dla większej czytelności (nie reprezentują wartości pośrednich). Wartość funkcji autokorelacji dla przesunięcia 0 wynosi 1, wychodząc poza wykres. Obserwujemy znaczącą ujemną korelację dwóch kolejnych zmian cen i brak istotnej korelacji, gdy przesunięcie jest większe od 1.



Rysunek 6.2: Histogram empiryczny obecnej oraz przedostatniej zmiany ceny. Przyjęte na rysunku oznaczenia są identyczne jak te na rysunku 5.1. Obecna i przedostatnia zmiana ceny przedzielone są ostatnią, niezobrazowaną na powyższych wykresach. Dobrze widoczny jest nie tylko centralny krzyż, ale również diagonalna i antydiagonalna. Ta obecność diagonalni stanowi zasadniczą różnicę w stosunku do histogramów przedstawionych na rysunku 5.1.



Rysunek 6.3: Histogram teoretyczny zmiany ceny i przedostatniej zmiany ceny powstały w ramach modelu z pamięcią jednokrokową. Przyjęte na rysunku oznaczenia są identyczne jak te na rysunku 5.1. Jak widać, histogram odtwarza tylko jedną (diagonalę) z dwóch widocznych na rysunku 6.2 przekątnych (diagonali i antydiagonali). Wynika stąd, że model z pamięcią jednokrokową jest niewystarczający do opisu zależności co drugiej zmiany ceny.



Rysunek 6.4: Histogram teoretyczny zmiany ceny i przedostatniej zmiany ceny powstały w oparciu o równanie (6.5) i postulat, że $\zeta = \epsilon^2$. Przyjęte na rysunku oznaczenia są identyczne jak te na rysunku 5.1. W przeciwieństwie do wyników modelu z pamięcią jednokrokową, w tym przypadku histogram odtwarza (przynajmniej jakościowo) obie przekątne (diagonalę i antydiagonalę) wyraźnie widoczne na histogramach empirycznych zamieszczonych na rysunku 6.2.

6.2 Definicja modelu

Przechodzimy teraz do skonstruowania modelu z pamięcią dwukrokową. Przypomnijmy najpierw jak wygląda zależność trzech kolejnych skoków w ramach modelu z pamięcią jednokrokową, opartą na równaniu (5.1). W tym celu wystarczy przepisać równanie (6.1), zamieniając (dla uproszczenia) zmienne (r_n, r_{n-1}, r_{n-2}) na (r_3, r_2, r_1) oraz podzielić je obustronnie przez $h(r_1)$

$$h(r_3, r_2 | r_1) = (1 - \epsilon)^2 h(r_3)h(r_2) + \epsilon^2 \delta(r_3 + r_2)\delta(r_2 + r_1) + (1 - \epsilon) \epsilon \delta(r_3 + r_2)h(r_2) + (1 - \epsilon) \epsilon h(r_3)\delta(r_2 + r_1). \quad (6.6)$$

Otrzymana w ten sposób gęstość prawdopodobieństwa warunkowego dwóch kolejnych skoków r_3, r_2 , pod warunkiem skoku je poprzedzającego r_1 , składa się z ważonej sumy czterech elementów. Każdy z nich opisuje inną sytuację. Kolejno:

- składnik zawierający iloczyn gęstości $h(r_3)h(r_2)$ odpowiada za dwa niezależne skoki r_2 i r_3 ,
- drugi składnik zawierający $\delta(r_3 + r_2)\delta(r_2 + r_1)$ opisuje przypadek powtórzenia przez skok r_2 skoku r_1 z przeciwnym zwrotem, po czym powtórzenia przez skok r_3 skoku r_2 również ze zmianą zwrotu,
- składnik zawierający iloczyn $\delta(r_3 + r_2)h(r_2)$ odpowiada za wykonanie niezależnego skoku r_2 , a następnie powtórzenie przez skok r_3 skoku r_2 z przeciwnym zwrotem,
- ostatni składnik zawierający iloczyn $h(r_3)\delta(r_2 + r_1)$ opowiada za przypadek powtórzenia przez skok r_2 skoku r_1 ze zmienionym zwrotem, po czym wykonania niezależnego skoku r_3 .

Jak pokazano w poprzednim podrozdziale, założenia modelu z pamięcią jednokrokową są niewystarczające do opisu zależności trzech kolejnych skoków. Mianowicie, postulowana w równaniu (6.5) fenomenologiczna postać zależności co drugiego skoku, wsparta argumentem wywodzącym się z analizy bid-ask spreadu, zawierała składnik proporcjonalny do $\delta(r_3 + r_1)$ (czyli $\delta(r_n + r_{n-2})$ wyrażony we wspomnianych powyżej nowych oznaczeniach) a brakujący w wyrażeniu (6.6). Zatem, w nowym, rozwijanym tutaj modelu gęstość warunkowa $h(r_3, r_2 | r_1)$ powinna, dodatkowo w stosunku do (6.6), posiadać

- składnik zawierający iloczyn $\delta(r_3 + r_1)h(r_2)$ odpowiedzialny za wykonanie niezależnego skoku r_2 , a następnie powtórzenie przez skok r_3 skoku r_1 z przeciwnym zwrotem.

Teraz można już zaproponować następującą, bardziej ogólną postać gęstości prawdopodobieństwa warunkowego trzech kolejnych zmian ceny:

$$\begin{aligned}
h(r_3, r_2 | r_1) = & (1 - 2\epsilon)h(r_3)h(r_2) + \zeta\delta(r_3 + r_2)\delta(r_2 + r_1) + \\
& + (\epsilon - \zeta)\delta(r_3 + r_2)h(r_2) + (\epsilon - \zeta)\delta(r_2 + r_1)h(r_3) + \\
& + \zeta\delta(r_3 + r_1)h(r_2), \tag{6.7}
\end{aligned}$$

gdzie dla $\zeta = \epsilon^2$ drugi, trzeci i czwarty składnik odtwarzają swoje odpowiedniki w wyrażeniu (6.6), natomiast piąty składnik został dodany kosztem wagi składnika pierwszego.

Oczywiście, mnożąc równanie (6.7) przez $h(r_1)$ otrzymujemy łączną trójskokową gęstość prawdopodobieństwa (trzech zmiennych)

$$\begin{aligned}
h(r_3, r_2, r_1) = & (1 - 2\epsilon)h(r_3)h(r_2)h(r_1) + \zeta\delta(r_3 + r_2)\delta(r_2 + r_1)h(r_1) + \\
& + (\epsilon - \zeta)\delta(r_3 + r_2)h(r_2)h(r_1) + (\epsilon - \zeta)\delta(r_2 + r_1)h(r_3)h(r_1) + \\
& + \zeta\delta(r_3 + r_1)h(r_2)h(r_1). \tag{6.8}
\end{aligned}$$

Wagi w równaniach (6.7) i (6.8) zostały tak dobrane, żeby nowy model był zgodny z poprzednim, jednokrokowym w takim sensie, aby dla dwóch kolejnych zmian ceny odtwarzał zależność (5.1). Można to sprawdzić obliczając na podstawie (6.8), następujące, łączne dwuskokowe gęstości prawdopodobieństwa (dwóch zmiennych)

$$h(r_2, r_1) = \int_{-\infty}^{\infty} dr_3 h(r_3, r_2, r_1) = (1 - \epsilon) h(r_2)h(r_1) + \epsilon \delta(r_2 + r_1)h(r_1) \tag{6.9}$$

oraz

$$h(r_3, r_2) = \int_{-\infty}^{\infty} dr_1 h(r_3, r_2, r_1) = (1 - \epsilon) h(r_3)h(r_2) + \epsilon \delta(r_3 + r_2)h(r_2). \tag{6.10}$$

Ponadto, z (6.8) otrzymujemy zależność co drugiego skoku daną przez

$$\begin{aligned}
h_2(r_3, r_1) = & \int_{-\infty}^{\infty} dr_2 h(r_3, r_2, r_1) = (1 - 2\zeta) h(r_3)h(r_1) + \\
& + \zeta \delta(r_3 + r_1)h(r_1) + \zeta \delta(r_3 - r_1)h(r_1), \tag{6.11}
\end{aligned}$$

czyli dokładnie (jak trzeba) postać fenomenologicznej zależności opisanej równaniem (6.5).

Należy podkreślić, że zgodność obu modeli ma miejsce jedynie na poziomie rozkładów dwuskokowych (dwóch zmiennych), czyli dwóch następujących bezpośrednio po sobie skoków. Na poziomie trójskokowym i zależności co drugiego skoku, model z pamięcią dwukrokową wprowadza nową nieredukowalną jakość. W ogólności, oznacza to nieredukowalność wyrażenia (6.7) do (6.6).

Redukowalność formalną można zapewnić, wprowadzając w wyrażeniu (6.8) trzeci, formalny parametr θ . Wtedy przybiera ono postać:

$$\begin{aligned} h(r_3, r_2, r_1) = & (1 - 2\epsilon + \theta)h(r_1)h(r_2)h(r_3) + \\ & + \zeta\delta(r_3 + r_2)\delta(r_2 + r_1)h(r_1) + (\zeta - \theta)\delta(r_3 + r_1)h(r_2)h(r_1) \quad (6.12) \\ & + (\epsilon - \zeta)\delta(r_2 + r_1)h(r_3)h(r_1) + (\epsilon - \zeta)\delta(r_3 + r_2)h(r_2)h(r_1). \end{aligned}$$

Widać, że dla $\theta = 0$ wyrażenie (6.12) redukuje się do postaci (6.8).

Przy rozwiązywaniu modelu (czyli wyprowadzeniu jawnej postaci propagatora) niezbędne jest dysponowanie w jawnej postaci wyrażeniem na gęstość prawdopodobieństwa ostatniego skoku r_3 , pod warunkiem poprzedzających go dwóch skoków r_2, r_1 , czyli $h(r_3 | r_2, r_1)$. Wspomnianą wielkość możemy uzyskać dzieląc $h(r_3, r_2, r_1)$ przez $h(r_2, r_1)$. Jednak, napotykamy wówczas na komplikację techniczną dzielenia przez rozkład zawierający składnik proporcjonalny do delty Diraca. Dzielenie takie jest możliwe do wykonania, jednak wynik oprócz delt Diraca będzie zawierał również delty Kroneckera.

Wprowadzając dla uproszczenia zapisu zmienne

$$A = \left(\frac{1 - 2\epsilon + \theta}{1 - \epsilon} \right), \quad (6.13)$$

$$B = \left(\frac{\zeta - \theta}{1 - \epsilon} \right), \quad (6.14)$$

$$C = \left(\frac{\epsilon - \zeta}{1 - \epsilon} \right), \quad (6.15)$$

$$D = \left(\frac{\epsilon - \zeta}{\epsilon} \right), \quad (6.16)$$

$$E = \left(\frac{\zeta}{\epsilon} \right), \quad (6.17)$$

możemy warunkową gęstość prawdopodobieństwa $h(r_3 | r_2, r_1)$ wyrazić w postaci

$$\begin{aligned} h(r_3 | r_2, r_1) = & (1 - \delta_{r_2, -r_1}) [Ah(r_3) + B\delta(r_3 + r_1) + C\delta(r_3 + r_2)] + \\ & + \delta_{r_2, -r_1} [Dh(r_3) + E\delta(r_3 + r_2)]. \quad (6.18) \end{aligned}$$

Jak widać, gdy skok $r_2 = -r_1$ to w wyrażeniu tym obecna jest tylko część zawarta w drugim nawiasie kwadratowym. Natomiast, gdy $r_2 \neq -r_1$ to mamy do czynienia tylko z częścią zawartą w pierwszym nawiasie kwadratowym.

Poprawność formuły (6.18) można sprawdzić, np. mnożąc ją przez $h(r_2, r_1)$ co, korzystając z zależności

$$(1 - \delta_{r_2, -r_1})\delta(r_2 + r_1) = 0, \quad (6.19)$$

$$\delta_{r_2, -r_1}\delta(r_2 + r_1) = \delta(r_2 + r_1), \quad (6.20)$$

daje (jak trzeba) wynik opisany przez wyrażenie (6.12). Aby uchwycić rolę parametru θ zauważmy, że dla $\theta = 0$ otrzymujemy model z pamięcią dwukrokową, natomiast podstawiając $\zeta = \theta = \epsilon^2$ otrzymujemy model z pamięcią jednokrokową, opisany w poprzednim rozdziale.

6.3 Pamięć dwukrokowa w ramach CTRW

W podrozdziale 4.2, rozważając formalizm CTRW z pamięcią jednokrokową, zwróciłem uwagę iż nie można połączyć zależnością rekurencyjną ostrego propagatora $Q_n(X, t)$ z ostrym propagatorem $Q_{n-1}(X, t)$. Wynika to z faktu, że propagator $Q_{n-1}(X, t)$ nie niesie informacji o ostatnim skoku, od którego zależy przecież rozkład kolejnego skoku. W celu rozwiązania tego problemu we wspomnianym podrozdziale została wprowadzona pomocnicza, bardziej złożona wielkość $Q_n(X, r_n; t | \xi)$. Okazało się to wystarczające, ale tylko dla modelu z pamięcią jednokrokową.

W przypadku modelu z pamięcią dwukrokową, wielkość $Q_n(X, r_n; t | \xi)$ jest wciąż niewystarczająca do wprowadzenia zależności rekurencyjnej opartej o warunkową gęstość prawdopodobieństwa $h(r_n | r_{n-1}, r_{n-2})$. Mianowicie, ostry propagator $Q_{n-1}(X, r_{n-1}; t | \xi)$ niesie informacje jedynie o skoku r_{n-1} podczas, gdy kolejny skok r_n zależy teraz od dwóch poprzedzających go skoków r_{n-1} i r_{n-2} . Celem uzyskania właściwej, nowej bazowej zależności rekurencyjnej wprowadzam bardziej złożony ostry propagator

- $Q_n(X, r_n, r_{n-1}; t | \xi_0, \xi_{-1})$ - gęstość prawdopodobieństwa tego, że **dokładnie w chwili t proces wykona n -ty przeskok** z wartości $X - r_n$ do wartości X (wykonując skok o r_n) a w dowolnej wcześniejszej chwili wykonał $n - 1$ skok z wartości $X - r_n - r_{n-1}$ do wartości $X - r_n$ (wykonując skok o r_{n-1}) pod warunkiem, że w chwili początkowej ($t = 0$) był w zerze ($X = 0$), do którego przybył wykonując kolejno skoki ξ_{-1} oraz ξ_0 .

Ten nowozdefiniowany, ostry propagator wiąże z wcześniej wprowadzonym ostrym propagatorem następujące zależności

$$Q_n(X, r_n, r_{n-1}; t) = \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} d\xi_{-1} Q_n(X, r_n, r_{n-1}; t | \xi_0, \xi_{-1}) h(\xi_0, \xi_{-1}) \quad (6.21)$$

oraz

$$Q_n(X, t) = \int_{-\infty}^{\infty} dr_n \int_{-\infty}^{\infty} dr_{n-1} Q_n(X, r_n, r_{n-1}; t). \quad (6.22)$$

Dla dowolnej postaci $h(r_n | r_{n-1}, r_{n-2})$, przy użyciu nowozdefiniowanego propagatora, mogę wprowadzić zasadnicze równanie rekurencyjne (analogiczne do równań (3.9) i (4.7)) w postaci

$$\begin{aligned} Q_n(X, r_n, r_{n-1}; t | \xi_0, \xi_{-1}) &= \int_0^t dt' \psi(t') \int_{-\infty}^{\infty} dr_{n-2} h(r_n | r_{n-1}, r_{n-2}) \times \\ &\quad \times Q_{n-1}(X - r_n, r_{n-1}, r_{n-2}; t - t' | \xi_0, \xi_{-1}), \\ n &= 3, 4, \dots \end{aligned} \quad (6.23)$$

Jeżeli proces dociera do punktu X w chwili t , wykonując n -ty skok o r_n i pamiętając poprzedzający skok r_{n-1} , oznacza to, że do punktu $X - r_n$ dotarł w $n - 1$ skokach, wykonując skok o r_{n-1} . Równanie (6.23) uwzględnia wszystkie możliwe skoki r_{n-2} poprzedzające skok r_{n-1} oraz fakt, że gęstość prawdopodobieństwa wystąpienia ostatniego skoku r_n jest uwarunkowana dwoma poprzedzającymi skokami, czyli r_{n-1} i r_{n-2} .

Rozwiązanie powyższego równania (6.23) dla postaci funkcji $h(r_n | r_{n-1}, r_{n-2})$ danej równaniem (6.18) jest znacznie trudniejsze niż w przypadku jednokrokowym i wymaga wielu uciążliwych przekształceń. Znalezienie ścisłej postaci ostrego propagatora w postaci zamkniętej było możliwe, jednak ze względu na swoją złożoność zostało ono przeniesione do dodatku A. Właśnie ta trudność była m.in. powodem tego, że dopiero w niniejszej pracy zagadnienie błędzenia losowego w czasie ciągłym z pamięcią dwukrokową zostało ściśle rozwiązane.

W następnym podrozdziale omawiam wariancję i funkcję autokorelacji prędkości procesu, bazując na wyniku uzyskanym w dodatku A.

6.4 Wariancja i funkcja autokorelacji prędkości

Przedstawione poniżej wyniki opierają się na znajomości jawnej postaci propagatora dla modelu z pamięcią dwukrokową. Otrzymałem go z równania rekurencyjnego (6.23) oraz zależności warunkowej trzech kolejnych zmian ceny danej równaniem (6.18). Wyprowadzenie przedstawiono w dodatku A. Uzyskano tam ścisłą postać propagatora $\tilde{Q}(k, s)$ daną równaniem (A.56). Stąd na mocy równania (3.22), wyznaczyłem miękki propagator $\tilde{P}(k, s)$. Następnie, korzystając z (3.30), wyznaczam transformatę Laplace'a zależnej od czasu wariancji

$$\begin{aligned} \tilde{m}_2(s) &= - \left. \frac{\partial^2 \tilde{P}(k, s)}{\partial k^2} \right|_{k=0} = - \frac{\mu_2}{s^2 \langle t \rangle} \times \\ &\times \frac{\tilde{\psi}^3 (\zeta - \epsilon^2) (\zeta - \theta) + (\epsilon - 1)^2 \tilde{\psi} (2\epsilon^2 - \zeta) + \epsilon \tilde{\psi}^2 (2(\epsilon - 1)\epsilon - \zeta + 1)(\zeta - \theta) - (\epsilon - 1)^2 \epsilon}{\tilde{\psi}^3 (\zeta - \epsilon^2) (\zeta - \theta) + \epsilon \tilde{\psi}^2 (-2\epsilon + \zeta + 1)(\zeta - \theta) + (\epsilon - 1)^2 \zeta \tilde{\psi} + (\epsilon - 1)^2 \epsilon}, \end{aligned} \quad (6.24)$$

gdzie oczywiście $\tilde{\psi} = \tilde{\psi}(s)$, natomiast pozostałe wielkości są parametrami i nie zależą od s . Na podstawie (3.35) otrzymujemy transformatę Laplace'a funkcji autokorelacji prędkości procesu, w której uwzględniamy zależność funkcji od parametrów ζ i θ

$$\begin{aligned} C(s, \zeta, \theta) &= - \frac{\mu_2}{2 \langle t \rangle} \times \\ &\times \frac{\tilde{\psi}^3 (\zeta - \epsilon^2) (\zeta - \theta) + (\epsilon - 1)^2 \tilde{\psi} (2\epsilon^2 - \zeta) + \epsilon \tilde{\psi}^2 (2(\epsilon - 1)\epsilon - \zeta + 1)(\zeta - \theta) - (\epsilon - 1)^2 \epsilon}{\tilde{\psi}^3 (\zeta - \epsilon^2) (\zeta - \theta) + \epsilon \tilde{\psi}^2 (-2\epsilon + \zeta + 1)(\zeta - \theta) + (\epsilon - 1)^2 \zeta \tilde{\psi} + (\epsilon - 1)^2 \epsilon}. \end{aligned} \quad (6.25)$$

Przypominam, że dla $\zeta = \theta = \epsilon^2$ model z pamięcią dwukrokową odtwarza wyniki modelu z pamięcią jednokrokową (patrz rozdział 5). Zatem, dla takich szczególnych wartości parametrów otrzymuję

$$C(s, \zeta = \epsilon^2, \theta = \epsilon^2) = \frac{\mu_2}{2 \langle t \rangle} \frac{1 - \epsilon \tilde{\psi}(s)}{1 + \epsilon \tilde{\psi}(s)}, \quad (6.26)$$

czyli wynik dany wcześniej równaniem (5.18).

Aby uzyskać funkcję autokorelacji dla modelu z pamięcią dwukrokową, zdefiniowanego zależnością (6.7), wystarczy przyjąć wartość parametru $\theta = 0$

$$C(s, \zeta, \theta = 0) = \frac{\mu_2}{2 \langle t \rangle} \times \frac{\zeta \tilde{\psi}^3 (\epsilon^2 - \zeta) - (\epsilon - 1)^2 \tilde{\psi} (2\epsilon^2 - \zeta) + \epsilon \zeta \tilde{\psi}^2 (-2(\epsilon - 1)\epsilon + \zeta - 1) + (\epsilon - 1)^2 \epsilon}{\zeta \tilde{\psi}^3 (\zeta - \epsilon^2) + \epsilon \zeta \tilde{\psi}^2 (-2\epsilon + \zeta + 1) + (\epsilon - 1)^2 \zeta \tilde{\psi} + (\epsilon - 1)^2 \epsilon}. \quad (6.27)$$

Jak widać powyższy wynik ma dosyć skomplikowaną postać. Z tego powodu przeprowadzenie odwrotnej transformaty Laplace'a dla przypadku podwójnie eksponencjalnego rozkładu czasów wyczekiwania (danego wzorem (5.28)) jest trudne do przeprowadzenia. Aby uprościć to zadanie i pozostać w obszarze ścisłych rozwiązań, zastosuję wspomniane wcześniej przybliżenie $\zeta = \epsilon^2$ (dyskutowane w podrozdziale 6.1). Przybliżenie to prowadzi do

$$C(s) = C(s, \zeta = \epsilon^2, \theta = 0) = \frac{\mu_2}{2 \langle t \rangle} \frac{1 - \epsilon \tilde{\psi}(s) (1 + \epsilon \tilde{\psi}(s))}{1 + \epsilon \tilde{\psi}(s) (1 + \epsilon \tilde{\psi}(s))}. \quad (6.28)$$

Powyższe wyrażenie można jeszcze uprościć rozkładając ułamek występujący po jego prawej stronie na ułamki proste

$$\begin{aligned} C(s) &= \frac{\mu_2}{2 \langle t \rangle} \frac{1 - \epsilon \tilde{\psi}(s) - \epsilon^2 \tilde{\psi}^2(s)}{1 + \epsilon \tilde{\psi}(s) + \epsilon^2 \tilde{\psi}^2(s)} = \frac{\mu_2}{2 \langle t \rangle} \left(1 - 2 \frac{\epsilon \tilde{\psi}(s) + \epsilon^2 \tilde{\psi}^2(s)}{1 + \epsilon \tilde{\psi}(s) + \epsilon^2 \tilde{\psi}^2(s)} \right) = \\ &= \frac{\mu_2}{2 \langle t \rangle} \left(1 - 2 \frac{\epsilon \tilde{\psi}(s)(\epsilon \tilde{\psi}(s) + 1)}{(\epsilon \tilde{\psi}(s) - j)(\epsilon \tilde{\psi}(s) - \bar{j})} \right) = \\ &= \frac{\mu_2}{2 \langle t \rangle} \left[1 - 2 \left(\frac{j}{j-1} \frac{\epsilon \tilde{\psi}(s)}{\epsilon \tilde{\psi}(s) - j} + \frac{-1}{j-1} \frac{\epsilon \tilde{\psi}(s)}{\epsilon \tilde{\psi}(s) - \bar{j}} \right) \right], \end{aligned} \quad (6.29)$$

gdzie

$$j = -\frac{1}{2} + i \frac{\sqrt{3}}{2}, \quad (6.30)$$

$$j^3 = 1, \quad (6.31)$$

$$j^2 = \bar{j} = \frac{1}{j}. \quad (6.32)$$

Przy czym \bar{j} oznacza wielkość sprzężoną do j . Zauważmy, że

$$\frac{j}{j-1} = \frac{1}{2} - i\frac{\sqrt{3}}{6} = \lambda, \quad (6.33)$$

$$\frac{-1}{j-1} = \frac{1}{2} + i\frac{\sqrt{3}}{6} = \bar{\lambda}, \quad (6.34)$$

dzięki czemu możemy zapisać unormowaną funkcję autokorelacji prędkości procesu w ostatecznej postaci

$$C^n(t) = \delta(t) - 2\epsilon \left(\lambda \mathcal{L}_t^{-1} \left\{ \frac{\tilde{\psi}(s)}{\epsilon \tilde{\psi}(s) - j} \right\} + \bar{\lambda} \mathcal{L}_t^{-1} \left\{ \frac{\tilde{\psi}(s)}{\epsilon \tilde{\psi}(s) - \bar{j}} \right\} \right). \quad (6.35)$$

Zaskakującym może wydawać się fakt, iż znacznie bardziej skomplikowany model, którego rozwiązanie jest bardziej złożone (patrz dodatek A) prowadzi do postaci unormowanej funkcji autokorelacji prędkości tylko nieco bardziej skomplikowanej w stosunku do przypadku jednokrokowego, danego równaniem (5.20). Ponadto, w rozważanym przybliżeniu ($\zeta = \epsilon^2$) kształt funkcji autokorelacji prędkości zależy dokładnie od tych samych parametrów co w przypadku jednokrokowym. Jest on zatem zdeterminowany rozkładem czasów wyczekiwania $\tilde{\psi}(s)$ oraz parametrem ϵ .

6.5 Porównanie z danymi empirycznymi

Uzyskanie unormowanej funkcji autokorelacji prędkości procesu, w postaci danej wzorem (6.35), umożliwi przeprowadzenie porównania przewidywań teoretycznych przedstawianego modelu z odpowiednimi danymi empirycznymi. Jak zauważono w poprzednim podrozdziale, funkcja autokorelacji prędkości w modelu z pamięcią dwukrokową, zależy od tych samych parametrów co w modelu z pamięcią jednokrokową. W podrozdziale 5.6 omówiłem już metodę wyznaczania z danych empirycznych, opisanych w podrozdziale 2.1, parametru ϵ oraz zaproponowałem funkcje modelujące rozkład czasów wyczekiwania. Dlatego w dalszej części analizy opieram się już na tych rezultatach. Zwróćmy uwagę, że uzyskane z danych empirycznych wielkości nie zależą od liczby kroków pamięci rozpatrywanej w modelu. W dalszych rozważaniach pomijam modelowanie rozkładu czasów wyczekiwania rozkładem eksponencjalnym, danym wzorem (5.26), gdyż nawet w modelu z pamięcią jednokrokową okazał się on niewystarczający do właściwego odtworzenia przebiegu empirycznej funkcji autokorelacji prędkości (patrz rysunek 5.6).

Rozpatrzmy teraz postać równania (6.35) dla rozkładu czasów wyczekiwania będącego ważoną sumą dwóch rozkładów eksponencjalnych, danego wzorem (5.28). Podstawiając (5.29) do (6.35) i dokonując odwrotnej transformaty Laplace'a uzyskujemy

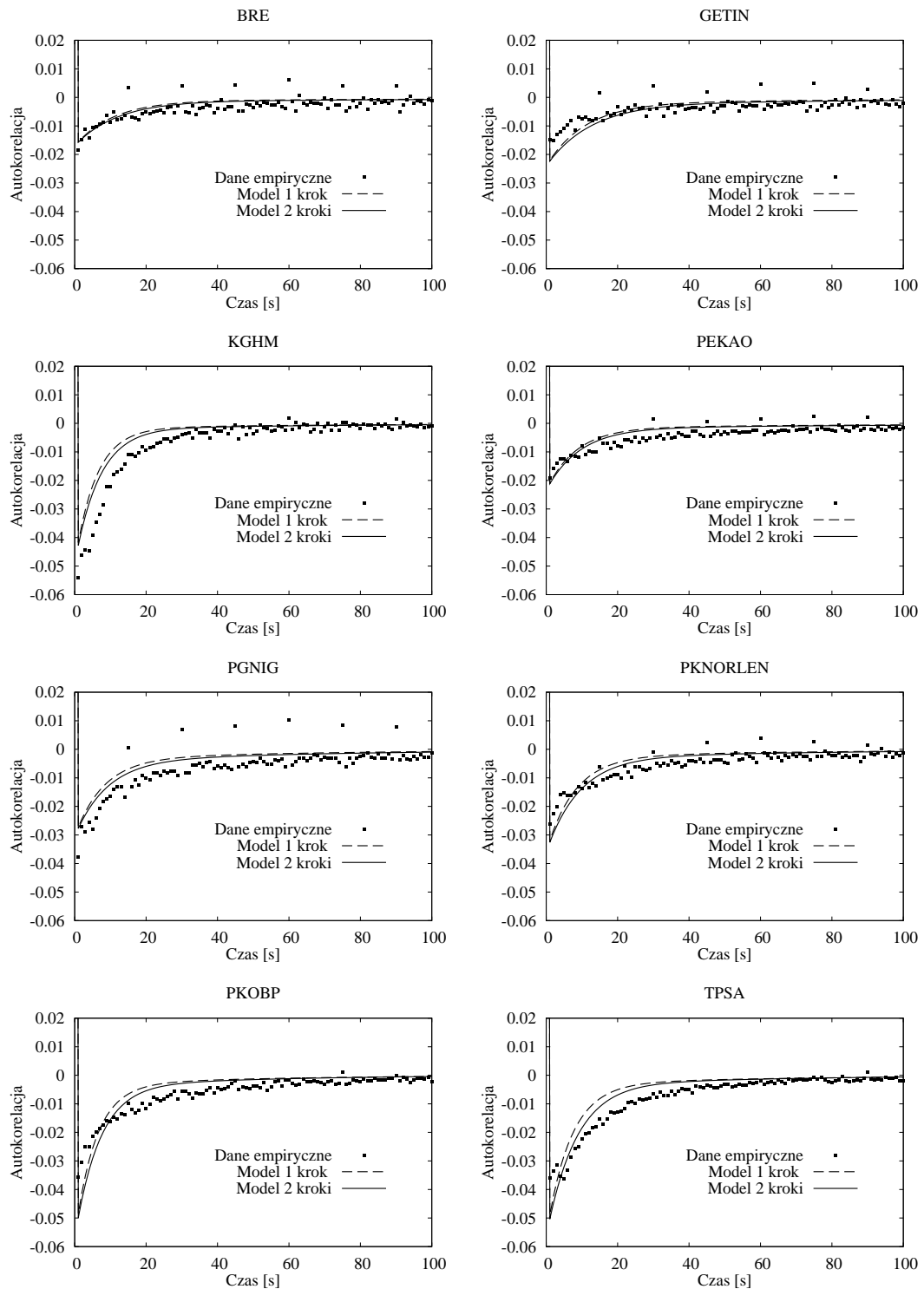
$$C(t) = \delta(t) - 2\epsilon \left[\lambda \left(A_1(j) e^{-\nu_1(j)t} + A_2(j) e^{-\nu_2(j)t} \right) + \bar{\lambda} \left(A_1(\bar{j}) e^{-\nu_1(\bar{j})t} + A_2(\bar{j}) e^{-\nu_2(\bar{j})t} \right) \right], \quad (6.36)$$

gdzie

$$\begin{aligned}
\nu_{1,2}(j) &= \frac{1}{2j} \left[j\omega_1 + j\omega_2 + \epsilon v \pm \sqrt{(j\omega_1 + j\omega_2 + \epsilon v)^2 - 4j\omega_1\omega_2(j + \epsilon)} \right], \\
A_1(j) &= \frac{\omega_1\omega_2 - \nu_1 v}{j(\nu_2 - \nu_1)}, \\
A_2(j) &= \frac{\omega_1\omega_2 - \nu_2 v}{j(\nu_1 - \nu_2)}, \\
\omega_1 &= 1/\tau_1, \\
\omega_2 &= 1/\tau_2, \\
v &= w\omega_1 + (1-w)\omega_2.
\end{aligned} \tag{6.37}$$

Jak widać podobieństwo równań (5.20) dla modelu z pamięcią jednokrokową i (6.35) dla modelu z pamięcią dwukrokową, pociąga za sobą podobieństwo rozwiązań (5.32) i (6.5), a nawet parametrów tych rozwiązań (5.33) i (6.37). Należy jednak zwrócić uwagę, że w obecnym przypadku parametry ν_1, ν_2, A_1, A_2 przyjmują wartości zespolone co komplikuje ich interpretację. Bazując na wartościach parametru ϵ podanych w drugiej kolumnie tabeli 5.1 i parametrach rozkładu czasów wyczekiwania z tabeli 5.2 możemy wykreślić, dla wszystkich badanych spółek, teoretyczną funkcję autokorelacji prędkości. Porównanie powyższej funkcji z danymi empirycznymi oraz przewidywaniami modelu z pamięcią jednokrokową (dla tego samego rozkładu czasów wyczekiwania danego wzorem (5.28)) przedstawia rysunek 6.5.

Na rysunku tym możemy zaobserwować poprawę przewidywań teoretycznych nowego modelu z pamięcią dwukrokową, w stosunku do modelu z pamięcią jednokrokową. O ile wartość autokorelacji prędkości dla wartości czasu bliskiej zeru ($t < 5$) oraz dla dłuższych czasów ($t > 40$) niemal nie uległa zmianie, to dla pośrednich czasów ($t \in [5, 40]$) **daje się zauważyć polepszenie, czyli zbliżenie przewidywań teoretycznych do danych empirycznych, zmniejszające tym samym systematyczne odstępstwo obserwowane w przypadku modelu z pamięcią jednokrokową.**



Rysunek 6.5: Unormowana funkcja autokorelacji prędkości ceny: empiryczna (czarne kwadraty), wyrażona wzorem (5.32) dla modelu z pamięcią jednokorokową (linia przerywana) i wyrażona wzorem (6.5) dla modelu z pamięcią dwukorokową (linia ciągła). Oba modele teoretyczne opierają się na podwójnie eksponencjalnym rozkładzie czasów wyczekiwania danym wzorem (5.28). Parametry krzywych teoretycznych zostały otrzymane osobnymi metodami estymacji. Zatem, przedstawione krzywe uzyskano bez potrzeby dopasowywania (fitowania) jakiegokolwiek parametru do danych empirycznych.

Rozdział 7

Zakończenie

7.1 Podsumowanie uzyskanych wyników

W niniejszym podrozdziale dokonuję podsumowania uzyskanych w rozprawie wyników. Składają się na nie wprowadzone metody i modele oraz dokonane porównania przewidywań tych modeli z danymi empirycznymi.

W rozdziale 2, poza opisem analizowanych w rozprawie danych empirycznych (podrozdział 2.1), przedstawiłem własną metodę obliczania empirycznej funkcji autokorelacji prędkości dla skokowych szeregów czasowych. We wspomnianej metodzie estymator funkcji autokorelacji prędkości wyraża się wzorem (2.7). Metoda ta posiada, w porównaniu z metodami dyskretyzującymi szereg czasowy ze stałym krokiem, kilka użytecznych cech. Przypomnę, że główne z nich to:

- możliwość działania z dowolną rozdzielczością czasową,
- dopuszczenie możliwości dwóch zmian wartości procesu w jednej chwili czasu,
- brak artefaktów powodowanych przez stały krok czasowy, czyli wyróżnienie częstotliwości odpowiadającej ustalonemu krokowi czasowemu oraz
- mniejsze zapotrzebowanie na pamięć operacyjną.

W podrozdziale 2.3 wprowadziłem oryginalne narzędzie do badania istotności uzyskanej funkcji autokorelacji, które nazwałem kwantylem autokorelacyjnym. Zaletą tego narzędzia jest brak jakichkolwiek założeń o postaci rozkładu uzyskiwanych wyników, tzn. przedział ufności można uzyskać bezpośrednio z danych przez ich podział na fragmenty. Uzyskany tą metodą przedział ufności może zależeć od wspomnianej fragmentyzacji. Jednak w przypadku analizowanych danych (opisanych w podrozdziale 2.1) istniał już naturalny podział na dni transakcyjne.

Rozdział 3 niniejszej rozprawy ma charakter monograficzny i ma na celu wprowadzenie w formalizm błędzenia losowego w czasie ciągłym (CTRW). Przedstawione tam wyprowadzenia, w szczególności stacjonaryzacja modelu, wykorzystywane są w następnych rozdziałach rozprawy do autorskich modeli opartych na formalizmie

CTRW. Zbudowanie nowych modeli było niezbędne, gdyż kanoniczny formalizm CTRW nie jest w stanie odtworzyć ujemnej części funkcji autokorelacji prędkości, przedstawionej na rysunku 2.2, a uzyskanej z analizowanych danych.

Rozdział 4 przedstawia przeprowadzoną przeze mnie, procedurę stacjonaryzacji modelu Montero i Masolivera [43] opartego na formalizmie CTRW. Model ten różni się od kanonicznego CTRW tym, że wprowadza zależność zwrotu kolejnej zmiany procesu, od zwrotu zmiany poprzedzającej. Zależność ta (4.3), będąca faktycznie pamięcią jednokrokową, nie jest jednak wystarczająca do odtworzenia empirycznej zależności dwóch kolejnych zmian wartości procesu, która dla badanych danych przedstawiona została na rysunku 5.1.

Dlatego w rozdziale 5 przedstawiam w pełni autorski model formalizmu CTRW z pamięcią jednokrokową, której postać dana jest wyrażeniem (5.1). Pragnę podkreślić, że motywacją do zbudowania modelu były zarówno, przedstawione w podrozdziale 5.1, wnioski wynikające bezpośrednio z danych empirycznych, jak i argumentacja teoretyczna podana w podrozdziale 5.2, wyjaśniająca na konkretnym przykładzie mechanizm powstawania zjawiska bid-ask bounce. Właśnie to zjawisko jest, moim zdaniem, źródłem tzw. ostrej korelacji powrotnej wyrażonej przez (5.1). Wspomniany model udało się ściśle rozwiązać wyprowadzając jawną postać propagatora, daną wyrażeniem (5.16), z której to uzyskałem średnie przesunięcie kwadratowe (5.17) oraz poszukiwaną funkcję autokorelacji prędkości procesu (5.20).

W dalszej części rozdziału 5 dyskutowałem estymację parametrów występujących w wyrażeniu na funkcję autokorelacji prędkości procesu ceny (5.20). Zaproponowałem m.in. postać funkcji modelującej rozkład czasów wyczekiwania, daną wzorem (5.28), będący ważoną sumą dwóch rozkładów eksponencjalnych. Dla zaproponowanej postaci wspomnianego rozkładu, funkcję autokorelacji prędkości uzyskałem w jawnej postaci wraz z parametrami wyznaczonymi bezpośrednio z danych empirycznych. Parametry te wyznaczyłem na rozłącznych zbiorach danych, dzięki temu uzyskałem funkcję, którą bezpośrednio porównałem z przedstawioną wcześniej empiryczną funkcją autokorelacji prędkości. Należy podkreślić, że porównanie to, przedstawione na rysunku 5.6, nie jest uzyskane metodą dopasowywania parametrów funkcji do wykresu. Zaproponowany model z pamięcią jednokrokową i podwójnie eksponencjalnym rozkładem czasów wyczekiwania odtwarza charakterystyczne zachowanie funkcji autokorelacji prędkości, przyjmującej ujemne wartości przez pierwszych kilkadziesiąt sekund. Niestety wciąż widoczne są systematyczne odstępstwa dla większości badanych spółek.

W związku z powyższym, w rozdziale 6 przedstawiłem kolejny autorski model, tym razem z pamięcią dwukrokową, czyli zależnością kolejnej zmiany wartości ceny od dwóch poprzedzających ją zmian. Model z pamięcią dwukrokową stanowi całkowicie oryginalną część teoretyczną niniejszej rozprawy, gdyż modele tego typu nie były dotychczas rozpatrywane. Motywacją do stworzenia modelu była zarówno chęć poprawienia zgodności przewidywań teoretycznych z danymi empirycznymi, jak również prawidłowy opis zależności co drugiej zmiany ceny. Dlatego zaproponowa-

łem gęstość prawdopodobieństwa trzech kolejnych zmian ceny w postaci wyrażenia (6.12).

W ramach tego modelu udało mi się uzyskać jawną postać propagatora procesu, wyprowadzoną w dodatku A. Co ciekawe, pomimo znacznie bardziej skomplikowanego wyprowadzenia, uzyskana w ramach modelu postać funkcji autokorelacji prędkości (wyrażenie (6.35)) jest tylko nieco bardziej skomplikowaną w porównaniu do postaci uzyskanej w modelu z pamięcią jednokrokową (wyrażenie (5.20)), zależąc dokładnie od tych samych wielkości. Umożliwiło to, analogicznie do poprzedniego modelu, przeprowadzenie porównania nowych przewidywań teoretycznych z danymi empirycznymi, co przedstawiono na rysunku 6.5. Rozszerzenie pamięci do dwukrokowej poprawiło obserwowaną zgodność przewidywań modelu z danymi empirycznymi.

7.2 Ocena wyników

Jak już wspomniałem, w niniejszej rozprawie przedstawiłem m.in. dwa w pełni oryginalne stochastyczne modele opisujące zachowanie ceny waloru giełdowego na najkrótszej możliwej skali czasu. Modele te opierają się na formalizmie błędzenia losowego w czasie ciągłym rozszerzając go o pamięć (zależność) jedno- i dwukrokową.

Na korzyść zaproponowanych modeli przemawia fakt, że ich założenia (głównie zależność dwóch kolejnych zmian ceny dana wyrażeniem (5.1)) są bezpośrednio inspirowane danymi empirycznymi. Z kolei formuły wyprowadzone w ramach modeli dobrze opisują analizowane dane. Niestety, ze względu na niedostępność danych, zawierających wszystkie składane na giełdzie zlecenia, mogłem tylko pośrednio wykazywać, iż źródłem obserwowanej zależności dwóch kolejnych zmian ceny jest faktycznie zjawisko bid-ask bounce, opisane w podrozdziale 5.2. Grupa fizyków z Uniwersytetu w Palermo i Instytutu Santa Fe dysponuje, co prawda, takimi danymi dla różnych giełd, ale zgodnie z warunkami na jakich dane te otrzymali, nie mogą ich udostępniać. Jednak wyniki ich analizy [64] potwierdzają, że ostra korelacja dwóch kolejnych skoków ma swoje źródło w odbijaniu się ceny pomiędzy krańcami bid-ask spreadu.

Kolejną istotną zaletą zbudowanych przeze mnie modeli jest fakt istnienia ścisłych rozwiązań. Ścisłe rozwiązania wraz z możliwością niezależnego wyznaczenia zawartych w nich parametrów, umożliwiają przeprowadzenie porównań przewidywań modelu z empirycznymi odpowiednikami. Porównania te **nie stanowią** już dopasowania (fitowania) parametrycznej zależności teoretycznej do wspomnianych odpowiedników. Pomimo, iż nie w każdym przypadku udało się dokładnie odtworzyć przebieg funkcji autokorelacji prędkości, co do kształtu i rzędu wielkości porównania te są co najmniej zadowalające, a w wielu przypadkach dobre. Wynika stąd, że **jeżeli inne czynniki, nieuwzględnione w rozwijanych przeze mnie modelach mają wpływ na kształt funkcji autokorelacji prędkości, to jest on wyższego rzędu niż uwzględniony wpływ efektu zależności dwóch kolejnych zmian ceny.**

Uzyskane przeze mnie postacie funkcji autokorelacji prędkości ((5.20) i (6.35)) pokazały jak funkcja ta zależy od rozkładu czasów wyczekiwania $\psi(t)$ oraz parametru ϵ , a także od czego ona nie zależy. Przykładowo, **brak w uzyskanych formułach zależności od postaci rozkładu pojedynczej zmiany ceny** jest interesujący, gdyż wynika z niego m.in. że funkcja autokorelacji znaku prędkości ceny powinna być identyczna z funkcją autokorelacji prędkości ceny. Porównanie tych funkcji przedstawiono w dodatku C. Jak widać, w połowie badanych przypadków wspomniane zależności mają zbliżony przebieg. W pozostałych przypadkach musi istnieć dodatkowy mechanizm powodujący rozbieżność.

Główne wnioski jakie można wyciągnąć z rozważań przedstawionych w niniejszej rozprawie są następujące:

- (a) zależność dwóch kolejnych zmian ceny jest główną przyczyną obserwowanej autokorelacji prędkości ceny. Jej źródłem jest zjawisko bid-ask bounce zatem, jest ono również przyczyną powstawania ujemnej autokorelacji. Inne czynniki mogą mieć jedynie wpływ wyższego rzędu. Nie jest to więc skutek reakcji graczy giełdowych na zmianę ceny, lecz samoistny efekt wynikający z mechanizmu transakcyjnego giełdy. Przypuszczenie, że obserwowana autokorelacja może być efektem interakcji między graczami jest dosyć powszechne [70].
- (b) Bid-ask spread jest możliwy do bezpośredniej obserwacji w czasie notowań ciągłych. Skoro obserwowana autokorelacja jest skutkiem efektu bid-ask bounce, widocznego w księdze zleceń, to zgodnie z zasadą eliminacji arbitrażu na giełdzie [32] na wspomnianej autokorelacji nie można uzyskać profitu.
- (c) Skoro źródło autokorelacji tkwi w księdze zleceń (patrz punkt (a)) to wszelkie modele giełdy bazujące na modelowaniu zachowań poszczególnych graczy giełdowych (tzw. agentów), o ile nie opierają się na księdze zleceń, nie powinny odtwarzać ujemnej części funkcji autokorelacji prędkości. Jeżeli taka część występuje to jest ona niepożądana.

Powyższe wnioski niosą bezpośrednie korzyści wynikające z rozważań przedstawionych w niniejszej rozprawie, gdyż są możliwe do bezpośredniego zastosowania w strategiach inwestycyjnych bądź w budowie nowych modeli giełdy.

Zbudowane przeze mnie modele mają dwa zasadnicze ograniczenia. Po pierwsze, mogą być zastosowane jedynie do opisu rynków opierających się na zasadach podwójnej aukcji z księgą zleceń. W przeciwnym wypadku efekt bid-ask bounce nie może wystąpić. Ma to np. miejsce na rynku walutowym FOREX. Drugim poważnym ograniczeniem jest założenie stacjonarności modelowanego szeregu czasowego. Założenie to jest bardzo silne zwłaszcza, że niestacjonarności giełdowego szeregu czasowego udokumentowana jest na różnych skalach czasu. W szczególności, na skali jednego dnia występuje tzw. efekt lunch'u, czyli większa zmienność (i częstsze transakcje) na początku i końcu dnia, a spowolnienie i mniejsza zmienności w porze lunchu (godz.

11-14). Wpływ tego efektu na wyniki przewidywań modeli omówiony został w dodatku B. Jego uwzględnienie nie ma istotnego znaczenia. Ponadto, przyjęte założenie o niezależności czasów wyczekiwania, zarówno między sobą, jak i ze zmianami ceny, jest silnym założeniem. Utrudniłoby ono dokładny opis procesu, w którym zależności takich zaniedbać nie można.

7.3 Konkluzje

Na podstawie przedstawionej powyżej oceny rozwiązania mogę stwierdzić, że **postawiony cel pracy, udało mi się osiągnąć**. Wprowadzone modele w sposób co najmniej zadowalający opisują, na poziomie funkcji autokorelacji prędkości, dynamikę ceny waloru giełdowego na najkrótszej skali czasowej.

W przedstawionych modelach udało się uchwycić, bazując na danych empirycznych, najistotniejsze zjawisko dominujące dynamikę stochastyczną na najkrótszej skali czasu rozpatrywaną poprzez autokorelację prędkości procesu. Zjawisko to leży u podstaw ściśle rozwiązywalnych modeli teoretycznych. Otwartym pytaniem pozostaje, czy jest to pełen opis zjawiska na tej skali czasowej? Mogą istnieć zjawiska, których nie widać na poziomie funkcji autokorelacji prędkości i do ich zbadania należy użyć np. nieliniowych autokorelacji, takich jak autokorelacja modułu prędkości procesu. Znalezienie wspomnianych autokorelacji w ramach zbudowanych przeze mnie modeli jest możliwe, co otwiera perspektywy dalszych badań w tej dziedzinie.

Wkład niniejszej rozprawy do rozwoju wiedzy można podzielić na trzy główne grupy:

- na wnioski zawarte w podrozdziale 7.2 dające możliwość bezpośredniego zastosowania w konstruowaniu nowych modeli opisujących dynamikę ceny spółek giełdowych.
- Wkład w rozwój teorii błędzenia losowego w czasie ciągłym z pamięcią, w szczególności wprowadzenie modelu z pamięcią dwukrokową. Modele tego typu mogą mieć szeroki zakres zastosowań, poza wymienionymi w pracy, w dziedzinach obecnie modelowanych kanonicznym formalizmem CTRW, jak również w opisie układów rządzonych regułą Le Chateliera–Browna.
- Wprowadzenie metody obliczania empirycznej funkcji autokorelacji prędkości dla skokowych szeregów czasowych, która w wielu przypadkach może okazać się efektywniejsza i dokładniejsza od metod stałokrokowych.

Dodatek A

Wyprowadzenie propagatora dla modelu z pamięcią dwukrokową

Wyprowadzenie ostrego propagatora dla modelu z pamięcią dwukrokową (zdefiniowaną poprzez wyrażenie (1.3)) jest kontynuacją rozważań przeprowadzonych w podrozdziale 6.3. Przedstawione jest tutaj rozwiązanie zależności rekurencyjnej danej wyrażeniem (6.23) dla przypadku stacjonarnego oraz warunkowej gęstości prawdopodobieństwa $h(r_3 | r_2, r_1)$ danej wyrażeniem (6.18).

Najpierw zapisuję bazowe równanie (6.23) dla ostrego propagatora, zdefiniowanego równaniem (6.21), w postaci

$$Q_n(X, r_n, r_{n-1}; t) = \int_0^t dt' \psi(t') \int_{-\infty}^{\infty} dr_{n-2} h(r_n | r_{n-1}, r_{n-2}) \times \\ \times Q_{n-1}(X - r_n, r_{n-1}, r_{n-2}; t - t'), \quad (\text{A.1})$$

gdzie obie strony równanie pomnożono przez $h(\xi_0, \xi_{-1})$ oraz wykonano całkowanie po zmiennych ξ_0 i ξ_{-1} . Dla skrócenia zapisu używać będę następującego oznaczenia

$$\tilde{Q}_n(r_n, r_{n-1}) \equiv \int_0^{\infty} dt \int_{-\infty}^{\infty} dX e^{ikX} e^{-st} Q_n(X, r_n, r_{n-1}; t), \quad (\text{A.2})$$

pomijającego jawną zależność ostrych propagatorów od zmiennych k i s .

Wyrażam teraz równanie (A.1) w zmiennych Fouriera–Laplace’a, używając wprowadzonej powyżej notacji

$$\tilde{Q}_n(r_n, r_{n-1}) = \tilde{\psi}(s) e^{ikr_n} \int_{-\infty}^{\infty} dr_{n-2} h(r_n | r_{n-1}, r_{n-2}) \tilde{Q}_{n-1}(r_{n-1}, r_{n-2}), \quad (\text{A.3})$$

co zamieniając zmienne, jak to zrobiono w podrozdziale 6.2, tzn. (r_n, r_{n-1}, r_{n-2}) na (r_3, r_2, r_1) , daje

$$\frac{1}{\tilde{\psi}(s)} \tilde{Q}_n(r_3, r_2) = e^{ikr_3} \int_{-\infty}^{\infty} dr_1 h(r_3 | r_2, r_1) \tilde{Q}_{n-1}(r_2, r_1). \quad (\text{A.4})$$

Kluczową operacją, umożliwiającą rozwiązanie równania (A.4) dla badanej w modelu postaci $h(r_3 | r_2, r_1)$ danej wyrażeniem (6.18), jest rozłożenie ostrego propagatora $\tilde{Q}_n(r_3, r_2)$ na dwie części, regularną i osobliwą (singularną), zdefiniowane odpowiednio przez

$$\tilde{Q}_n^R(r_2, r_1) = (1 - \delta_{r_2, -r_1}) \tilde{Q}_n(r_2, r_1), \quad (\text{A.5})$$

$$\tilde{Q}_n^S(r_2) = \int_{-\infty}^{\infty} dr_1 \delta_{r_2, -r_1} \tilde{Q}_n(r_2, r_1). \quad (\text{A.6})$$

Z tak zdefiniowanych propagatorów można odtworzyć $\tilde{Q}_n(r_3, r_2)$, korzystając z równania

$$\tilde{Q}_n(r_2, r_1) = \tilde{Q}_n^R(r_2, r_1) + \delta(r_2 + r_1) \tilde{Q}_n^S(r_2). \quad (\text{A.7})$$

Przekształcamy teraz równanie (A.4), podstawiając postać wyrażenia (6.18) oraz korzystając z definicji (A.5) i (A.6), do postaci

$$e^{-ikr_3} \frac{1}{\tilde{\psi}(s)} \tilde{Q}_n(r_3, r_2) = \tilde{Q}_{n-1}^S(r_2) [Dh(r_3) + E\delta(r_3 + r_2)] + B\tilde{Q}_{n-1}^R(r_2, -r_3) + \int_{-\infty}^{\infty} dr_1 \tilde{Q}_{n-1}^R(r_2, r_1) [Ah(r_3) + C\delta(r_3 + r_2)]. \quad (\text{A.8})$$

Po prawej stronie równania (A.8) występują już tylko propagatory regularne i osobliwe, jednak po lewej stronie mamy wciąż sumaryczny propagator $\tilde{Q}_n(r_3, r_2)$. Dążymy do uzyskania zależności rekurencyjnej dla propagatorów regularnych i osobliwych.

Rozłóżmy zatem powyższe równanie na dwie części. Najpierw pomnożmy obie strony równania przez $(1 - \delta_{r_3, -r_2})$

$$e^{-ikr_3} \frac{1}{\tilde{\psi}(s)} (1 - \delta_{r_3, -r_2}) \tilde{Q}_n(r_3, r_2) = \tilde{Q}_{n-1}^S(r_2) (1 - \delta_{r_3, -r_2}) Dh(r_3) + B(1 - \delta_{r_3, -r_2}) \tilde{Q}_{n-1}^R(r_2, -r_3) + (1 - \delta_{r_3, -r_2}) Ah(r_3) \int_{-\infty}^{\infty} dr_1 \tilde{Q}_{n-1}^R(r_2, r_1), \quad (\text{A.9})$$

przy czym, wyrazy zawierające $(1 - \delta_{r_3, -r_2}) \delta(r_3 + r_2)$ znikają. Następnie, pomnożmy obie strony równania przez $\delta_{r_3, -r_2}$, co daje

$$e^{-ikr_3} \frac{1}{\tilde{\psi}(s)} \delta_{r_3, -r_2} \tilde{Q}_n(r_3, r_2) = \tilde{Q}_{n-1}^S(r_2) [D\delta_{r_3, -r_2} h(r_3) + E\delta(r_3 + r_2)] + B\delta_{r_3, -r_2} \tilde{Q}_{n-1}^R(r_2, -r_3) + \int_{-\infty}^{\infty} dr_1 \tilde{Q}_{n-1}^R(r_2, r_1) [A\delta_{r_3, -r_2} h(r_3) + C\delta(r_3 + r_2)]. \quad (\text{A.10})$$

Przepisuję równania (A.9) i (A.10) w nowych zmiennych, całkując drugie z równań po zmiennej r_2 z pominięciem zbiorów miary zero¹

$$e^{-ikr_3} \frac{1}{\tilde{\psi}(s)} \tilde{Q}_n^R(r_3, r_2) = Dh(r_3) \tilde{Q}_{n-1}^S(r_2) + B \tilde{Q}_{n-1}^R(r_2, -r_3) + Ah(r_3) \int_{-\infty}^{\infty} dr_1 \tilde{Q}_{n-1}^R(r_2, r_1) \quad (\text{A.11})$$

oraz

$$e^{-ikr_3} \frac{1}{\tilde{\psi}(s)} \tilde{Q}_n^S(r_3) = E \tilde{Q}_{n-1}^S(-r_3) + C \int_{-\infty}^{\infty} dr_1 \tilde{Q}_{n-1}^R(-r_3, r_1). \quad (\text{A.12})$$

W ten oto sposób uzyskaliśmy układ równań rekurencyjnych na ostre propagatory regularne i osobliwe.

W celu rozwiązania tego układu równań wprowadzamy dodatkowe zmienne

$$\mathbf{R}_n(b, a) = \Re \int_{-\infty}^{\infty} dr_3 e^{ibkr_3} \int_{-\infty}^{\infty} dr_2 e^{iakr_2} \tilde{Q}_n^R(r_3, r_2), \quad (\text{A.13})$$

$$\mathbf{S}_n(a) = \Re \int_{-\infty}^{\infty} dr e^{iakr} \tilde{Q}_n^S(r), \quad (\text{A.14})$$

$$H(a) = \Re \int_{-\infty}^{\infty} dx e^{iakx} h(x), \quad (\text{A.15})$$

gdzie operator \Re zwraca część rzeczywistą liczby zespolonej. Zauważmy, że dzięki tym definicjom zachodzą następujące równości

$$\begin{aligned} \mathbf{R}_n(b, a) &= \mathbf{R}_n(-b, a) = \mathbf{R}_n(b, -a) = \mathbf{R}_n(-b, -a), \\ \mathbf{S}_n(a) &= \mathbf{S}_n(-a), \\ H(a) &= H(-a), \end{aligned} \quad (\text{A.16})$$

czyli istotny jest tutaj jedynie moduł argumentów powyższych funkcji.

Działając na równanie (A.11) obustronnie operatorem $\Re \int_{-\infty}^{\infty} dr_3 e^{ibkr_3} \int_{-\infty}^{\infty} dr_2 e^{iakr_2}$, a na równanie (A.12) operatorem $\Re \int_{-\infty}^{\infty} dr_3 e^{iakr_3}$ oraz używając zdefiniowanych wy-

¹Zakładamy tutaj, że rozkład $h(x)$ nie ma członów proporcjonalnych do delt Diraca. Jeżeli takie człony występują, należy uzyskać wynik przez przejście graniczne do rozkładu z osobliwościami przez rozkłady bez osobliwości.

rażeniami (A.13) i (A.14) zmiennych, uzyskujemy

$$\frac{1}{\tilde{\psi}(s)} \mathbf{R}_n(b-1, a) = D H(b) \mathbf{S}_{n-1}(a) + B \mathbf{R}_{n-1}(a, -b) + A H(b) \mathbf{R}_{n-1}(a, 0), \quad (\text{A.17})$$

$$\frac{1}{\tilde{\psi}(s)} \mathbf{S}_n(b-1) = E \mathbf{S}_{n-1}(-b) + C \mathbf{R}_{n-1}(-b, 0). \quad (\text{A.18})$$

Zauważmy, że na podstawie wyrażenia (6.22) w zmiennych Fouriera–Laplace’a, wyrażenia (A.7) oraz definicji (A.13), (A.14), zachodzi

$$\begin{aligned} \tilde{Q}_n(k, s) &= \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_1 \tilde{Q}_n(k, r_2, r_1; s) = \\ &= \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_1 \tilde{Q}_n^R(r_2, r_1) + \int_{-\infty}^{\infty} dr_2 \tilde{Q}_n^S(r_2) = \mathbf{R}_n(0, 0) + \mathbf{S}_n(0). \end{aligned} \quad (\text{A.19})$$

Zatem szukany, ostry propagator $\tilde{Q}_n(k, s)$ jesteśmy już w stanie wyrazić poprzez funkcje \mathbf{R}_n i \mathbf{S}_n ze wszystkimi argumentami równymi zero. Stąd, dążymy do takiego przekształcenia wyrażeń (A.17) i (A.18) aby występowały w nich powyższe funkcje jedynie od argumentów równych zero.

Zacznijmy od rozważenia równania (A.17) w czterech przypadkach, korzystając z własności (A.16)

$$\begin{aligned} \begin{pmatrix} a=0 \\ b=0 \end{pmatrix} &\Rightarrow \frac{1}{\tilde{\psi}(s)} \mathbf{R}_n(1, 0) = D H(0) \mathbf{S}_{n-1}(0) + B \mathbf{R}_{n-1}(0, 0) + A H(0) \mathbf{R}_{n-1}(0, 0), \\ \begin{pmatrix} a=0 \\ b=1 \end{pmatrix} &\Rightarrow \frac{1}{\tilde{\psi}(s)} \mathbf{R}_n(0, 0) = D H(1) \mathbf{S}_{n-1}(0) + B \mathbf{R}_{n-1}(0, 1) + A H(1) \mathbf{R}_{n-1}(0, 0), \\ \begin{pmatrix} a=1 \\ b=0 \end{pmatrix} &\Rightarrow \frac{1}{\tilde{\psi}(s)} \mathbf{R}_n(1, 1) = D H(0) \mathbf{S}_{n-1}(1) + B \mathbf{R}_{n-1}(1, 0) + A H(0) \mathbf{R}_{n-1}(1, 0), \\ \begin{pmatrix} a=1 \\ b=1 \end{pmatrix} &\Rightarrow \frac{1}{\tilde{\psi}(s)} \mathbf{R}_n(0, 1) = D H(1) \mathbf{S}_{n-1}(1) + B \mathbf{R}_{n-1}(1, 1) + A H(1) \mathbf{R}_{n-1}(1, 0). \end{aligned} \quad (\text{A.20})$$

W tym miejscu uzasadnionym jest przedefiniowanie współczynników A, B, C, D, E

tak, aby zawierały czynnik $\tilde{\psi}(s)$, czyli

$$\begin{aligned}
A &= \tilde{\psi}(s) \left(\frac{1 - 2\epsilon + \theta}{1 - \epsilon} \right), \\
B &= \tilde{\psi}(s) \left(\frac{\zeta - \theta}{1 - \epsilon} \right), \\
C &= \tilde{\psi}(s) \left(\frac{\epsilon - \zeta}{1 - \epsilon} \right), \\
D &= \tilde{\psi}(s) \left(\frac{\epsilon - \zeta}{\epsilon} \right), \\
E &= \tilde{\psi}(s) \left(\frac{\zeta}{\epsilon} \right).
\end{aligned} \tag{A.21}$$

Zamieniamy teraz numerację w układzie równań (A.20), mnożąc dodatkowo pierwsze, trzecie i czwarte równanie przez odpowiednio dobrane współczynniki

$$\begin{aligned}
B^3 \mathbf{R}_{n-3}(1, 0) &= B^3 D H(0) \mathbf{S}_{n-4}(0) + B^4 \mathbf{R}_{n-4}(0, 0) + B^3 A H(0) \mathbf{R}_{n-4}(0, 0), \\
\mathbf{R}_n(0, 0) &= D H(1) \mathbf{S}_{n-1}(0) + B \mathbf{R}_{n-1}(0, 1) + A H(1) \mathbf{R}_{n-1}(0, 0), \\
B^2 \mathbf{R}_{n-2}(1, 1) &= B^2 D H(0) \mathbf{S}_{n-3}(1) + B^3 \mathbf{R}_{n-3}(1, 0) + B^2 A H(0) \mathbf{R}_{n-3}(1, 0), \\
B \mathbf{R}_{n-1}(0, 1) &= B D H(1) \mathbf{S}_{n-2}(1) + B^2 \mathbf{R}_{n-2}(1, 1) + B A H(1) \mathbf{R}_{n-2}(1, 0).
\end{aligned} \tag{A.22}$$

Sumując powyższe równania stronami uzyskujemy

$$\begin{aligned}
\mathbf{R}_n(0, 0) &= B^3 D H(0) \mathbf{S}_{n-4}(0) + B^4 \mathbf{R}_{n-4}(0, 0) + B^3 A H(0) \mathbf{R}_{n-4}(0, 0) + \\
&+ D H(1) \mathbf{S}_{n-1}(0) + A H(1) \mathbf{R}_{n-1}(0, 0) + \\
&+ B^2 D H(0) \mathbf{S}_{n-3}(1) + B^2 A H(0) \mathbf{R}_{n-3}(1, 0) \\
&+ B D H(1) \mathbf{S}_{n-2}(1) + B A H(1) \mathbf{R}_{n-2}(1, 0).
\end{aligned} \tag{A.23}$$

Jak widać współczynniki występujące w wyrażeniach (A.20), przez które pomnożyliśmy powyższe równania, zostały tak dobrane, aby odpowiednie składniki zawierające funkcję \mathbf{R} od argumentów różnych od zera uległy skróceniu. W ostatnich dwóch wierszach wyrażenia (A.23) pozostały jeszcze dwa takie składniki (zawierające $\mathbf{R}_{n-3}(1, 0)$ oraz $\mathbf{R}_{n-2}(1, 0)$), które powtórnie możemy wyrazić, korzystając z (A.20), poprzez

$$\begin{aligned}
B^2 A H(0) \mathbf{R}_{n-3}(1, 0) &= A B^2 D H^2(0) \mathbf{S}_{n-4}(0) + A B^3 H(0) \mathbf{R}_{n-4}(0, 0) + \\
&+ A^2 B^2 H^2(0) \mathbf{R}_{n-4}(0, 0)
\end{aligned} \tag{A.24}$$

oraz

$$\begin{aligned}
B A H(1) \mathbf{R}_{n-2}(1, 0) &= A B D H(0) H(1) \mathbf{S}_{n-3}(0) + A B^2 H(1) \mathbf{R}_{n-3}(0, 0) + \\
&+ A^2 B H(0) H(1) \mathbf{R}_{n-3}(0, 0).
\end{aligned} \tag{A.25}$$

Zauważmy, że z faktu iż $H(0) = 1$ oraz podstawienia powyższych wyrażeń do (A.23), wynika, że

$$\begin{aligned} \mathbf{R}_n(0, 0) = & B^4 \mathbf{R}_{n-4}(0, 0) + AB^3 \mathbf{R}_{n-4}(0, 0) + A H(1) \mathbf{R}_{n-1}(0, 0) + \\ & + B^3 D \mathbf{S}_{n-4}(0) + D H(1) \mathbf{S}_{n-1}(0) + B^2 D \mathbf{S}_{n-3}(1) + BD H(1) \mathbf{S}_{n-2}(1) + \\ & + AB^3 \mathbf{R}_{n-4}(0, 0) + A^2 B^2 \mathbf{R}_{n-4}(0, 0) + AB^2 D \mathbf{S}_{n-4}(0) + \\ & + AB^2 H(1) \mathbf{R}_{n-3}(0, 0) + A^2 B H(1) \mathbf{R}_{n-3}(0, 0) + ABD H(1) \mathbf{S}_{n-3}(0), \end{aligned} \quad (\text{A.26})$$

czyli postać równania, w której funkcja \mathbf{R} występuje już tylko od zerowych wartości argumentów. Po uporządkowaniu podobnych składników, uzyskujemy:

$$\begin{aligned} \mathbf{R}_n(0, 0) = & + A H(1) \mathbf{R}_{n-1}(0, 0) + B^2(A + B)^2 \mathbf{R}_{n-4}(0, 0) + \\ & + AB(A + B) H(1) \mathbf{R}_{n-3}(0, 0) + \\ & + AB^2 D \mathbf{S}_{n-4}(0) + ABD H(1) \mathbf{S}_{n-3}(0) + \\ & + DB^3 \mathbf{S}_{n-4}(0) + D H(1) \mathbf{S}_{n-1}(0) + \\ & + DB^2 \mathbf{S}_{n-3}(1) + DB H(1) \mathbf{S}_{n-2}(1). \end{aligned} \quad (\text{A.27})$$

Powyższa postać zawiera jeszcze składniki proporcjonalne do funkcji \mathbf{S} od argumentu równego 1. Możemy je inaczej wyrazić używając równania (A.18) dla $b = 0$ (z nowymi współczynnikami zdefiniowanymi wyrażeniem (A.21)), czyli

$$\mathbf{S}_n(1) = E \mathbf{S}_{n-1}(0) + C \mathbf{R}_{n-1}(0, 0). \quad (\text{A.28})$$

Zatem

$$DB^2 \mathbf{S}_{n-3}(1) = DB^2 E \mathbf{S}_{n-4}(0) + DB^2 C \mathbf{R}_{n-4}(0, 0), \quad (\text{A.29})$$

$$DB H(1) \mathbf{S}_{n-2}(1) = DBE H(1) \mathbf{S}_{n-3}(0) + DBC H(1) \mathbf{R}_{n-3}(0, 0), \quad (\text{A.30})$$

co po podstawieniu do wyrażenia (A.27) daje

$$\begin{aligned} \mathbf{R}_n(0, 0) = & + A H(1) \mathbf{R}_{n-1}(0, 0) + B^2(A + B)^2 \mathbf{R}_{n-4}(0, 0) + \\ & + AB(A + B) H(1) \mathbf{R}_{n-3}(0, 0) + \\ & + AB^2 D \mathbf{S}_{n-4}(0) + ABD H(1) \mathbf{S}_{n-3}(0) + \\ & + DB^3 \mathbf{S}_{n-4}(0) + D H(1) \mathbf{S}_{n-1}(0) + \\ & + DB^2 E \mathbf{S}_{n-4}(0) + DB^2 C \mathbf{R}_{n-4}(0, 0) + \\ & + DBE H(1) \mathbf{S}_{n-3}(0) + DBC H(1) \mathbf{R}_{n-3}(0, 0). \end{aligned} \quad (\text{A.31})$$

Powtórne uporządkowanie podobnych składników prowadzi do

$$\begin{aligned} \mathbf{R}_n(0, 0) = & + A H(1) \mathbf{R}_{n-1}(0, 0) + D H(1) \mathbf{S}_{n-1}(0) \\ & + [B^2(A + B)^2 + DB^2 C] \mathbf{R}_{n-4}(0, 0) + \\ & + [AB(A + B) + DBC] H(1) \mathbf{R}_{n-3}(0, 0) + \\ & + DB^2 (A + B + E) \mathbf{S}_{n-4}(0) + BD (A + E) H(1) \mathbf{S}_{n-3}(0), \end{aligned} \quad (\text{A.32})$$

czyli do żądanej postaci, w której wszystkie funkcje \mathbf{R} i \mathbf{S} występują jedynie od zerowych argumentów.

Aby wyprowadzić drugie, niezbędne równanie rekurencyjne, przepisuję zależność (A.18) (z nowo zdefiniowanymi współczynnikami) w postaci

$$\mathbf{S}_n(b-1) = E \mathbf{S}_{n-1}(-b) + C \mathbf{R}_{n-1}(-b, 0). \quad (\text{A.33})$$

Dla $b = 1$ daje to

$$\mathbf{S}_n(0) = E \mathbf{S}_{n-1}(1) + C \mathbf{R}_{n-1}(1, 0), \quad (\text{A.34})$$

gdzie składnik zawierający $\mathbf{S}_{n-1}(1)$ uzyskujemy z wyrażenia (A.28). Prowadzi to, do wyrażenia

$$\mathbf{S}_n(0) = E (E \mathbf{S}_{n-2}(0) + C \mathbf{R}_{n-2}(0, 0)) + C \mathbf{R}_{n-1}(1, 0). \quad (\text{A.35})$$

Składnik zawierający $\mathbf{R}_{n-1}(1, 0)$, jak poprzednio, możemy wyznaczyć z wyrażenia (A.20)

$$\mathbf{R}_{n-1}(1, 0) = D \mathbf{S}_{n-2}(0) + (B + A) \mathbf{R}_{n-2}(0, 0), \quad (\text{A.36})$$

co prowadzi do ostatecznej postaci drugiego, brakującego równania rekurencyjnego

$$\mathbf{S}_n(0) = E (E \mathbf{S}_{n-2}(0) + C \mathbf{R}_{n-2}(0, 0)) + C (D \mathbf{S}_{n-2}(0) + (B + A) \mathbf{R}_{n-2}(0, 0)). \quad (\text{A.37})$$

Udało się zatem uzyskać ostateczny układ dwóch równań rekurencyjnych, w których wszystkie funkcje \mathbf{R} i \mathbf{S} występują jedynie od zerowych argumentów, a mianowicie

$$\begin{aligned} \mathbf{S}_n(0) &= (E^2 + CD)\mathbf{S}_{n-2}(0) + C(A + B + E) \mathbf{R}_{n-2}(0, 0), \\ \mathbf{R}_n(0, 0) &= + A H(1) \mathbf{R}_{n-1}(0, 0) + D H(1) \mathbf{S}_{n-1}(0) \\ &+ [B^2(A + B)^2 + DB^2C] \mathbf{R}_{n-4}(0, 0) + \\ &+ [AB(A + B) + DBC] H(1)\mathbf{R}_{n-3}(0, 0) + \\ &+ DB^2 (A + B + E) \mathbf{S}_{n-4}(0) + BD (A + E) H(1) \mathbf{S}_{n-3}(0). \end{aligned} \quad (\text{A.38})$$

Dalej, wprowadzamy upraszczające oznaczenia

$$\mathbf{S}_n \equiv \mathbf{S}_n(0), \quad (\text{A.39})$$

$$\mathbf{R}_n \equiv \mathbf{R}_n(0, 0), \quad (\text{A.40})$$

$$\mathbf{S} = \sum_{n=1}^{\infty} \mathbf{S}_n, \quad (\text{A.41})$$

$$\mathbf{R} = \sum_{n=1}^{\infty} \mathbf{R}_n, \quad (\text{A.42})$$

oraz przeprowadzamy sumowanie stronami układu równań (A.38), w którym pierwsze równanie sumujemy od $n = 3$ do ∞ , natomiast drugie od $n = 5$ do ∞ . Uzyskujemy

$$\begin{aligned}
(\mathbf{S} - \mathbf{S}_1 - \mathbf{S}_2) &= (E^2 + CD)\mathbf{S} + C(A + B + E)\mathbf{R}, \\
(\mathbf{R} - \mathbf{R}_1 - \mathbf{R}_2 - \mathbf{R}_3 - \mathbf{R}_4) &= + A H(1) (\mathbf{R} - \mathbf{R}_1 - \mathbf{R}_2 - \mathbf{R}_3) + \\
&\quad + D H(1) (\mathbf{S} - \mathbf{S}_1 - \mathbf{S}_2 - \mathbf{S}_3) \\
&\quad + [B^2(A + B)^2 + DB^2C]\mathbf{R} + \\
&\quad + [AB(A + B) + DBC] H(1) (\mathbf{R} - \mathbf{R}_1) + \\
&\quad + DB^2(A + B + E)\mathbf{S} + BD(A + E)H(1) (\mathbf{S} - \mathbf{S}_1).
\end{aligned} \tag{A.43}$$

Po uporządkowaniu i wprowadzeniu $\tilde{h} = \tilde{h}(k) = H(1)$ uzyskujemy układ dwóch równań w postaci

$$\begin{aligned}
0 &= \mathbf{S}_1 + \mathbf{S}_2 + (E^2 + CD - 1)\mathbf{S} + C(A + B + E)\mathbf{R}, \\
0 &= \mathbf{R} \left[-1 + A \tilde{h} + B^2(A + B)^2 + DB^2C + (AB(A + B) + DBC) \tilde{h} \right] + \\
&\quad + \mathbf{S} \left[D \tilde{h} + DB^2(A + B + E) + BD(A + E) \tilde{h} \right] + \\
&\quad + \mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3 + \mathbf{R}_4 + \\
&\quad + A (-\mathbf{R}_1 - \mathbf{R}_2 - \mathbf{R}_3) \tilde{h} + [AB(A + B) + DBC] (-\mathbf{R}_1) \tilde{h} + \\
&\quad + D (-\mathbf{S}_1 - \mathbf{S}_2 - \mathbf{S}_3) \tilde{h} + BD(A + E) (-\mathbf{S}_1) \tilde{h}.
\end{aligned} \tag{A.44}$$

Wyznaczając w postaci jawnej niewiadome wielkości $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \mathbf{R}_4, \mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$ występujące w powyższym równaniu, sprowadzamy znalezienie propagatora do rozwiązania układu dwóch równań z dwiema niewiadomymi \mathbf{R} i \mathbf{S} .

W tym celu, niezbędna jest znajomość pierwszych czterech ostrych propagatorów. Analogicznie do pierwszych dwóch propagatorów ostrych wyprowadzonych w podrozdziale 5.4 (bazując na schemacie przedstawionym na rysunku 5.4), dają się one wyrazić w przestrzeni Laplace'a, poprzez

$$\begin{aligned}
\tilde{Q}_1(X, s | \xi_0, \xi_{-1}) &= \tilde{\psi}_1(s)h(X | \xi_0, \xi_{-1}), \\
\tilde{Q}_2(X, s | \xi_0, \xi_{-1}) &= \tilde{\psi}_1(s)\tilde{\psi}(s) \int_{-\infty}^{\infty} dr_1 h(X - r_1 | r_1, \xi_0)h(r_1 | \xi_0, \xi_{-1}), \\
\tilde{Q}_3(X, s | \xi_0, \xi_{-1}) &= \tilde{\psi}_1(s)\tilde{\psi}^2(s) \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 h(X - r_2 - r_1 | r_2, r_1) \times \\
&\quad \times h(r_2 | r_1, \xi_0)h(r_1 | \xi_0, \xi_{-1}), \\
\tilde{Q}_4(X, s | \xi_0, \xi_{-1}) &= \tilde{\psi}_1(s)\tilde{\psi}^3(s) \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 h(X - r_3 - r_2 - r_1 | r_3, r_2) \times \\
&\quad \times h(r_3 | r_2, r_1)h(r_2 | r_1, \xi_0)h(r_1 | \xi_0, \xi_{-1}).
\end{aligned} \tag{A.45}$$

Uśredniamy teraz powyższy układ równań po skokach poprzedzających chwilę $t = 0$, czyli ξ_0 i ξ_{-1} . Prowadzi to, do

$$\begin{aligned}
\frac{\tilde{Q}_1(X, s)}{\tilde{\psi}_1(s)} &= \int_{-\infty}^{\infty} d\xi_{-1} \int_{-\infty}^{\infty} d\xi_0 h(X, \xi_0, \xi_{-1}), \\
\frac{\tilde{Q}_2(X, s)}{\tilde{\psi}_1(s)\tilde{\psi}(s)} &= \int_{-\infty}^{\infty} d\xi_{-1} \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dr_1 h(X - r_1 | r_1, \xi_0)h(r_1, \xi_0, \xi_{-1}), \\
\frac{\tilde{Q}_3(X, s)}{\tilde{\psi}_1(s)\tilde{\psi}^2(s)} &= \int_{-\infty}^{\infty} d\xi_{-1} \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 h(X - r_2 - r_1 | r_2, r_1) \times \\
&\quad \times h(r_2 | r_1, \xi_0)h(r_1, \xi_0, \xi_{-1}), \\
\frac{\tilde{Q}_4(X, s)}{\tilde{\psi}_1(s)\tilde{\psi}^3(s)} &= \int_{-\infty}^{\infty} d\xi_{-1} \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 h(X - r_3 - r_2 - r_1 | r_3, r_2) \times \\
&\quad \times h(r_3 | r_2, r_1)h(r_2 | r_1, \xi_0)h(r_1, \xi_0, \xi_{-1}).
\end{aligned} \tag{A.46}$$

Po wykonaniu niektórych całek uzyskujemy poniższą postać

$$\begin{aligned}
\frac{\tilde{Q}_1(X, s)}{\tilde{\psi}_1(s)} &= \int_{-\infty}^{\infty} d\xi_0 h(X, \xi_0), \\
\frac{\tilde{Q}_2(X, s)}{\tilde{\psi}_1(s)\tilde{\psi}(s)} &= \int_{-\infty}^{\infty} dr_1 h(X - r_1, r_1), \\
\frac{\tilde{Q}_3(X, s)}{\tilde{\psi}_1(s)\tilde{\psi}^2(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 h(X - r_2 - r_1, r_2, r_1), \\
\frac{\tilde{Q}_4(X, s)}{\tilde{\psi}_1(s)\tilde{\psi}^3(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 h(X - r_3 - r_2 - r_1 | r_3, r_2)h(r_3, r_2, r_1).
\end{aligned} \tag{A.47}$$

Dalej, przechodząc do zmiennej Fouriera otrzymujemy

$$\begin{aligned}
\frac{\tilde{Q}_1(k, s)}{\tilde{\psi}_1(s)} &= \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dX e^{ikX} h(X, \xi_0), \\
\frac{\tilde{Q}_2(k, s)}{\tilde{\psi}_1(s)\tilde{\psi}(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} h(X, r_1), \\
\frac{\tilde{Q}_3(k, s)}{\tilde{\psi}_1(s)\tilde{\psi}^2(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} h(X, r_2, r_1), \\
\frac{\tilde{Q}_4(k, s)}{\tilde{\psi}_1(s)\tilde{\psi}^3(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} e^{ikr_3} h(X | r_3, r_2) h(r_3, r_2, r_1),
\end{aligned} \tag{A.48}$$

co daje nam postać umożliwiającą, obliczenie szukanych $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \mathbf{R}_4, \mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$. Na podstawie (A.39) i (A.40), (A.13) i (A.14), (A.5) i (A.6) oraz (A.2), wyrażają się one kolejno przez

$$\begin{aligned}
\frac{\mathbf{S}_1}{\tilde{\psi}_1(s)} &= \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dX e^{ikX} \delta_{X, -\xi_0} h(X, \xi_0) = \\
&= \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dX e^{ikX} \epsilon \delta(X + \xi_0) h(\xi_0) = \epsilon \int_{-\infty}^{\infty} dX e^{ikX} h(-X) = \epsilon \tilde{h},
\end{aligned} \tag{A.49}$$

$$\begin{aligned}
\frac{\mathbf{R}_1}{\tilde{\psi}_1(s)} &= \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dX e^{ikX} (1 - \delta_{X, -\xi_0}) h(X, \xi_0) = \\
&= \int_{-\infty}^{\infty} d\xi_0 \int_{-\infty}^{\infty} dX e^{ikX} (1 - \epsilon) h(X) h(\xi_0) = (1 - \epsilon) \tilde{h},
\end{aligned} \tag{A.50}$$

$$\begin{aligned}
\frac{\mathbf{S}_2}{\tilde{\psi}_1(s)\tilde{\psi}(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} \delta_{X,-r_1} h(X, r_1) = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} \epsilon \delta(X + r_1) h(r_1) = \\
&= \epsilon \int_{-\infty}^{\infty} dr_1 e^{ik(-r_1)} e^{ikr_1} h(r_1) = \epsilon,
\end{aligned} \tag{A.51}$$

$$\begin{aligned}
\frac{\mathbf{R}_2}{\tilde{\psi}_1(s)\tilde{\psi}(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} (1 - \delta_{X,-r_1}) h(X, r_1) = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} (1 - \epsilon) h(X) h(r_1) = (1 - \epsilon) \tilde{h}^2,
\end{aligned} \tag{A.52}$$

$$\begin{aligned}
\frac{\mathbf{S}_3}{\tilde{\psi}_1(s)\tilde{\psi}^2(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} \delta_{X,-r_2} h(X, r_2, r_1) = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} \\
&\quad [\zeta \delta(X + r_2) \delta(r_2 + r_1) h(r_1) + (\epsilon - \zeta) \delta(X + r_2) h(r_2) h(r_1)] = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 e^{ikr_1} [\zeta \delta(r_2 + r_1) h(r_1) + (\epsilon - \zeta) h(r_2) h(r_1)] = \tilde{\epsilon} \tilde{h},
\end{aligned} \tag{A.53}$$

$$\begin{aligned}
\frac{\mathbf{R}_3}{\tilde{\psi}_1(s)\tilde{\psi}^2(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} (1 - \delta_{X,-r_2}) h(X, r_2, r_1) = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} [(1 - 2\epsilon + \theta) h(r_1) h(r_2) h(X) + \\
&\quad (\zeta - \theta) \delta(X + r_1) h(r_2) h(r_1) + (\epsilon - \zeta) \delta(r_2 + r_1) h(X) h(r_1)] = \\
&= (1 - 2\epsilon + \theta) \tilde{h}^3 + (\zeta - \theta) \tilde{h} + (\epsilon - \zeta) \tilde{h} = (1 - 2\epsilon + \theta) \tilde{h}^3 + (\epsilon - \theta) \tilde{h},
\end{aligned} \tag{A.54}$$

$$\begin{aligned}
\frac{\mathbf{R}_4}{\tilde{\psi}_1(s)\tilde{\psi}^3(s)} &= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} e^{ikr_3} \\
&\quad (1 - \delta_{X,-r_3}) h(X | r_3, r_2) h(r_3, r_2, r_1) = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} e^{ikr_3} (1 - \delta_{X,-r_3}) \\
&\quad \{(1 - \delta_{r_3,-r_2}) [Ah(X) + B\delta(X + r_2) + C\delta(X + r_3)] + \\
&\quad + \delta_{r_3,-r_2} [Dh(X) + E\delta(X + r_3)]\} \\
&\quad [(1 - 2\epsilon + \theta)h(r_1)h(r_2)h(r_3) + \zeta\delta(r_3 + r_2)\delta(r_2 + r_1)h(r_1) + \\
&\quad + (\zeta - \theta)\delta(r_3 + r_1)h(r_2)h(r_1) + \\
&\quad + (\epsilon - \zeta)\delta(r_2 + r_1)h(r_3)h(r_1) + (\epsilon - \zeta)\delta(r_3 + r_2)h(r_2)h(r_1)] = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} e^{ikr_3} \\
&\quad \{(1 - \delta_{r_3,-r_2}) [Ah(X) + B\delta(X + r_2)] + \delta_{r_3,-r_2} [Dh(X)]\} \\
&\quad [(1 - 2\epsilon + \theta)h(r_1)h(r_2)h(r_3) + \zeta\delta(r_3 + r_2)\delta(r_2 + r_1)h(r_1) + \\
&\quad + (\zeta - \theta)\delta(r_3 + r_1)h(r_2)h(r_1) + \\
&\quad + (\epsilon - \zeta)\delta(r_2 + r_1)h(r_3)h(r_1) + (\epsilon - \zeta)\delta(r_3 + r_2)h(r_2)h(r_1)] = \\
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} e^{ikr_3} \\
&\quad \{[Ah(X) + B\delta(X + r_2)] (1 - 2\epsilon + \theta)h(r_1)h(r_2)h(r_3) + \\
&\quad + [Dh(X)] \zeta\delta(r_3 + r_2)\delta(r_2 + r_1)h(r_1) + \\
&\quad + [Ah(X) + B\delta(X + r_2)] (\zeta - \theta)\delta(r_3 + r_1)h(r_2)h(r_1) + \\
&\quad + [Ah(X) + B\delta(X + r_2)] (\epsilon - \zeta)\delta(r_2 + r_1)h(r_3)h(r_1) + \\
&\quad + [Dh(X)] (\epsilon - \zeta)\delta(r_3 + r_2)h(r_2)h(r_1)\} =
\end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} dr_1 \int_{-\infty}^{\infty} dr_2 \int_{-\infty}^{\infty} dr_3 \int_{-\infty}^{\infty} dX e^{ikX} e^{ikr_1} e^{ikr_2} e^{ikr_3} \\
&\quad [A(1 - 2\epsilon + \theta)h(X)h(r_1)h(r_2)h(r_3) + \\
&\quad + B(1 - 2\epsilon + \theta)\delta(X + r_2)h(r_1)h(r_2)h(r_3) + \\
&\quad + D\zeta h(X)\delta(r_3 + r_2)\delta(r_2 + r_1)h(r_1) + \\
&\quad + A(\zeta - \theta)h(X)\delta(r_3 + r_1)h(r_2)h(r_1) + \\
&\quad + B(\zeta - \theta)\delta(X + r_2)\delta(r_3 + r_1)h(r_2)h(r_1) + \\
&\quad + A(\epsilon - \zeta)h(X)\delta(r_2 + r_1)h(r_3)h(r_1) + \\
&\quad + B(\epsilon - \zeta)\delta(X + r_2)\delta(r_2 + r_1)h(r_3)h(r_1) + \\
&\quad + D(\epsilon - \zeta)h(X)\delta(r_3 + r_2)h(r_2)h(r_1)] = \\
&= A(1 - 2\epsilon + \theta)\tilde{h}^4 + B(1 - 2\epsilon + \theta)\tilde{h}^2 + D\zeta\tilde{h}^2 + A(\zeta - \theta)\tilde{h}^2 + \\
&\quad + B(\zeta - \theta) + A(\epsilon - \zeta)\tilde{h}^2 + B(\epsilon - \zeta)\tilde{h}^2 + D(\epsilon - \zeta)\tilde{h}^2 = \\
&= A(1 - 2\epsilon + \theta)\tilde{h}^4 + B(\zeta - \theta) + \\
&\quad \tilde{h}^2 [B(1 - 2\epsilon + \theta) + D\zeta + A(\zeta - \theta) + (A + B + D)(\epsilon - \zeta)].
\end{aligned} \tag{A.55}$$

Podstawiając uzyskaną wielkość do układu równań (A.44) i wspomagając się przy rozwiązaniu programem Mathematica, otrzymuję, na podstawie (3.8), (A.19) oraz (A.42) i (A.41)

$$\begin{aligned}
\tilde{Q}(k, s) &= \mathbf{R} + \mathbf{S} = \\
&= \tilde{\psi}_1 \left[-(\epsilon - 1)\tilde{\psi}^3(\zeta - \theta) \left(\tilde{h}^2(\epsilon^2 - \zeta)(\epsilon^2(\zeta - \theta + 1) - \zeta) - (\epsilon - 1)^2\epsilon^2(\zeta - \theta) \right) + \right. \\
&\quad + (\epsilon - 1)^3\epsilon^2\tilde{\psi} \left(\left(\tilde{h}^2 - 1 \right) \epsilon^2 + \tilde{h}^2(-\theta) + \epsilon \right) - \tilde{h}(\epsilon - 1)^3\tilde{\psi}^2(\epsilon^2(\theta - 2\zeta) + \zeta^2) + \\
&\quad \tilde{h}(\epsilon - 1)^4\epsilon^2 - \tilde{h}\tilde{\psi}^4(\epsilon^2 - \zeta)(\zeta - \theta)(\epsilon^3 + \epsilon^2(-\zeta + \theta - 1) + \epsilon\zeta(\zeta - \theta - 1) + \zeta) + \\
&\quad \left. + \tilde{\psi}^5(\epsilon^2 - \zeta)^2(\zeta - \theta)^2 \right] \\
&\quad \left[\tilde{h}(\epsilon - 1)\tilde{\psi}^5(\epsilon^2 - \zeta)(\zeta - \theta)(\epsilon^2 + \zeta(\zeta - \theta - 1)) + \tilde{h}(\epsilon - 1)^3\epsilon^2\tilde{\psi}(2\epsilon - \theta - 1) \right. \\
&\quad - \tilde{h}(\epsilon - 1)\tilde{\psi}^3(\epsilon^4(3\zeta - 2\theta) + 2\epsilon^3(\zeta^2 - 2\zeta(\theta + 1) + \theta^2 + \theta) + \\
&\quad - \epsilon^2((\zeta + 1)\theta^2 - 2\zeta(\zeta + 2)\theta + \zeta(\zeta(\zeta + 4) - 2) + \theta) + 2\epsilon\zeta^2 + \zeta^2(\zeta - \theta - 1)) + \\
&\quad - \tilde{\psi}^6(\epsilon^2 - \zeta)^2(\zeta - \theta)^2 - (\epsilon - 1)^4\epsilon^2 + \epsilon\tilde{\psi}^4(3\epsilon^3 - \epsilon^2(2\zeta + 3) + \epsilon + \zeta^2)(\zeta - \theta)^2 + \\
&\quad \left. - (\epsilon - 1)^3\tilde{\psi}^2(\epsilon^3 - 2\epsilon^2\zeta + \zeta^2) \right]^{-1}.
\end{aligned} \tag{A.56}$$

Zauważmy, że w równaniu (A.56) zależność od argumentu s występuje jedynie poprzez $\tilde{\psi} = \tilde{\psi}(s)$ i $\tilde{\psi}_1 = \tilde{\psi}_1(s)$, natomiast zależność od argumentu k poprzez $\tilde{h} = \tilde{h}(k)$. Udało się w ten sposób znaleźć jawną postać ostrego propagatora dla procesu z pamięcią dwukrokową.

Dodatek B

Wpływ trendu dziennego na autokorelację prędkości

Analizując wewnątrzdzienną dynamikę stochastyczną ceny waloru giełdowego, napotykamy uniwersalny trend. W pracach [52–54] (i referencje tamże) zwany jest on trendem 'U' lub efektem lunchu. Niezależnie od analizowanej wielkości charakteryzującej aktywność graczy giełdowych (np. wariancja zmian ceny, wolumen obrotów, odwrotność średniego czasu międzytransakcyjnego), jej przebieg w ciągu dnia będzie przyjmował charakterystyczny kształt litery 'U' (właśnie to mamy na myśli mówiąc o uniwersalności). Oznacza to, większą aktywność giełdy na początku i końcu sesji, a mniejszą w środku dnia. Okres spowolnienia odpowiada porze lunchu (godz. 11-14), stąd nazwa zjawiska. Można zatem przypuszczać, że zmienność aktywności wewnątrz dnia powoduje niestacjonarność szeregu czasowego.

W interesującym okresie wewnątrz dnia (godz. 11-14), oprócz mniejszej aktywności giełdy, zaobserwować możemy również większą stabilność wspomnianych charakterystyk. Można zatem przypuszczać, że dane pochodzące ze wspomnianego okresu w środku dnia, z lepszym przybliżeniem niż w przypadku całego dnia, mogą być traktowane jako stacjonarne.

Możemy sprawdzić, na ile trend dzienny wpływa na wyniki analizy przeprowadzonej w niniejszej rozprawie, a w szczególności na funkcję autokorelacji prędkości. W tym celu, z danych opisanych w podrozdziale 2.1 wycinam jedynie rekordy danych odpowiadające transakcjom zawartym pomiędzy godziną 11.00 a 14.59 (gdyż, jak wynika z analizy przedstawionej w [53], dla GPW można, z dobrym przybliżeniem, przedłużyć okres lunchu do godziny 15.00). Zmniejsza to niestety liczbę dostępnych transakcji, co pokazano w tabeli B.1 (dla całości danych odpowiada jej tabela 2.1).

Przeprowadźmy, dla tak przygotowanych danych, powtórnie wszystkie analizy wykonane w podrozdziale 5.6. Wyniki przeprowadzonej estymacji parametru ϵ przedstawia tabela B.2 (odpowiadając tabeli 5.1). Wyniki dopasowania parametrów rozkładów czasów wyczekiwania przedstawia tabela B.3 (odpowiadająca tabeli 5.2). Natomiast rezultaty tego dopasowania przedstawione są na rysunku B.1 (odpowiada on rysunkowi 5.5). Jak widać, średnie czasy wyczekiwania $\langle t \rangle$ uległy wydłużeniu, co

jednak nie miało wpływu na jakość dopasowania do danych empirycznych.

Wyznaczone na podstawie powyższych danych parametry funkcji autokorelacji prędkości ceny danej wyrażeniem (5.32) (uzyskanej w ramach modelu z pamięcią jednokrokową i podwójnie eksponencjalnym rozkładem czasów wyczekiwania), przedstawia tabela B.4 (odpowiada ona tabeli 5.3). Porównanie empirycznej funkcji autokorelacji prędkości z przewidywaniami modeli jedno- i dwukrokowego przedstawia rysunek B.2. Porównując wykresy na rysunkach B.2 i 6.5 nie widać znaczącego wpływu trendu dziennego na zgodność przewidywań omawianych w rozprawie modeli z danymi empirycznymi (należy pamiętać, że mamy tutaj do czynienia z mniejszą statystyką).

Tablica B.1: Lista ośmiu badanych spółek wchodzących w skład indeksu WIG20 wraz z liczbą dni i transakcji. Uwzględniono jedynie transakcje zawarte w godzinach 11.00-14.59.

Skrót nazwy	Pełna nazwa	Dni	Zmian ceny
BRE	BRE Bank	2684	264058
GETIN	Getin Holding	2559	303427
KGHM	KGHM Polska Miedź	2686	1006272
PEKAO	Bank Pekao	2686	571779
PGNIG	Polskie Górnictwo Naftowe i Gazownictwo	1450	362238
PKNORLEN	PKN Orlen	2686	762184
PKOBP	PKO Bank Polski	1688	781213
TPSA	Telekomunikacja Polska	2686	967209

Tablica B.2: Wyniki estymacji parametru ϵ dla badanych spółek z WIG20. W drugiej kolumnie przedstawiono wyniki uzyskane metodą największej wiarygodności (maksymalizacji funkcjonału (5.24)), w trzeciej kolumnie – wyniki uzyskane metodą najmniejszych kwadratów (minimalizacji funkcjonału (5.25)), natomiast w czwartej – metodą korelacyjną (opartą o równanie (5.22)). Uwzględniono jedynie transakcje zawarte w godzinach 11.00-14.59.

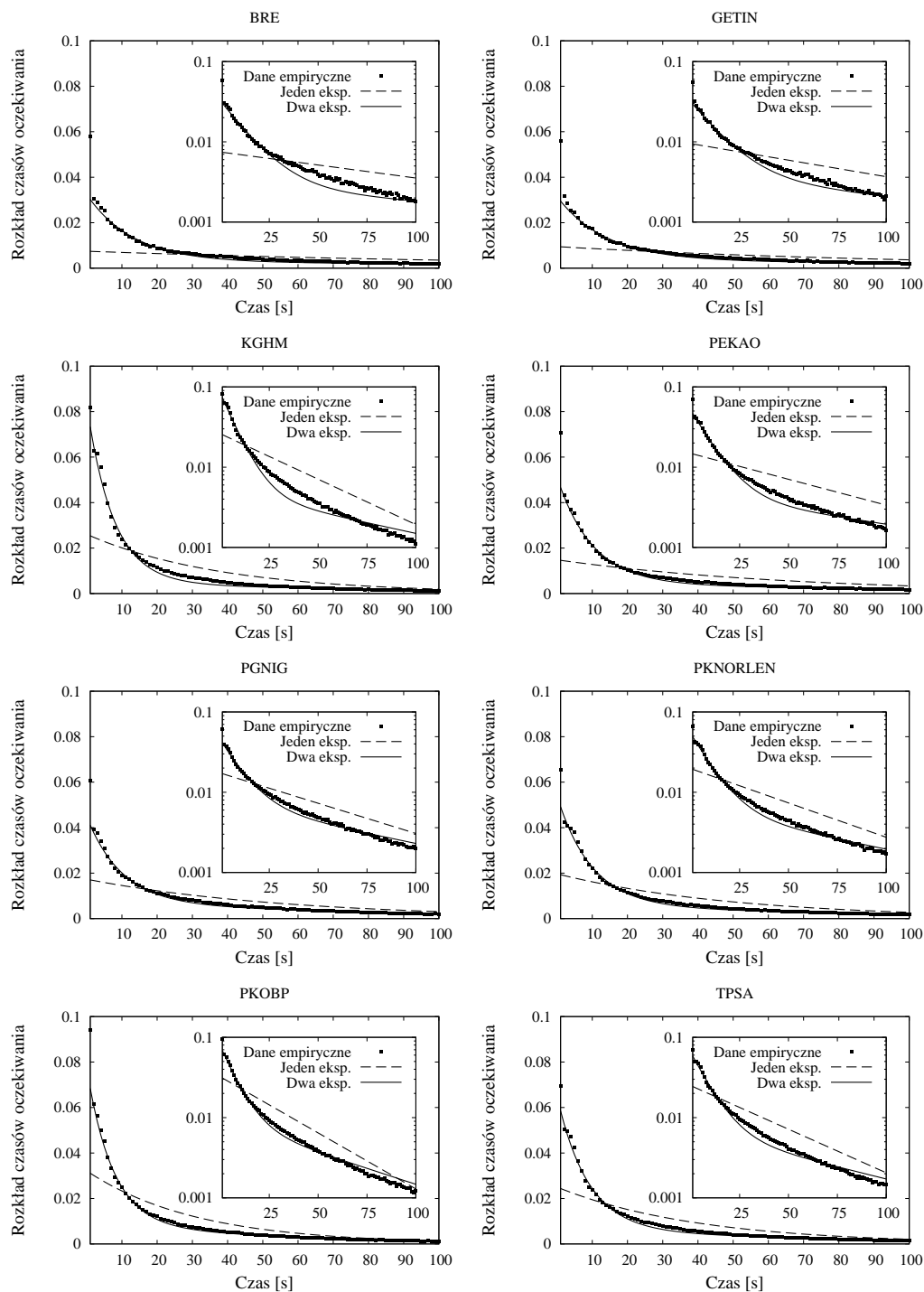
Skrót nazwy	m.n.wiaryg.	m.n.kwadratów	korelacja
BRE	0.224	0.227	0.330
GETIN	0.312	0.318	0.295
KGHM	0.256	0.233	0.416
PEKAO	0.206	0.203	0.357
PGNIG	0.296	0.257	0.426
PKNORLEN	0.300	0.280	0.377
PKOBP	0.320	0.308	0.396
TPSA	0.387	0.363	0.476

Tablica B.3: Wyniki estymacji parametrów rozkładów modelujących rozkład czasów oczekiwania dla badanych spółek z WIG20. W drugiej kolumnie zamieszczono średni czas między transakcjami używany w rozkładzie (5.26), w trzeciej i czwartej kolumnie znajdują się cząstkowe czasy relaksacji τ_1 i τ_2 rozkładu podwójnie eksponencjalnego (5.28), natomiast w kolumnie piątej waga w . Uwzględniono jedynie transakcje zawarte w godzinach 11.00-14.59.

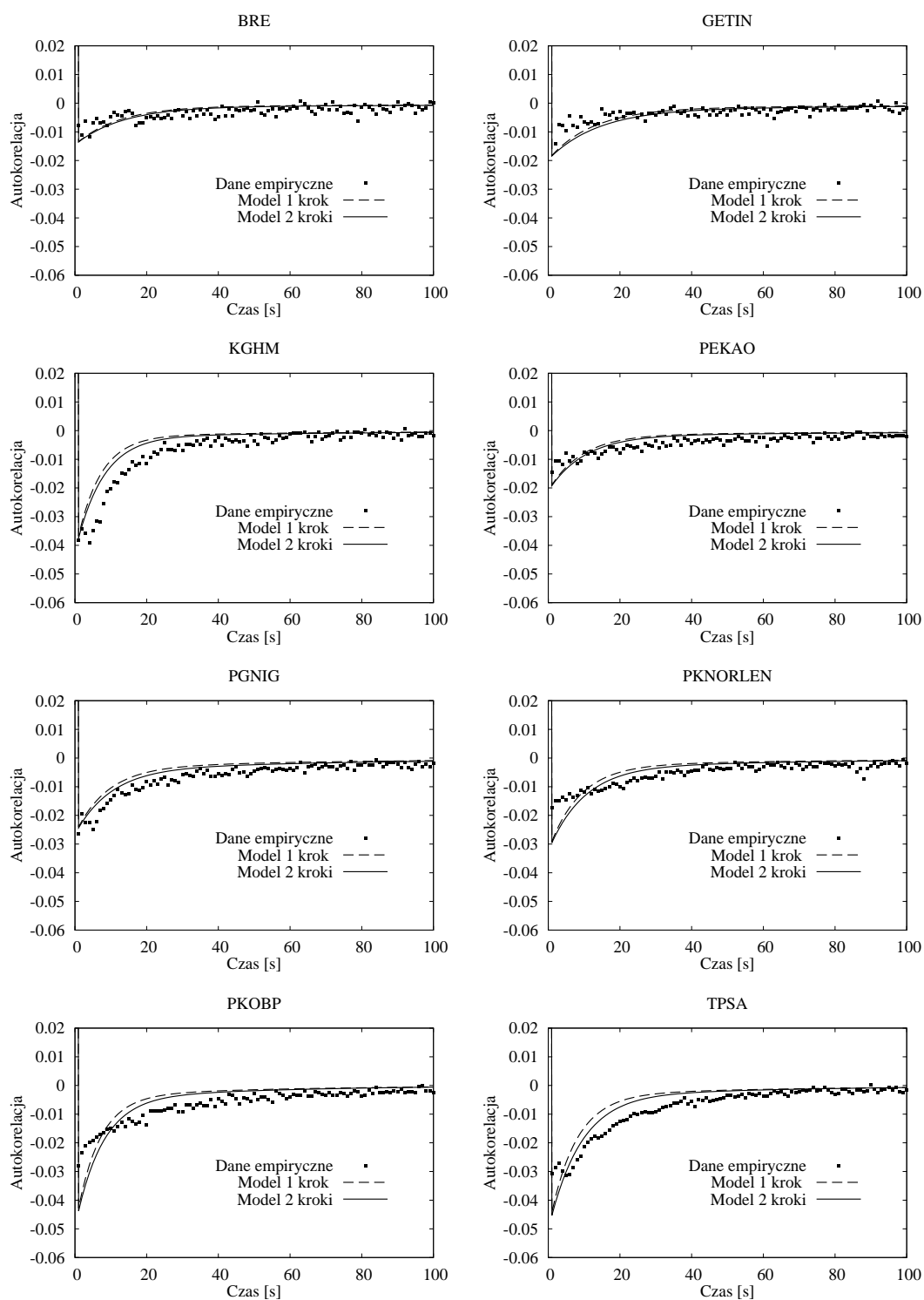
Skrót nazwy	$\langle t \rangle$	τ_1	τ_2	w
BRE	133.915	13.135	210.197	0.387
GETIN	105.337	13.725	161.170	0.379
KGHM	38.375	7.489	81.847	0.585
PEKAO	67.407	9.988	117.353	0.465
PGNIG	57.734	9.823	85.573	0.368
PKNORLEN	50.787	9.728	86.331	0.464
PKOBP	31.114	7.317	54.744	0.498
TPSA	40.003	8.681	71.768	0.504

Tablica B.4: Wartości parametrów ($1/\nu_1$, $1/\nu_2$, A_1 i A_2) teoretycznej funkcji autokorelacji prędkości opisanej równaniem (5.32) ośmiu badanych spółek wchodzących w skład indeksu WIG20. Parametry wyznaczono według wzorów (5.33), bazując na estymowanych parametrach, których wartości podano w tabelach B.2 (metoda największej wiarygodności) i B.3. W ostatniej kolumnie podano również wartość współczynnika μ_1^2/μ_2 występującego w stacjonaryzowanym modelu Montero i Masolivera (4.32). Uwzględniono jedynie transakcje zawarte w godzinach 11.00-14.59.

Skrót nazwy	$1/\nu_1$	$1/\nu_2$	A_1	A_2	μ_1^2/μ_2
BRE	12.079	186.753	0.02995	0.00244	0.197
GETIN	12.252	137.568	0.02844	0.00301	0.181
KGHM	6.508	75.018	0.07942	0.00373	0.225
PEKAO	9.109	106.737	0.04741	0.00373	0.210
PGNIG	8.839	73.360	0.03899	0.00582	0.310
PKNORLEN	8.520	75.812	0.04930	0.00460	0.233
PKOBP	6.292	48.224	0.07079	0.00647	0.221
TPSA	7.238	62.040	0.06034	0.00458	0.236



Rysunek B.1: Porównanie empirycznego histogramu czasów wyczekiwania (małe czarne kwadraty) z dopasowanymi przewidywaniami modeli opisującymi odpowiadający histogramowi rozkład. Linią przerywaną przedstawiony jest rozkład eksponencjalny wyrażony wzorem (5.26), natomiast linią ciągłą rozkład podwójnie eksponencjalny (5.28). Wewnątrz wykresów wykonanych w skali liniowej znajdują się ich zmniejszone odpowiedniki w skali pół-logarytmicznej. Uwzględniono jedynie transakcje zawarte w godzinach 11.00-14.59.

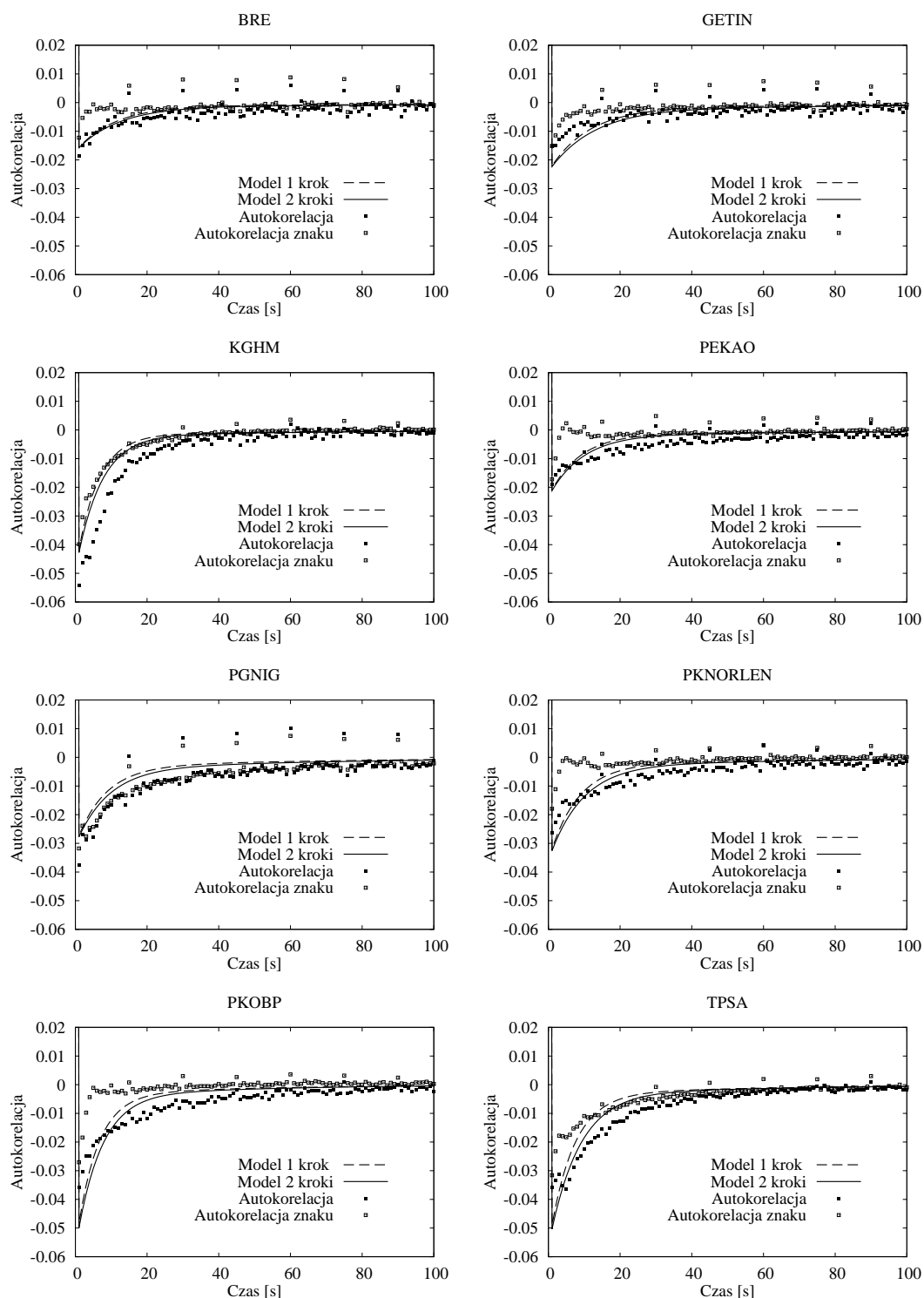


Rysunek B.2: Unormowana funkcja autokorelacji prędkości ceny: empiryczna (czarne kwadraty), wyrażona wzorem (5.32) dla modelu z pamięcią jednokorokową (linia przerywana) i wyrażona wzorem (6.5) dla modelu z pamięcią dwukorokową (linia ciągła). Oba modele teoretyczne opierają się na podwójnie eksponencjalnym rozkładzie czasów wyczekiwania danym wzorem (5.28). Uwzględniono jedynie transakcje zawarte w godzinach 11.00-14.59.

Dodatek C

Autokorelacja znaku prędkości ceny

Otrzymane w rozprawie funkcje autokorelacji prędkości ((5.20) i (6.35)) nie zależą od postaci rozkładu pojedynczej zmiany ceny. Jeżeli rozwinięte tutaj modele faktycznie poprawnie opisują dynamikę cen, to własność ta powinna być również obecna w danych empirycznych. Zatem, można przeprowadzić test polegający, na zamianie w danych empirycznych wszystkich zmian ceny na wartości $-1, 0, 1$, odpowiednio dla zmian ujemnych, zerowych i dodatnich. Z przewidywań modelu wynika, że funkcja autokorelacji prędkości tak otrzymanego procesu (czyli funkcja autokorelacji znaku prędkości) powinna pokrywać się z oryginalną funkcją autokorelacji prędkości (przedstawioną jako małe czarne kwadraty np. na rysunku 6.5). Wyniki tego porównania, przedstawiono na rysunku C.1 poprzez nałożenie na rysunek 6.5 unormowanej funkcji autokorelacji znaku prędkości (białe kwadraty). Faktycznie, pomijając wyniki dla spółek PEKAO, PKNORLEN i PKOBP, przebiegi porównywanych funkcji są do siebie zbliżone, chociaż widoczna jest tendencja szybszego zbiegania autokorelacji znaku prędkości do zera. W trzech wymienionych przypadkach musi istnieć, nieuwzględniony w modelach, dodatkowy mechanizm powodujący obserwowaną rozbieżność.



Rysunek C.1: Unormowana funkcja autokorelacji znaku prędkości ceny (białe kwadraty), funkcja autokorelacji prędkości ceny (czarne kwadraty), wyrażona wzorem (5.32) dla modelu z pamięcią jednokorokową (linia przerywana) i wyrażona wzorem (6.5) dla modelu z pamięcią dwukorokową (linia ciągła). Oba modele teoretyczne opierają się na podwójnie eksponencjalnym rozkładzie czasów wyczekiwania danym wzorem (5.28). Autokorelacji znaku prędkości ceny, w porównaniu do autokorelacji prędkości ceny, szybciej zbiega do zera.

Bibliografia

- [1] E. W. Montroll and G. H. Weiss. Random walks on lattices. II. *Journal of Mathematical Physics*, 6(2):167–181, 1965.
- [2] H. Scher and M. Lax. Stochastic transport in a disordered solid. I. Theory. *Phys. Rev. B*, 7(10):4491–4502, 1973.
- [3] E. J. Moore. On random walks and the ac conductivity of hopping systems. *Journal of Physics C: Solid State Physics*, 7(2):339, 1974.
- [4] J. K. E. Tunaley. Theory of ac conductivity based on random walks. *Phys. Rev. Lett.*, 33(17):1037–1039, 1974.
- [5] J. K. E. Tunaley. Asymptotic solutions of the continuous-time random walk model of diffusion. *Journal of Statistical Physics*, 11:397–408, 1974.
- [6] C. Monthus and J.-P. Bouchaud. Models of traps and glass phenomenology. *Journal of Physics A: Mathematical and General*, 29(14):3847, 1996.
- [7] E. Barkai and Y.-Ch. Cheng. Aging continuous time random walks. *The Journal of Chemical Physics*, 118(14):6167–6178, 2003.
- [8] W. Dieterich and P. Maass. Constant dielectric loss in disordered ionic conductors: Theoretical aspects. *Solid State Ionics*, 180(6-8):446 – 450, 2009. The 2nd International Conference on Physics of Solid State Ionics (2nd ICPSSI).
- [9] R. Klages, G. Radons, and I. M. Sokolov (eds.). Anomalous transport: Foundations and applications. *Journal of Statistical Physics*, 135:389–391, 2009.
- [10] F. Boano, A.-I. Packman, A. Cortis, R. Revelli, and L. Ridolfi. A continuous time random walk approach to the stream transport of solutes. *Water Resources Research*, 431:W10425, 2007.
- [11] A. Helmstetter and D. Sornette. Diffusion of epicenters of earthquake aftershocks, Omori’s law, and generalized continuous-time random walk models. *Phys. Rev. E*, 66(6):061104, 2002.
- [12] W. Feller. *Wstęp do rachunku prawdopodobieństwa. Tom 1*. Wydawnictwo Naukowe PWN, Warszawa, 2006.

- [13] V. S. Barbu and N. Limnios. *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications: Their Use in Reliability and DNA Analysis*. Lecture Notes in Statistics. Springer, 2008.
- [14] R. Hilfer. *Applications of fractional calculus in physics*. World Scientific, 2000.
- [15] O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer Series in Statistics, Springer Science+Business Media Inc., New York, 2005.
- [16] E. Scalas, R. Gorenflo, and F. Mainardi. Fractional calculus and continuous-time finance. *Physica A: Statistical Mechanics and its Applications*, 284(1-4):376 – 384, 2000.
- [17] F. Mainardi, M. Raberto, R. Gorenflo, and E. Scalas. Fractional calculus and continuous-time finance ii: the waiting-time distribution. *Physica A: Statistical Mechanics and its Applications*, 287(3-4):468 – 481, 2000.
- [18] M. Raberto, E. Scalas, and F. Mainardi. Waiting-times and returns in high-frequency financial data: an empirical study. *Physica A: Statistical Mechanics and its Applications*, 314(1-4):749 – 755, 2002.
- [19] E. Scalas, R. Gorenflo, and F. Mainardi. Uncoupled continuous-time random walks: Solution and limiting behavior of the master equation. *Phys. Rev. E*, 69(1):011107, 2004.
- [20] E. Scalas. The application of continuous-time random walks in finance and economics. *Physica A: Statistical Mechanics and its Applications*, 362(2):225 – 239, 2006.
- [21] R. Kutner and F. Świtłała. Stochastic simulations of time series within Weierstrass - Mandelbrot walks. *Quantitative Finance*, 3(3):201–211, 2003.
- [22] J. Masoliver, M. Montero, and G. H. Weiss. Continuous-time random-walk model for financial distributions. *Phys. Rev. E*, 67(2):021112, 2003.
- [23] P. Repetowicz and P. Richmond. Modeling share price evolution as a continuous time random walk (ctrw) with non-independent price changes and waiting times. *Physica A: Statistical Mechanics and its Applications*, 344(1-2):108 – 111, 2004. Applications of Physics in Financial Analysis 4 (APFA4).
- [24] J. Masoliver, M. Montero, J. Perelló, and G. H. Weiss. The continuous time random walk formalism in financial markets. *Journal of Economic Behavior & Organization*, 61(4):577 – 598, 2006.

- [25] J. Masoliver, M. Montero, J. Perello, and G. H. Weiss. The CTRW in finance: Direct and inverse problems with some generalizations and extensions. *Physica A: Statistical Mechanics and its Applications*, 379(1):151 – 167, 2007.
- [26] R. S. Tsay. *Analysis of financial time series*. Wiley Series in Probability and Statistics. Wiley, 2005.
- [27] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceed. IEEE*, 77(2):257–286, 1989.
- [28] R. Bhar and S. Hamori. *Hidden Markov models: applications to financial economics in Advanced studies in theoretical and applied econometrics*. Kluwer Academic Publishers, 2004.
- [29] S. L. Scott. Bayesian methods for hidden markov models. *Journal of the American Statistical Association*, 97(457):337–351, 2002.
- [30] P.W. Atkins. *Chemia fizyczna*. Wydawnictwo Naukowe PWN, 2001.
- [31] P. A. Samuelson. *Foundations of Economic Analysis*. Harvard University Press, 1983.
- [32] R. N. Mantegna and H. E. Stanley. *Ekonofizyka. Wprowadzenie*. Wydawnictwo Naukowe PWN, Warszawa, 2001.
- [33] J. W. Haus and K. W. Kehr. Random walk model with correlated jumps: Self-correlation function and frequency-dependent diffusion coefficient. *Journal of Physics and Chemistry of Solids*, 40(12):1019 – 1025, 1979.
- [34] K. W. Kehr, R. Kutner, and K. Binder. Diffusion in concentrated lattice gases. self-diffusion of noninteracting particles in three-dimensional lattices. *Phys. Rev. B*, 23(10):4931–4945, 1981.
- [35] R. Kutner. Correlated hopping in honeycomb lattice: tracer diffusion coefficient at arbitrary lattice gas concentration. *Journal of Physics C: Solid State Physics*, 18(34):6323, 1985.
- [36] J. W. Haus and K. W. Kehr. Diffusion in regular and disordered lattices. *Physics Reports*, 150(5-6):263 – 406, 1987.
- [37] J. R. Manning. *Diffusion Kinetics for Atoms in Crystals*. D. van Nostrand Comp, Inc., Princeton, 1968.
- [38] K. Funke. Jump relaxation in solid electrolytes. *Progress in solid state chemistry*, 22(2):111–195, 1993.
- [39] Ch. Kittel. *Introduction to Solid State Physics*. J. Wiley & Sons, Inc., New York, 1996.

- [40] C.P. Flynn. *Point defects and diffusion*. International series of monographs on physics. Clarendon Press, Oxford, 1972.
- [41] C. P. Flynn and A. M. Stoneham. Quantum theory of diffusion with application to light interstitials in metals. *Phys. Rev. B*, 1:3966–3978, 1970.
- [42] M. Montero, J. Perelló, J. Masoliver, F. Lillo, S. Miccichè, and R. N. Mantegna. Scaling and data collapse for the mean exit time of asset prices. *Phys. Rev. E*, 72(5):056101, 2005.
- [43] M. Montero and J. Masoliver. Nonindependent continuous-time random walks. *Phys. Rev. E*, 76(6):061115, 2007.
- [44] J. Villarroel and M. Montero. On properties of continuous-time random walks with non-poissonian jump-times. *Chaos, Solitons and Fractals*, 42(1):128 – 137, 2009.
- [45] T. Gubiec and R. Kutner. Share price evolution as stationary, dependent continuous-time random walk. *Acta Phys. Pol. A*, 117(4):669–672, 2010.
- [46] D. Fudenberg and J. Tirole. *Game Theory*. The MIT Press, Cambridge in Massachusetts, 1991.
- [47] D. P. Friedman and J. Rust (Eds.). *The Double Auction Market. Institutions, Theories, AND Evidence*. Proceedings of Santa Fe Institute Studies in the Sciences of Complexity.
- [48] S. Gjerstad and J. Dickhaut. Price formation in double auctions,. *Games and Economic Behavior*, 22(1):1 – 29, 1998.
- [49] R. Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39(4):1127–1139, 1984.
- [50] T. Gubiec and R. Kutner. Backward jump continuous-time random walk: An application to market trading. *Phys. Rev. E*, 82(4):046119, 2010.
- [51] T. Gubiec and R. Kutner. Continuous time random walk with memory over two successive jumps. An application to high-frequency stock market dynamics. *W przygotowaniu do druku*, 2011.
- [52] D. Sornette. Dragon-kings, black swans and the prediction of crises. *International Journal of Terraspace Science and Engineering*, 2(1):1–18, 2009.
- [53] A. Kasprzak. *Analiza wielofraktalnej struktury czasów międzytransakcyjnych za pomocą modelu błędzenia losowego w czasie ciągłym*. Rozprawa doktorska, promotor: prof. dr hab. R. Kutner, Wydział Fizyki, Uniwersytet Warszawski, 2010.

- [54] L. A. N. Amaral, P. Cizeau, P. Gopikrishnan, Y. Liu, M. Meyer, C.-K. Peng, and H. E. Stanley. Econophysics: can statistical physics contribute to the science of economics? *Computer Physics Communications*, 121:145–152, 1999.
- [55] <http://bossa.pl>.
- [56] <http://www.equis.com/>.
- [57] *Systemy informatyczne GPW*, <http://www.gpw.pl>.
- [58] N. G. van Kampen. *Procesy stochastyczne w fizyce i chemii*. Państwowe Wydawnictwo Naukowe, Warszawa, 1990.
- [59] R. Kutner and F. Świtala. Remarks on the possible universal mechanism of the non-linear long-term autocorrelations in financial time-series. *Physica A: Statistical Mechanics and its Applications*, 344(1-2):244 – 251, 2004.
- [60] R. Kutner. Hierarchical spatio-temporal coupling in fractional wanderings. (i) continuous-time weierstrass flights. *Physica A: Statistical and Theoretical Physics*, 264(1-2):84 – 106, 1999.
- [61] G. Bel and E. Barkai. Weak ergodicity breaking in the continuous-time random walk. *Phys. Rev. Lett.*, 94(24):240602, 2005.
- [62] R. Kubo, M. Toda, and N. Hashitsume. *Fizyka statystyczna II. Mechanika statystyczna stanów nierównowagowych*. Wydawnictwo Naukowe PWN, Warszawa, 1991.
- [63] R. F. Engle and J. R. Russell. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66:1127–1162, 1998.
- [64] F. Lillo. Informacja prywatna.
- [65] R. Nowak. *Statystyka dla fizyków*. Wydawnictwo Naukowe PWN, Warszawa, 2002.
- [66] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.
- [67] Alicja Zalewska. *Analiza autokorelacji krótkookresowych w finansowych szeregach czasowych*. Praca licencjacka, opiekun: prof. dr hab. R. Kutner, Wydział Fizyki, Uniwersytet Warszawski, 2010.
- [68] F. Mainardi, M. Raberto, R. Gorenflo, and E. Scalas. Fractional calculus and continuous-time finance ii: the waiting-time distribution. *Physica A: Statistical Mechanics and its Applications*, 287(3-4):468 – 481, 2000.

- [69] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.
- [70] T. Preis. Econophysics - complex correlations and trend switchings in financial time series. *Eur. Phys. J. Special Topics*, 194:5–86, 2011.