



# SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

## Visibility Graphs för tidsserier: Matematiska egenskaper och tillämpningar

av

Melvin Segerman

2025 - No K1

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM



# Visibility Graphs för tidsserier: Matematiska egenskaper och tillämpningar

Melvin Segerman

---

Självständigt arbete i  
matematik 15 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2025



## Abstract

This paper presents the mathematical foundations of visibility-based graph representations of time series, focusing primarily on Natural Visibility Graphs and Horizontal Visibility Graphs, while also introducing Invisibility Graphs as a complementary construction. Key theoretical properties are formulated and proven, including invariance under strictly monotone transformations and the exact degree distribution of Horizontal Visibility Graphs for i.i.d. stochastic processes. These results are then used to contrast theoretical behaviour of periodic, stochastic and chaotic time series, with empirical focus on stochastic, chaotic and financial data.

Building on this theory, the paper develops local visibility- and invisibility-based indicators designed to quantify concave and convex price dynamics within rolling windows. The methodology is applied to daily SPY (SP 500 ETF) price data and compared to simulated reference series generated from white noise and the logistic map in a chaotic regime. The empirical findings show that SPY exhibits graph characteristics that differ from both purely stochastic and low-dimensional chaotic systems, with convex indicators in particular highlighting periods of rapid market stress.

The results show that the visibility graphs grasp structural asymmetries in financial time series while being a descriptive tool of trend shape analysis. Possible extensions involve assessing predictability and embedding the proposed indicators within trading or risk-management schemes.

## Sammanfattning

Detta arbete behandlar de matematiska grunderna för visibility-baserade grafrepresentationer av tidsserier, med fokus på Natural Visibility Graphs (NVG) och Horizontal Visibility Graphs (HVG), samt introducerar Invisibility Graphs (IVG) som ett kompletterande begrepp. Centrala teoretiska egenskaper formuleras och bevisas, däribland invarians under strikt monotona transformationer och den exakta gradfördelningen för HVG konstruerad från i.i.d. slumpserier. Dessa resultat används därefter för att kontrastera teoretiskt beteende hos periodiska, stokastiska och kaotiska tidsserier, medan den empiriska analysen fokuserar på stokastiska, kaotiska och finansiella data.

Med utgångspunkt i denna teori utvecklas lokala indikatorer baserade på visibility- och invisibility-grafer, konstruerade för att kvantifiera konkava respektive konvexa prisrörelser inom rullande tidsfönster. Metoden tillämpas på dagliga prisdata för SPY (S&P 500 ETF) och jämförs med simulerade referensserier genererade från vitt brus och den logistiska avbildningen i en kaotisk regim. De empiriska resultaten visar att SPY uppvisar grafegenskaper som skiljer sig från både rent stokastiska och lågdimensionella kaotiska system, där särskilt de konvexa indikatorerna tydligt fångar perioder av snabb marknadsstress.

Sammantaget visar resultaten att visibility-baserade grafer kan fånga strukturella asymmetrier i finansiella tidsserier samtidigt som de utgör ett deskriptivt verktyg för analys av trendform. Möjliga fortsättningar innefattar att undersöka indikatorernas prognosförmåga samt att integrera dem i handelsstrategier eller riskhanteringsramverk.

## Innehållsförteckning

# Innehåll

<b>1</b>	<b>Introduktion</b>	<b>6</b>
<b>2</b>	<b>Definitioner</b>	<b>8</b>
2.1	Tidsserier . . . . .	8
2.2	Deterministiska och stokastiska tidsserier . . . . .	9
2.3	Grundläggande graftermer . . . . .	10
2.4	Visibility Graphs . . . . .	13
2.4.1	Natural Visibility Graphs . . . . .	13
2.4.2	Horizontal Visibility Graphs . . . . .	14
2.4.3	Invisibility Graphs . . . . .	15
2.4.4	Algoritmisk representation av Visibility Graphs . . . . .	15
2.4.5	Sambandet mellan NVG och HVG . . . . .	17
2.5	Exempel på deterministiskt kaos: den logistiska avbildningen . . . . .	19
<b>3</b>	<b>Teori för Visibility Graphs</b>	<b>20</b>
3.1	Invariants under strikt växande transformationer för HVG . . . . .	20
3.2	Gradfördelning i HVG . . . . .	21
3.3	Visibility graphs för periodiska, stokastiska och kaotiska tidsserier . . . . .	24
<b>4</b>	<b>Metod</b>	<b>25</b>
4.1	Utgångspunkt: tidsserie . . . . .	25
4.2	Konstruktion av synlighetsgrafer . . . . .	25
4.3	Begränsat siktfönster . . . . .	25
4.4	Konkava och konvexa indikatorer . . . . .	25
4.5	Nätverksmått . . . . .	28
4.6	Simulerade referensserier: vitt brus och deterministiskt kaos . . . . .	28
4.7	Rullande analys . . . . .	29
<b>5</b>	<b>Finansiella Tillämpningar</b>	<b>29</b>
5.1	Data och empirisk uppställning . . . . .	30
5.2	SPY . . . . .	31
5.3	Konkava och konvexa årsindikatorer för SPY . . . . .	32

5.4	Jämförelse med stokastiska och kaotiska referensserier . . . . .	34
5.5	Diskussion av indikatorerna . . . . .	37
	<b>Referenser</b>	<b>39</b>
	<b>Generative AI disclosure</b>	<b>39</b>

# 1 Introduktion

## Bakgrund

Tidsserier är sekvenser av observationer ordnade i tiden och används inom en mängd olika områden, såsom finans, klimatforskning och fysik. Traditionell tidsserieanalys bygger ofta på statistiska verktyg som autokorrelation, spektralanalys eller regressionsmodeller. Dessa metoder är kraftfulla men bygger vanligen på antaganden om stationaritet eller linjära samband, vilket inte alltid är uppfyllt i verklig data.

Under de senaste två decennierna har så kallade *visibility graphs* introducerats som ett alternativt, strukturellt angreppssätt. En visibility graph omvandlar en tidsserie till en graf genom att för varje observation skapa en nod och koppla samman noder med kanter baserat på ett geometriskt synlighetskriterium. Detta gör det möjligt att studera tidsseriens form genom graftermer såsom gradfördelning, klustringskoefficienter och symmetrier.

Metoden har studerats i flera centrala arbeten. Luque et al. (2010) härledde exakta uttryck för gradfördelningen i Horizontal Visibility Graphs (HVG) för i.i.d. slumpserier och visade att dessa grafer uppvisar universella egenskaper som inte beror på den underliggande tätheten. Lacasa och Flanagan (2015) använde visibility graphs för att analysera tidsreversibilitet och visade hur grafernas struktur kan avslöja icke-stationära processer. Inom finans har Liu och Chen (2020) visat att visibility-baserade grafer kan användas för att kvantifiera asymmetrier i upp- och nedgångar i aktiepriser.

Mot denna bakgrund framträder visibility graphs som ett verktyg som både är matematiskt intressant och praktiskt användbart i empiriska sammanhang.

## Syfte

Syftet med detta arbete är tvådelat. För det första undersöks matematiska egenskaper hos Natural Visibility Graphs (NVG) och Horizontal Visibility Graphs (HVG), däribland invarians under transformationer och gradfördelningar. För det andra används visibility- och invisibility-baserade grafer för att konstruera indikatorer som

kvantifierar konkava och konvexa prisrörelser i finansiella tidsserier.

## **Arbetets struktur**

Arbetet är organiserat på följande sätt. I Kapitel 2 introduceras de matematiska grundbegrepp som uppsatsen bygger på, däribland definitioner av tidsserier, stokastisk respektive deterministisk dynamik samt graftermer och konstruktionen av NVG, HVG och IVG. Kapitel 3 behandlar teoretiska egenskaper hos dessa grafer, inklusive formella bevis av invariansresultat, kända resultat om gradfördelningar och en kort introduktion till deterministiskt kaos sedd genom visibility graphs. Kapitel 4 beskriver den metod som används för att analysera både simulerade och empiriska tidsserier, med särskilt fokus på indikatorer baserade på konkava och konvexa synlighetskriterier, samt en jämförelse mellan stokastiska och deterministiska referensmodeller. I Kapitel 5 tillämpas metoden på prisdata för SPY (S&P 500), där indikatorernas utveckling studeras över tid och relateras till större marknadshändelser.

## 2 Definitioner

I grund och botten är Visibility Graphs ett verktyg för att analysera data i tidsserier. I en Visibility Graph representeras varje observation i en tidsserie som en nod i en graf, där kopplingarna mellan noderna bestäms utifrån ett geometriskt synlighetskriterium. På så sätt kan man överföra tidsseriens struktur till ett nätverk och därefter analysera den med hjälp av metoder från grafteori.

Syftet med detta tillvägagångssätt är att fånga de underliggande dynamiska egenskaperna i en tidsserie. Det har till exempel visats att periodiska, slumpmässiga och fraktala tidsserier ger upphov till olika typer av grafer, från regelbundna till exponentiella och skalfria nätverk.

I detta arbete fokuserar vi på de två vanligaste typerna av Visibility Graphs: Natural Visibility Graphs och Horizontal Visibility Graphs, som beskrivs i följande avsnitt.

### 2.1 Tidsserier

För att kunna bygga upp begreppet Visibility Graph behöver vi först definiera vad vi menar med en tidsserie och hur dessa observationer kan beskrivas matematiskt.

En tidsserie är en sekvens datapunkter som är observerade över en given tid. Tidsserier används för att beskriva och analysera utvecklingen av en viss variabel över tid, exempelvis aktiepriser, BNP-tillväxt eller temperaturer.

#### Definition 2.1. Tidsserie

En tidsserie är en ordnad mängd observationer

$$\{x_t\}_{t=1}^T = (x_1, x_2, \dots, x_T),$$

där  $x_t \in \mathbb{R}$  betecknar värdet på en variabel vid tidpunkt  $t$ , och  $T$  är det totala antalet observationer.

Om observationerna sker vid diskreta tidpunkter kallas serien diskret, medan en

tidsserie som är kontinuerligt definierad i tiden betecknas

$$x(t), \quad t \in [0, T]$$

och kallas då en kontinuerlig tidsserie.

Dessa distinktioner är viktiga eftersom de påverkar hur synlighetskriterierna definieras i grafen, särskilt vid beräkning av kanter mellan datapunkter.

### Notation 2.1

I detta arbete kommer  $x_t$  att beteckna observationer vid tidpunkt  $t$ , och väntevärdet och variansen vid tidpunkt  $t$  betecknas med

$$\mathbb{E}[x_t] \quad \text{och} \quad \text{Var}(x_t).$$

## 2.2 Deterministiska och stokastiska tidsserier

I den enklaste formen kan en tidsserie beskrivas antingen som resultatet av en stokastisk process eller av ett deterministiskt dynamiskt system.

### Definition 2.2. Stokastisk tidsserie

En stokastisk tidsserie  $\{X_t\}_{t=1}^T$  är en följd av slumpvariabler definierade på ett gemensamt sannolikhetsrum. Egenskaper som väntevärde, varians och autokorrelation bestäms då av processens sannolikhetslagar. Exempel är vitt brus, random walk och ARMA-processer.

### Definition 2.3. Deterministisk tidsserie

En deterministisk tidsserie är en följd av värden  $x_t$  där nästa värde helt bestäms av en funktion av det föregående:

$$x_{t+1} = F(x_t),$$

där  $F$  är en (eventuellt icke-linjär) funktion. Givet ett startvärde  $x_0$  är hela banan  $\{x_t\}$  entydigt bestämd utan någon slump.

En viktig punkt är att deterministiska system kan ge upphov till tidsserier som ser slumpmässiga ut, trots att de styrs av en enkel regel. Detta kallas *deterministiskt kaos*. En kaotisk bana uppvisar bland annat:

- *känslighet för begynnelsevillkor* (small skillnad i  $x_0$  ger snabbt helt olika banor),
- *aperiodicitet* (ingen exakt upprepning även på lång sikt),
- en rik geometrisk struktur i fastrummet (t.ex. fraktala attraktorer).

Ur ett rent observationsperspektiv är det därför inte självklart om en till synes komplex tidsserie är styrd av stokastiskt brus eller deterministiskt kaos. I denna uppsats undersöks om visibility graphs kan bidra med strukturmått som hjälper till att skilja dessa fall åt.

#### **Definition 2.4. Vitt brus**

En (diskret) vitbrusprocess  $\{\varepsilon_t\}_{t=1}^T$  är en följd av oberoende och identiskt fördelade slumpvariabler med

$$\mathbb{E}[\varepsilon_t] = 0, \quad \text{Var}(\varepsilon_t) = \sigma^2 < \infty,$$

och utan autokorrelation,

$$\text{Cov}(\varepsilon_t, \varepsilon_s) = 0 \quad \text{för } t \neq s.$$

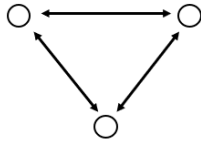
I detta arbete används som referensmodell normalfördelat vitt brus med  $\varepsilon_t \sim \mathcal{N}(0, 1)$ .

### **2.3 Grundläggande graftermer**

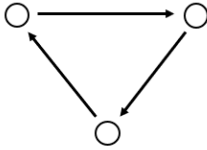
När vi nu definierat tidsseriernas matematiska struktur kan vi gå vidare till den grafteoretiska grunden som Visibility Graphs bygger på. En graf är en matematisk struktur som används för att beskriva relationer mellan objekt.

#### **Definition 2.5. Oriktad graf**

En oriktad graf består av noder och kanter som förbinder dem, där kanterna saknar riktning (tillskillnad från en riktad graf).



Figur 1: Illustration av oriktad graf.



Figur 2: Illustration av riktad graf.

Denna typ av graf definieras som

$$G = (V, E),$$

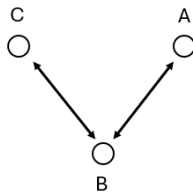
där  $V = \{v_1, v_2, \dots, v_n\}$  är mängden av *noder* (även kallat hörn) och  $E \subseteq \{\{u, v\} \mid u, v \in V, u \neq v\}$  är mängden av *kanter*. En kant  $\{u, v\} \in E$  betyder att noderna  $u$  och  $v$  är sammanlänkade.

### Definition 2.6. Incident

En kant är incident till en nod om den ansluter till, eller går ut från, den noden.

### Exempel

Anta att du har en enkel graf med noderna  $V = \{A, B, C\}$  och kanterna  $E = \{\{A, B\}, \{B, C\}\}$ .



Figur 3: Graf med noderna  $V = \{A, B, C\}$  och kanterna  $E = \{\{A, B\}, \{B, C\}\}$ .

Då gäller:

- Noden A har en incident kant:  $\{A, B\}$
- Noden B har två incidenta kanter:  $\{A, B\}$  och  $\{B, C\}$
- Noden C har en incident kant:  $\{B, C\}$

### Definition 2.7. Gradfördelning

Graden av en nod  $v \in V$ , betecknad  $k(v)$ , är antalet kanter som är incidenta till  $v$ :

$$k(v) = |\{u \in V \mid \{u, v\} \in E\}|.$$

Alltså är gradfördelningen i figur 3:

$$k(A) = 1, k(B) = 2, k(C) = 1$$

Gradfördelningen,  $P(k)$ , beskriver sannolikheten att en slumpmässig vald nod har grad  $k$ :

$$P(k) = \frac{N_k}{N},$$

där  $N_k$  är antalet noder med grad  $k$  och  $N = |V|$  är totala antalet noder.

### Definition 2.8. Kortaste vägen

Den kortaste vägen mellan två noder  $u, v \in V$  betecknas  $d(u, v)$  och motsvarar det minsta antalet kanter som måste passeras för att ta sig från  $u$  till  $v$ .

Den genomsnittliga längden av de kortaste vägarna i hela grafen, medellängden, definieras som

$$L = \frac{1}{N(N-1)} \sum_{u \neq v} d(u, v).$$

Diametern av grafen,  $\text{diam}(G)$ , är den största kortaste vägen mellan två noder:

$$\text{diam}(G) = \max_{u, v \in V} d(u, v).$$

## Definition 2.9. Klustringskoefficienten

Den lokala klustringskoefficienten för en nod  $v$ , betecknad  $C_i$ , mäter hur sammankopplade  $v$ :s grannar är och ges av

$$C_i = \frac{2e_i}{k_i(k_i - 1)}.$$

där  $e_i$  är antalet kanter mellan noderna i mängden  $v$ :s grannar.

Den globala klustringskoefficienten definieras som medelvärdet över alla noder

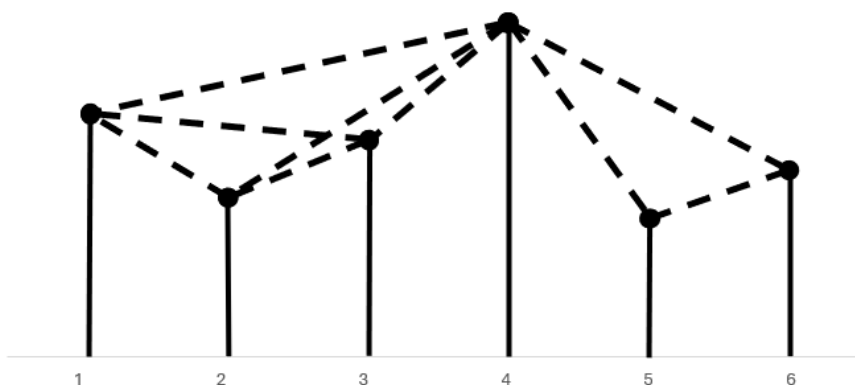
$$C = \frac{1}{N} \sum_{i=1}^N C_i.$$

## 2.4 Visibility Graphs

Med denna teoretiska grund kan vi nu definiera själva huvudobjektet för arbetet, Visibility Graphs, och se hur dessa konstrueras från en tidsserie.

### 2.4.1 Natural Visibility Graphs

En Natural Visibility Graph (NVG) är när varje observation i en tidsserie representeras som en nod i en graf. Varje nod i grafen motsvarar en tidsstämpel  $t$  i tidsserien, så noderna är ordnade i serie. Två noder är sammankopplade om det finns en siktlinje mellan de motsvarande datastaplar som inte avbryts.



Figur 4: Illustration av en Natural Visibility Graph (NVG).

## Definition 2.10. Natural Visibility Graph

Givet en tidserie  $\{(t_i, y_i)\}_{i=1}^N$  där  $t_i$  är tidpunkten och  $y_i$  är värdet, konstruerar man en NVG så här:

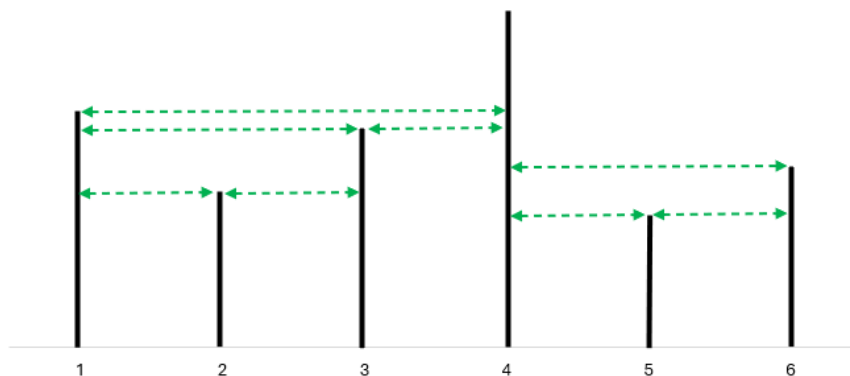
- Varje datapunkt  $(t_i, y_i)$  representerar en nod i grafen.
- Två noder  $i$  och  $j$  (då  $i < j$ ) är anslutna med en kant om man kan dra en rak linje mellan punkterna  $(t_i, y_i)$  och  $(t_j, y_j)$  utan att linjen korsas av någon mellanliggande datapunkt.

$(t_i, y_i)$  och  $(t_j, y_j)$  är kopplade om

$$y_k < y_i + \frac{y_j - y_i}{t_j - t_i}(t_k - t_i) \quad \forall t_k \in (t_i, t_j).$$

### 2.4.2 Horizontal Visibility Graphs

En Horizontal Visibility Graph (HVG) är en förenklad version av NVG, där endast horisontella siktlinjer tillåts. Linjerna behöver i denna metod inte nå stapels topp, utan man kan "se" från en stapel till en annan så länge ingen mellanliggande stapel är lika hög eller högre.



Figur 5: Illustration av en Horizontal Visibility Graph (HVG).

### Definition 2.11. Horizontal Visibility Graph

I HVG är två noder  $i$  och  $j$  (med  $i < j$ ) kopplade om det inte finns någon datapunkt mellan dem som har ett högre eller lika högt värde.

$(t_i, y_i)$  och  $(t_j, y_j)$  är kopplade om

$$y_k < \min(y_i, y_j) \quad \forall t_k \in (t_i, t_j).$$

### 2.4.3 Invisibility Graphs

Utöver Visibility Graphs kan man definiera en *Invisibility Graph* (IVG), som bygger på samma geometriska idé men med ett omvänt synlighetskriterium. Givet en tidsserie  $\{(t_i, x_i)\}_{i=1}^T$  med  $t_1 < \dots < t_T$  betraktar vi som tidigare vertikala staplar vid varje tidpunkt  $t_i$  med höjd  $x_i$ .

### Definition 2.12. Invisibility Graph

Två datapunkter  $(t_a, x_a)$  och  $(t_b, x_b)$  med  $a < b$  är förbundna med en kant i IVG om den räta linjen mellan dem *skärs* av minst en mellanliggande datapunkt. Formellt krävs att det finns ett index  $c$  med  $a < c < b$  sådant att

$$x_c > x_b + \frac{x_a - x_b}{t_b - t_a} (t_b - t_c).$$

Om detta villkor är uppfyllt säger vi att  $(t_a, x_a)$  och  $(t_b, x_b)$  är *osynliga* för varandra i den vanliga visibility-meningen, och knyter därför samman dem i en Invisibility Graph.

I analogi med NVG/HVG kan man därefter införa riktade eller tidsfönsterbegränsade IVG-varianter, där man t.ex. endast tillåter länkar bakåt i tiden ( $t_a < t_b$ ) och begränsar avståndet  $|a - b|$  till ett givet siktfönster  $S$ .

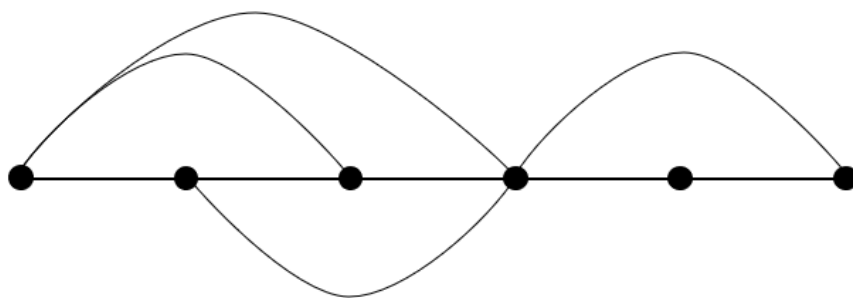
### 2.4.4 Algoritmisk representation av Visibility Graphs

För att förstå hur en Visibility Graph byggs upp kan man betrakta två nära relaterade representationer: den geometriska konstruktionen och den algoritmiska grafrepresentationen.

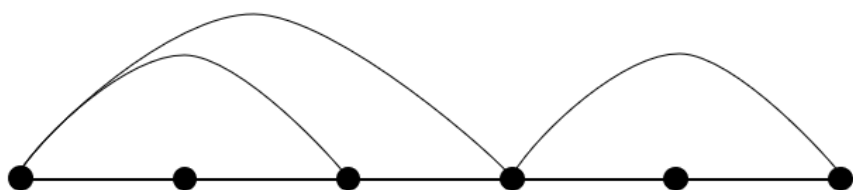
Den geometriska konstruktionen, vilket visats i Figur 6 för NVG och Figur 7 för HVG, utgår ifrån en tidsserien i Figur 4 och Figur 5. Varje observation  $(t_i, y_i)$  visas som en lodrät stapel, och streckade linjer dras mellan de datapunkter som uppfyller synlighetskriteriet.

Den algoritmiska representationen visar resultatet av denna process. Här representeras varje observation endast som en nod placerad längs en horisontell axel och bågar ovanför axeln visar vilka noder som är sammankopplade i den färdiga grafen.

Denna form används för att analysera de strukturella egenskaperna hos nätverket, till exempel gradfördelning, klustring och kortaste vägar.



Figur 6: Algoritmisk representation av NVG

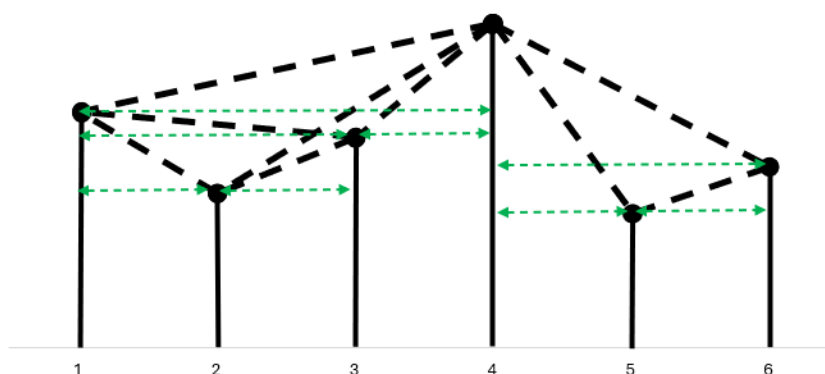


Figur 7: Algoritmisk representation av HVG

Genom att analysera de två algoritmiska representationerna blir skillnaden mellan de två Visibility Graph varianterna tydligare. I NVG av tidsserien i exemplet finns det en koppling mellan nod 2 och nod 4 som inte finns i HVG varianten av samma tidsserie.

### 2.4.5 Sambandet mellan NVG och HVG

När man ser på konceptet Visibility Graphs i sin helhet är det NVG som bäst fångar dess grundidé. HVG är en enklare version av NVG vilket kommer med en del fördelar, de är snabbare att beräkna men har också ett par unika egenskaper vilket NVG saknar.



Figur 8: NVG och HVG applicerat på samma tidsserie.

I grund och botten är HVG en subgraf av NVG, alla HVG-kanter finns alltså i NVG men inte tvärtom.

#### Proposition 2.1

För varje tidsserie  $S$  gäller att

$$E(\text{HVG}(S)) \subseteq E(\text{NVG}(S))$$

dsv, HVG är alltid en delgraf av NVG.

#### Bevis

Antag att  $(i, j)$  med  $i < j$  är en HVG-kant. Per definition betyder det att

$$y_k < \min\{y_i, y_j\} \quad \text{för alla } t_k \in (t_i, t_j) \text{ (HVG).}$$

Vi ska också visa att NVG-villkoret då också gäller, dvs för alla  $t_k \in (t_i, t_j)$ ,

$$y_k < y_i + \frac{y_j - y_i}{t_j - t_i}(t_k - t_i) \quad (\text{NVG}).$$

Betrakta den räta linjen  $L$  genom punkterna  $(t_i, y_i)$  och  $(t_j, y_j)$ . För ett givet  $t_k \in (t_i, t_j)$  finns en unik  $\lambda \in (0, 1)$  sådan att

$$t_k = (1 - \lambda)t_i + \lambda t_j, \quad \lambda = \frac{t_k - t_i}{t_j - t_i}.$$

Linjevärdet i  $t_k$  kan skrivas som den konvexa kombinationen

$$L(t_k) = (1 - \lambda)y_i + \lambda y_j = y_i + \frac{y_j - y_i}{t_j - t_i}(t_k - t_i).$$

Eftersom  $\lambda \in (0, 1)$  gäller

$$L(t_k) \in (\min\{y_i, y_j\}, \max\{y_i, y_j\}) \text{ om } y_i \neq y_j, \quad L(t_k) = y_i = y_j \text{ om } y_i = y_j.$$

I båda fallen får vi

$$L(t_k) \geq \min\{y_i, y_j\},$$

med strikt  $>$  om  $y_i \neq y_j$  och  $t_k \in (t_i, t_j)$ .

Kombinera nu detta med HVG-villkoret

$$y_k < \min\{y_i, y_j\} \leq L(t_k),$$

vilket ger NVG-villkoret

$$y_k < y_i + \frac{y_j - y_i}{t_j - t_i}(t_k - t_i).$$

Eftersom NVG gäller för varje mellanliggande  $t_k \in (t_i, t_j)$  är  $(i, j)$  en NVG-kant. Alltså  $E(\text{HVG}(S)) \subseteq E(\text{NVG}(S))$ .

### Anmärkning

Om  $y_i = y_j$  är  $L$  horisontell med nivå  $y_i$ . Då implicerar  $y_k < y_i = L(t_k)$  direkt

NVG. Om  $y_i \neq y_j$  ligger  $L(t_k)$  strikt över  $\min\{y_i, y_j\}$  för alla  $t_k \in (t_i, t_j)$ , så slutsatsen är ännu starkare.

## 2.5 Exempel på deterministiskt kaos: den logistiska avbildningen

Ett klassiskt exempel på ett enkelt system som kan ge upphov till kaotisk dynamik är den logistiska avbildningen

$$x_{t+1} = f_r(x_t) = rx_t(1 - x_t), \quad x_t \in (0, 1), \quad r \in (0, 4].$$

Parametern  $r$  styr styrkan i den icke-linjära återkopplingen och fungerar som en bifurkationsparameter: när  $r$  ändras kan systemets långsiktiga beteende skifta mellan konvergens, periodicitet och kaos.

För att beskriva konvergensen är det naturligt att introducera *fixpunkter*. En fixpunkt är ett värde  $x^*$  som lämnas oförändrat av dynamiken, dvs.

$$x^* = f_r(x^*).$$

För den logistiska avbildningen ger detta två fixpunkter

$$x_1^* = 0, \quad x_2^* = 1 - \frac{1}{r} \quad (r > 1).$$

Vilken fixpunkt (om någon) banan faktiskt konvergerar mot beror på stabilitet. I ett endimensionellt diskret dynamiskt system är en fixpunkt stabil om

$$|f'_r(x^*)| < 1,$$

där  $f'_r(x) = r(1 - 2x)$ . För  $x_1^* = 0$  fås  $|f'_r(0)| = r$ , så  $x_1^*$  är stabil för  $0 < r < 1$ . För  $x_2^* = 1 - \frac{1}{r}$  fås  $f'_r(x_2^*) = 2 - r$ , vilket innebär att  $x_2^*$  är stabil för  $1 < r < 3$ . I dessa parameterintervall konvergerar alltså tidsserien mot en konstant nivå (en stabil fixpunkt).

När  $r$  ökar förbi 3 förlorar fixpunkten  $x_2^*$  sin stabilitet och systemet går över till periodiska banor. Därefter följer en periodfördubblingskaskad (period 2, 4, 8, ...).

Efter ackumulationspunkten

$$r_\infty \approx 3.56995$$

uppträder kaotisk dynamik för ett stort intervall av parameter-värden (med vissa “fönster” av periodicitet).  $r$  väljs i en kaotisk regim, typiskt  $r \gtrsim 3.57$ ; ett vanligt val är  $r = 4$ , vilket ger fullt utvecklat kaos.

Den logistiska avbildningen används i detta arbete som en prototyp för deterministiskt kaos. Genom att applicera NVG och HVG på tidsserier genererade av den logistiska avbildningen i kaotisk regim kan vi jämföra deras grafegenskaper med motsvarande för stokastiska tidsserier (t.ex. vitt brus) och finansiella data.

### 3 Teori för Visibility Graphs

I detta kapitel studeras några av de matematiska egenskaperna hos Visibility Graphs, däribland invariants under transformationer och fördelningar av nodernas grader. Syftet är att visa hur strukturen hos en tidsserie påverkar grafens formella egenskaper.

#### 3.1 Invariants under strikt växande transformationer för HVG

##### Sats 3.1

Låt  $S = \{x_1, \dots, x_n\}$  vara en tidsserie och låt  $G_h(S)$  beteckna den HVG som konstrueras från  $S$  enligt kriteriet

$$x_i, x_j > x_k \quad \text{för alla } i < k < j.$$

Låt  $f : \mathbb{R} \rightarrow \mathbb{R}$  vara en strikt växande funktion och sätt  $f(S) = \{f(x_1), \dots, f(x_n)\}$ . Då gäller

$$G_h(S) = G_h(f(S)),$$

dvs, HVG är oförändrad under alla strikt monotona transformationer av värdeaxeln. I synnerhet beror HVG:ns struktur endast på rangordningen av datapunkterna och inte på deras exakta värden.

##### Bevis

Antag  $i < j$ . En kant  $(i, j)$  finns i  $G_h(S)$  om och endast om

$$x_k < \min(x_i, x_j) \quad \text{för alla } i < k < j.$$

Om  $f$  är strikt växande är ordningen bevarad:  $x_a < x_b \Leftrightarrow f(x_a) < f(x_b)$ . Alltså gäller

$$x_k < \min\{x_i, x_j\} \Leftrightarrow f(x_k) < \min\{f(x_i), f(x_j)\}.$$

vilket innebär att  $(i, j)$  är i kant i  $G_h(f(S))$  precis när det är kant i  $G_h(S)$ . Om  $f$  är strikt avtagande byter alla olikheter riktning men samma ordningsrelation på intervallet  $i < k < j$  avgörs av jämförelser mot  $\max\{x_i, x_j\}$ , och eftersom HVG villkoret använder min-tröskeln symmetriskt för  $i$  och  $j$  erhålls ekvivalent:

$$x_k < \min\{x_i, x_j\} \Leftrightarrow f(x_k) > \max\{f(x_i), f(x_j)\}.$$

Men detta är samma kantvillkor uttryckt på den transformerade skalan, så att  $(i, j)$  är kant i  $G_h(f(S))$  om och endast om  $(i, j)$  är kant i  $G_h(S)$ . Alltså sammanfaller kantmängderna och  $G_h(S) = G_h(f(S))$ .

## 3.2 Gradfördelning i HVG

### Definition 3.1 Concatenation

En concatenation i en HVG avser den rekursiva sammansättningen av delintervall som uppstår när inre noder mellan två givna noder läggs till under bibehållen siktlinje.

Formellt betyder detta att en sekvens av inre noder  $x_{i_1}, x_{i_2}, \dots, x_{i_m}$  är concatenated om varje nod uppfyller synlighetsvillkoret

$$x_{i_j} < \min(x_{i_{j-1}}, x_{i_{j+1}}), \quad j = 2, \dots, m - 1.$$

Varje sådan sekvens bidrar med en multiplikativ integralterm  $[I]$  i sannolikhetsberäkningen.

### Lemma 3.1 Concatenation-egenskaper

Låt  $G$  vara den HVG som konstrueras från en oberoende och identiskt fördelad (i.i.d.) sekvens av kontinuerliga slumpvariabler.

Då kan sannolikheten att en fixerad nod har grad  $k$  uttryckas som summan av bidrag från  $k - 1$  disjunkta konfigurationer  $C_j$ :

$$P(K = k) = \sum_{j=0}^{k-2} [S][B]^2 [I]^j [I]^{k-2-j},$$

där  $[S]$  representerar bidraget från den fixerade noden ("seed"),  $[B]$  grannvillkoren ("bounding terms"), och  $[I]$  concatenation-integralen som beskriver de inre synliga noderna.

Beviset bygger på induktionen över  $k$ ; se Luque et al. (2009).

### Sats 3.2

Låt  $\{X_i\}_{i \in \mathbb{Z}}$  vara oberoende och identiskt fördelade med kontinuerlig fördelning, och låt  $G$  vara HVG konstruerad från serien enligt kriteriet

$$x_i, x_j > x_n \quad \text{för alla } i < n < j.$$

Beteckna med  $K$  graden för en godtycklig nod i  $G$ . Då gäller

$$\mathbb{P}(K = k) = \frac{1}{3} \left(\frac{2}{3}\right)^{k-2}, \quad k = 2, 3, \dots$$

Oberoende av den underliggande tätheten  $f$ . I synnerhet är  $\mathbb{E}[K] = 4$ .

### Bevis

Eftersom HVG-kriteriet beror på ordningen mellan observationerna, kan vi (vid kontinuerlig fördelning) utan förlust anta stöd  $[0, 1]$  med täthet  $f$  och fördelningsfunktion  $F$ .

Vi fixerar en nod  $x_0$  och beräknar sannolikheten för grad  $k$  via integraler i  $x_0$ .

*Steg 1: Fallet  $k = 2$ .*

Den minsta möjliga grad i en HVG är 2. Villkoret för exakt grad 2 är att båda grannarna ligger över  $x_0$ , så att inga ytterligare inre noder blir synliga. Detta ger

$$\mathbb{P}(K = 2) = \int_0^1 f(x_0) \left( \int_{x_0}^1 f(x_1) dx_1 \right) \left( \int_{x_0}^1 f(x_{-1}) dx_{-1} \right) dx_0 = \int_0^1 f(x_0)[1-F(x_0)]^2 dx_0.$$

Med standardidentiteten  $f(x)F^{n-1}(x) = \frac{1}{n} \frac{d}{dx} F^n(x)$  blir

$$\int_0^1 [1 - F(x_0)]^2 f(x_0) dx_0 = \int_0^1 (1 - u)^2 du = \frac{1}{3}.$$

Alltså  $\mathbb{P}(K = 2) = \frac{1}{3}$ .

*Steg 2: Allmänt  $k \geq 3$ .*

För grad  $k$  finns  $k - 2$  ”inre” synliga noder mellan grannarna. Dessa kan placeras med  $j$  på vänster sida och  $k - 2 - j$  på höger sida, vilket ger  $k - 1$  disjunkta konfigurationer  $C_j$ . Enligt Lemma 3.1 kan varje konfiguration uttryckas som produkten av byggblocken  $[S]$ ,  $[B]^2$  och  $[I]^{k-2}$ , där concatenation-reglerna ger upphov till logaritm faktorer i integralen. Efter korrekt vägning av varje konfiguration med symmetrifaktorn

$$p_j = \frac{(k-2)!}{j!(k-2-j)!} 2^{-(k-2)},$$

reduceras summan till

$$P(K = k) = \sum_{j=0}^{k-2} \frac{(-1)^{k-2}}{j!(k-2-j)!} \int_0^1 f(x_0)[1 - F(x_0)]^2 [\ln(1 - F(x_0))]^{k-2} dx_0.$$

*Steg 3: Slutledning*

Med substitutionen  $u = 1 - F(x_0)$  och binomial-summering fås

$$P(K = k) = \frac{1}{3} \left( \frac{2}{3} \right)^{k-2}, \quad k \geq 2.$$

Eftersom  $f$  endast förekommer via  $F$  beror resultatet inte på den specifika tätheten.

Slutligen ges

$$\mathbb{E}[K] = \sum_{k=2}^{\infty} k \cdot \frac{1}{3} \left(\frac{2}{3}\right)^{k-2} = 4.$$

Alltså har en slumpmässig HVG alltid förväntad grad 4, oberoende av fördelningen för den underliggande tidsserien. Därmed är satsen bevisad.

### Anmärkning

I en ändlig tidserie har de två yttersta datapunkterna endast en granne. Det vill säga  $K = 1$ . I härledningen ovan antas däremot en oändlig tidserie, vilket gör att varje nod har minst två grannar och därmed  $k_{\min} = 2$ .

## 3.3 Visibility graphs för periodiska, stokastiska och kaotiska tidsserier

Tidigare resultat visar att olika typer av tidsserier ger upphov till principiellt olika strukturer i HVG och NVG. Förenklat kan man säga att

- *periodiska* tidsserier ger upphov till mycket regelbundna grafer, ofta med begränsat antal gradvärden och stark lokal struktur,
- *stokastiska* tidsserier (t.ex. i.i.d. vitt brus) ger grafer med exponentiell gradfördelning enligt satsen ovan, där kantstrukturen i huvudsak saknar långdistanskorrelationer,
- *kaotiska* tidsserier hamnar mellan dessa extremfall: graferna är varken helt regelbundna eller helt slumpmässiga, och uppvisar typiskt både icke-trivial gradfördelning och klustringsstruktur.

En central fråga i denna uppsats är om sådana skillnader i grafegenskaper kan användas för att skilja stokastiska tidsserier från deterministiskt kaotiska system, särskilt när det tillförs additivt brus. I nästa kapitel formaliseras denna fråga som en jämförelse mellan:

1. HVG/NVG konstruerade från vitt brus,
2. HVG/NVG konstruerade från den logistiska avbildningen i en kaotisk regim,
3. samma deterministiska serie med tillagt vitt brus,
4. finansiella tidsserier (SPY).

## 4 Metod

I detta kapitel beskrivs den allmänna metod som används för att analysera tidsserier med hjälp av Visibility Graphs. Metoden är formulerad så att den kan tillämpas på olika typer av tidsserier (finansiella, biologiska, klimatdata m.m.) och är därför oberoende av specifik datakälla.

### 4.1 Utgångspunkt: tidsserie

Låt  $\{x_t\}_{t=1}^T$  beteckna en diskret tidsserie. Vid behov kan serien förbehandlas genom t.ex. logaritmering, differensbildning eller normalisering. Denna förbehandling påverkar inte konstruktionen av HVG när transformationen är strikt växande (jfr Sats 3.1).

### 4.2 Konstruktion av synlighetsgrafer

För varje tidsserie konstrueras två grafer:

- en NVG enligt Definition 2.10
- en HVG enligt Definition 2.11

I båda fallen erhålls en odirigerad, oviktad graf  $G = (V, E)$  där varje tidssteg motsvarar en nod.

### 4.3 Begränsat siktfönster

För vissa analyser används ett tidsfönster av längd  $S$ , dvs. endast observationer inom intervallet  $[t - S, t]$  (eller ett symmetriskt fönster kring  $t$ ) tillåts bidra till synlighetsrelationen. Detta gör det möjligt att definiera lokala indikatorer som beskriver den aktuella formen på tidsserien.

### 4.4 Konkava och konvexa indikatorer

För att kunna skilja mellan olika typer av rörelser (accelererade respektive avtagande uppgångar och nedgångar) introducerar vi lokala indikatorer som mäter hur väl ett tidsfönster kan approximeras av en konkav respektive konvex rörelse.

Låt  $\{x_t\}_{t=1}^T$  vara en diskret tidsserie och fixera en tidpunkt  $b$  samt ett siktfönster

$S \in \mathbb{N}$  med  $S < T/2$ . Vi betraktar endast observationer inom intervallet  $[b - S, b]$ , så att avståndet mellan  $t_a$  och  $t_b$  uppfyller

$$|a - b| \leq S.$$

### Konkaviteten, konvexiteten och accelererande rörelser

Valet att beskriva indikatorerna som konkava respektive konvexa hänger direkt samman med hur lutningen förändras över tid. I kontinuerlig notation ges lutning av den första derivatan  $x'(t)$ , medan den andra derivatan  $x''(t)$  mäter hur lutningen förändras (acceleration). En *konvex* (convex) funktion har  $x''(t) > 0$ , vilket innebär att lutningen ökar: en uppgång blir allt brantare (accelererande uppgång), medan en nedgång bromsar in. En *konkav* (concave) funktion har i stället  $x''(t) < 0$ , så att lutningen minskar: en uppgång planar ut (avtagande uppgång) och en nedgång blir successivt brantare (accelererande nedgång).

### Uppgång och nedgång

För att skilja på uppgångar och nedgångar delar vi in grannarna till  $b$  i två mängder, beroende på om  $x_b$  är högre eller lägre än  $x_a$ :

- *Uppgång* (rise):  $x_b > x_a$ ,
- *Nedgång* (fall):  $x_b < x_a$ .

Vi kommer att beteckna dessa två fall med index  $X \in \{R, F\}$ .

### Konkava indikatorer via Visibility Graph

Utgångspunkten är en (modifierad) Visibility Graph, där vi tillåter länkar endast bakåt i tiden (från  $b$  till  $a < b$ ) och endast inom fönstret  $S$ . Givet NVG-kriteriet innebär detta att en nod  $(t_a, x_a)$  är synlig från  $(t_b, x_b)$  om

$$x_c < x_b + \frac{x_a - x_b}{t_b - t_a} (t_b - t_c) \quad \text{för alla } a < c < b.$$

Vi definierar först mängden av alla synliga noder till  $b$ :

$$A = \left\{ (t_a, x_a) \mid x_c < x_b + \frac{x_a - x_b}{t_b - t_a} (t_b - t_c) \text{ för alla } a < c < b, 1 \leq a < b \leq T \right\}.$$

Med de extra villkoren för uppgång/nedgång och det begränsade siktfönstret fås därefter

$$A_R = \{(t_a, x_a) \in A \mid x_b > x_a, |a - b| \leq S\},$$

$$A_F = \{(t_a, x_a) \in A \mid x_b < x_a, |a - b| \leq S\}.$$

Graden för nod  $b$  i den konkava uppgångs- respektive nedgångsgrafan definieras som

$$D_X^{\text{VG}}(b) = |A_X|, \quad X \in \{R, F\}.$$

Den *konkava indikatorn* vid tidpunkt  $b$  ges sedan av

$$I_X^{\text{CC}}(b) = \frac{D_X^{\text{VG}}(b)}{S}, \quad X \in \{R, F\}.$$

Här tar  $I_X^{\text{CC}}(b)$  värden i intervallet  $[0, 2]$  (i praktiken ofta  $[0, 1]$  för rimliga fönster), där höga värden motsvarar en ”renare” konkav uppgång (för  $X = R$ ) eller konkav nedgång (för  $X = F$ ) i fönstret kring  $b$ .

### Konvexa indikatorer via Invisibility Graph

På motsvarande sätt kan vi definiera *konvexa* indikatorer med hjälp av en Invisibility Graph, där vi istället kopplar ihop datapunkter vars förbindelselinje skärs av mellanliggande punkter. Här vänder vi alltså på olikheten och kräver

$$x_c > x_b + \frac{x_a - x_b}{t_b - t_a} (t_b - t_c) \quad \text{för minst ett } a < c < b.$$

Vi definierar mängden osynliga noder till  $b$  som

$$A' = \{(t_a, x_a) \mid x_c > x_b + \frac{x_a - x_b}{t_b - t_a} (t_b - t_c) \text{ för något } a < c < b, 1 \leq a < b \leq T\}.$$

Med samma uppgång/nedgång-indelning och fönsterbegränsning får vi

$$A'_R = \{(t_a, x_a) \in A' \mid x_b > x_a, |a - b| \leq S\},$$

$$A'_F = \{(t_a, x_a) \in A' \mid x_b < x_a, |a - b| \leq S\}.$$

Graden för nod  $b$  i de konvexa upp- och nedgångsgraferna är

$$D_X^{\text{IVG}}(b) = |A'_X|, \quad X \in \{R, F\},$$

och den *konvexa indikatorn* definieras som

$$I_X^{\text{CV}}(b) = \frac{D_X^{\text{IVG}}(b)}{S}, \quad X \in \{R, F\}.$$

Höga värden på  $I_R^{\text{CV}}(b)$  motsvarar då en tydligt konvex (avtagande) uppgång i fönstret kring  $b$ , medan höga värden på  $I_F^{\text{CV}}(b)$  beskriver en tydligt konvex (accelererad) nedgång.

## Tolkning

Sammanfattningsvis ger de fyra indikatorerna

$$I_R^{\text{CC}}, I_F^{\text{CC}}, I_R^{\text{CV}}, I_F^{\text{CV}}$$

ett lokalt mått på hur väl tidsserien följer en konkav respektive konvex uppgång eller nedgång inom ett siktfönster av längd  $S$ . Genom att studera deras fördelningar över tid kan man kvantifiera asymmetrier mellan accelererade och avtagande prisrörelser.

## 4.5 Nätverksmått

För varje konstruerad graf beräknas därefter standardmått från grafteori:

- grad och gradfördelning  $\hat{P}(k)$
- medellängd för kortaste vägar  $L$
- diameter  $\text{diam}(G)$
- lokal och global klustringskoefficient

Dessa mått kan senare jämföras mellan olika tidsserier eller mot referensmodeller.

## 4.6 Simulerade referensserier: vitt brus och deterministiskt kaos

För att belysa skillnaderna mellan stokastiska och deterministiskt kaotiska tidsserier används två enkla referensmodeller:

1. **Vitt brus:** en i.i.d. sekvens  $\{\varepsilon_t\}$  med  $\varepsilon_t \sim \mathcal{N}(0, 1)$ .

2. **Logistisk avbildning:** en deterministisk serie  $\{x_t\}$  given av

$$x_{t+1} = rx_t(1 - x_t),$$

där  $r$  väljs i en kaotisk regim (t.ex.  $r = 4$ ) och  $x_0 \in (0, 1)$ .

För att undersöka effekten av brus konstrueras dessutom en serie

$$y_t = x_t + \sigma\varepsilon_t,$$

där  $\{\varepsilon_t\}$  är vitt brus och  $\sigma > 0$  styr brusstyrkan.

På varje serie (vitt brus, logistisk utan brus, logistisk med brus) konstrueras NVG och HVG enligt Kapitel 2. För dessa grafer beräknas:

- gradfördelning,
- klustringskoefficient,
- de konkava och konvexa indikatorerna  $I_X^{CC}$  och  $I_X^{CV}$ .

Dessa resultat används som jämförelsebas när vi analyserar SPY i Kapitel 5. Särskilt intressant är om visibility- och invisibility-graferna för SPY liknar mer de stokastiska eller de deterministiskt kaotiska referensfallen, och hur detta påverkas av tillfört brus.

## 4.7 Rullande analys

Slutligen kan metoden tillämpas i rullande tidsfönster för att studera hur grafegenskaper förändras över tid. För varje fönster konstrueras en graf och dess mått lagras, vilket ger tidsserier av grafindikatorer.

## 5 Finansiella Tillämpningar

Tidsserieanalys med Visibility Graphs kan användas inom en mängd olika områden. I detta arbete presenteras, undersöks och diskuteras en finansiell tillämpning, mer specifikt hur metoden kan användas på ett aktiediagram.

Att analysera en aktie enbart genom att betrakta dess prisgraf, och alltså helt bortse från bolagets fundamentala prestation, kallas *teknisk analys*. Enkelt uttryckt använder investerare som tillämpar teknisk analys olika typer av indikatorer för att avgöra om det är läge att köpa eller sälja en aktie. Det finns en mängd olika indikatorer med varierande komplexitetsnivå, men ett av de vanligaste och mest intuitiva exemplen är glidande medelvärde (moving average).

Det fungerar så att man beräknar ett genomsnitt av de senaste datapunkterna, ofta de senaste 50 eller 200 dagarna. Detta skapar en ny graf som sedan kan jämföras med den ursprungliga aktiekursgraf. Handlas aktien under sitt glidande medelvärde kan det tolkas som ett svaghetstecken (och därmed en säljsignal), medan handel över medelvärdet kan indikera styrka (och därmed ett köpläge). I praktiken används ofta mer avancerade strategier, där flera olika indikatorer kombineras för att ge en samlad signal om köp eller sälj.

Huruvida teknisk analys faktiskt fungerar är dock mycket omdebatterat. Kritiker menar att den saknar verkligt värde och att investerare som enbart förlitar sig på denna metod snarare ägnar sig åt ett slags casinospel, där slumpen avgör om man tjänar pengar eller inte.

Oavsett om man är förespråkare eller kritiker av teknisk analys kan man dock enas om en sak: alla skulle vara intresserade av en indikator som på ett tillförlitligt sätt kan förutsäga framtida uppgångar eller nedgångar i en aktie. I detta arbete undersöks därför om Visibility Graphs kan användas för att skapa en sådan indikator, samt hur väl den fångar marknadens upp- och nedgångar i praktiken. Fokus ligger inte på att konstruera en fullständig handelsstrategi, utan på att kvantifiera formen på prisrörelser och belysa asymmetrier mellan upp- och nedgångar.

## 5.1 Data och empirisk uppställning

I den empiriska delen analyseras prisutvecklingen i SPY, en börshandlad fond (ETF) som följer utvecklingen av det amerikanska aktieindexet S&P 500. När man köper SPY köper man i praktiken en liten del av de 500 största börsnoterade företagen i USA, såsom Apple, Microsoft och Amazon. SPY skapades för att göra det enkelt för investerare att investera i hela den amerikanska aktiemarknaden utan att behöva köpa varje aktie separat. Eftersom SPY mycket nära speglar utvecklingen på S&P

500 används den ofta som ett proxy för hela aktiemarknaden.

I analysen används dagliga prisdata för SPY under perioden

[januari 2000 till november 2025].

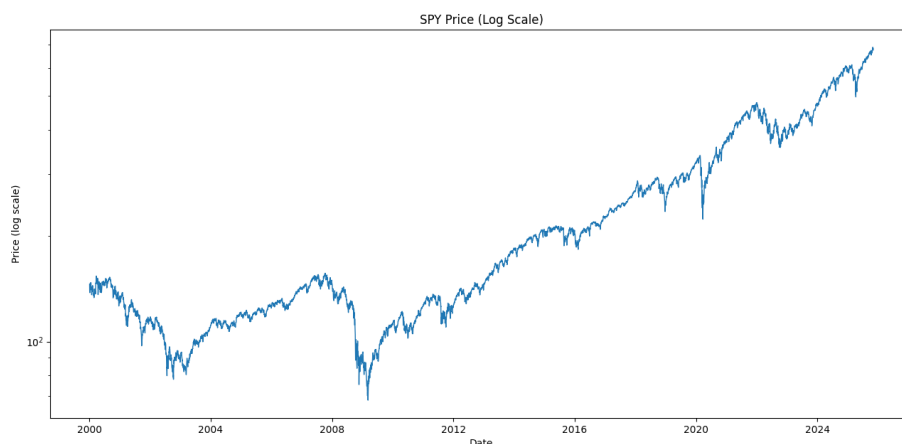
Som prisvariabel används *justerad stängningskurs* (“adjusted close”), vilket innebär att utdelningar, aktiesplittar och andra bolagshändelser har justerats för. På så sätt motsvarar serien en investerare som återinvesterar utdelningar i fonden.

Endast handelsdagar ingår i serien, vilket innebär att avståndet mellan två efterföljande observationer inte alltid motsvarar en kalenderdag (helger och helgdagar saknar observationer). Detta är standard inom finansiell tidsserieanalys och påverkar inte konstruktionen av visibility-graferna, eftersom dessa bygger på ordningen mellan observationerna och inte på exakt tidsavstånd i kalendertiden.

Priser ritas i log-skala för att underlätta visuell jämförelse mellan tidiga perioder med lägre prisnivåer och senare perioder då SPY handlas på betydligt högre nivåer. Själva visibility-graferna konstrueras dock på de faktiska prisnivåerna  $x_t$ , i enlighet med definitionerna i Kapitel 2.

## 5.2 SPY

I detta arbete kommer vi att använda SPY som ett konkret exempel på en tillgång där form på prisrörelser är av intresse. Eftersom SPY kan betraktas som ett proxy för den amerikanska aktiemarknaden i stort är den väl lämpad för vårt syfte, då vi vill undersöka om Visibility Graphs kan användas för att skapa indikatorer som fångar marknadens upp- och nedgångar. Figur 9 visar prisutvecklingen i SPY under den studerade perioden. De mest framträdande rörelserna är IT-bubblan i början av 2000-talet, finanskrisen 2008–2009 och fallet i samband med Covid-19 våren 2020, följt av en kraftig återhämtning. Ett återkommande tema i finansiell litteratur är att uppgångar ofta är relativt långsamma och gradvisa, medan nedgångar tenderar att vara snabbare och mer koncentrerade i tiden. Ett av målen i denna del är att undersöka om detta mönster kan fångas och kvantifieras med hjälp av Visibility- och Invisibility-grafer.



Figur 9: Utveckling i SPY sedan år 2000.

### 5.3 Konkava och konvexa årsindikatorer för SPY

Med utgångspunkt i den metod som beskrivs i föregående kapitel beräknas nu de fyra indikatorerna

$$I_R^{CC}, I_F^{CC}, I_R^{CV}, I_F^{CV}$$

för den dagliga prisserien i SPY. I den empiriska implementeringen används ett tidsfönster om

$$S = 262$$

observationer, vilket motsvarar ungefär antalet handelsdagar under ett år. Varje indikatorvärde vid tidpunkt  $b$  kan därför tolkas som ett *årsbaserat*, lokalt mått på hur konkav eller konvex uppgången respektive nedgången har varit under det senaste handelsåret.

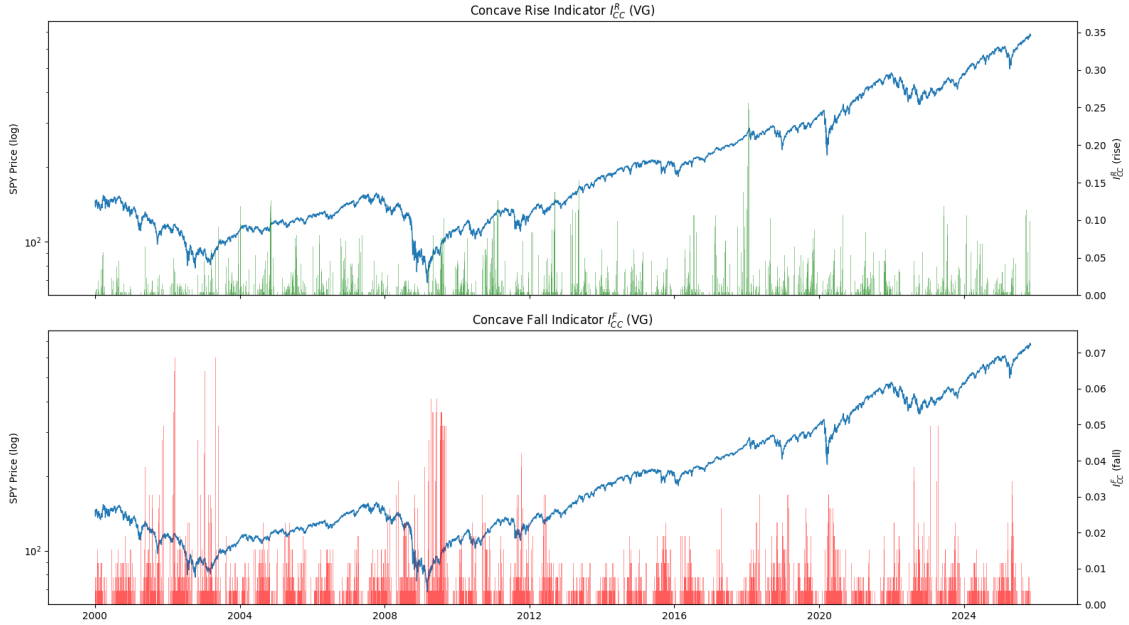
Beräkningen följer direkt definitionerna i avsnitt 4.4. För varje tidpunkt  $b$  betraktas endast observationer bakåt i tiden,  $a < b$ , med  $|a - b| \leq S$ . Detta gör indikatorerna strikt bakåtblickande: de beskriver formen på prisbanan som är känd för en investerare vid datum  $b$ , utan att använda någon information från framtida datapunkter. För varje sådan granne  $a$  kontrolleras först om NVG-kriteriet (konkav siktlinje) respektive IVG-kriteriet (konvex "osynlighetslinje") är uppfyllt, och därefter om rörelsen mellan  $a$  och  $b$  är en uppgång ( $x_b > x_a$ ) eller nedgång ( $x_b < x_a$ ). Antalet synliga/osynliga grannar summeras till

$$D_X^{VG}(b), \quad D_X^{IVG}(b),$$

och normaliseras sedan med fönsterlängden  $S$ :

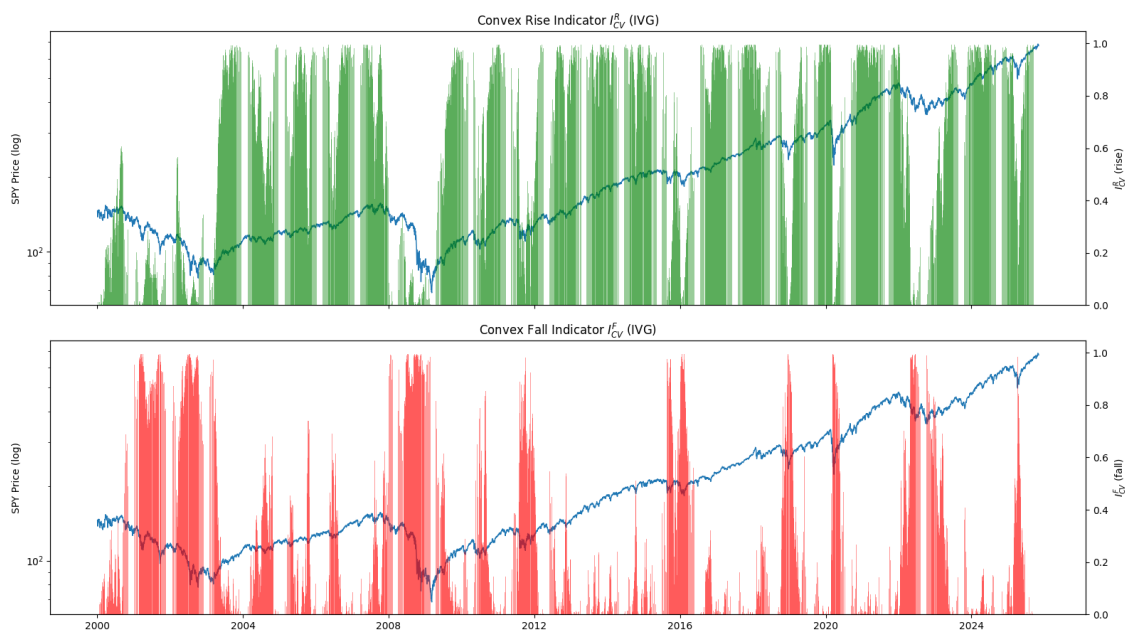
$$I_X^{\text{CC}}(b) = \frac{D_X^{\text{VG}}(b)}{S}, \quad I_X^{\text{CV}}(b) = \frac{D_X^{\text{IVG}}(b)}{S}, \quad X \in \{R, F\}.$$

Eftersom högst  $S$  bakåtriktade noder kan bidra ligger indikatorerna i intervallet  $[0, 1]$  (upp till små randeffekter i början av serien).



Figur 10: Konkava upp- och nedgångsindikatorer  $I_R^{\text{CC}}$  respektive  $I_F^{\text{CC}}$  för SPY med års-fönster  $S = 262$ .

I Figur 10 visas de konkava indikatorerna baserade på Visibility Graph. De gröna staplarna  $I_R^{\text{CC}}$  markerar perioder där priset under det senaste handelsåret huvudsakligen rört sig i en mer ”utplanande” uppgång mot dagens nivå. De röda staplarna  $I_F^{\text{CC}}$  fångar i stället konkava nedgångar, det vill säga utdragna fall som bromsar in mot slutet.



Figur 11: Konvexa upp- och nedgångsindikatorer  $I_R^{CV}$  respektive  $I_F^{CV}$  för SPY med års-fönster  $S = 262$ .

Figur 11 visar de konvexa indikatorerna baserade på Invisibility Graph. Höga värden i  $I_R^{CV}$  (gröna staplar) motsvarar snabba, accelererande uppgångar där en stor del av prISRörelsen under året koncentreras till de senaste månaderna. Motsvarande gäller för  $I_F^{CV}$  (röda staplar), som fångar accelererande nedgångar.

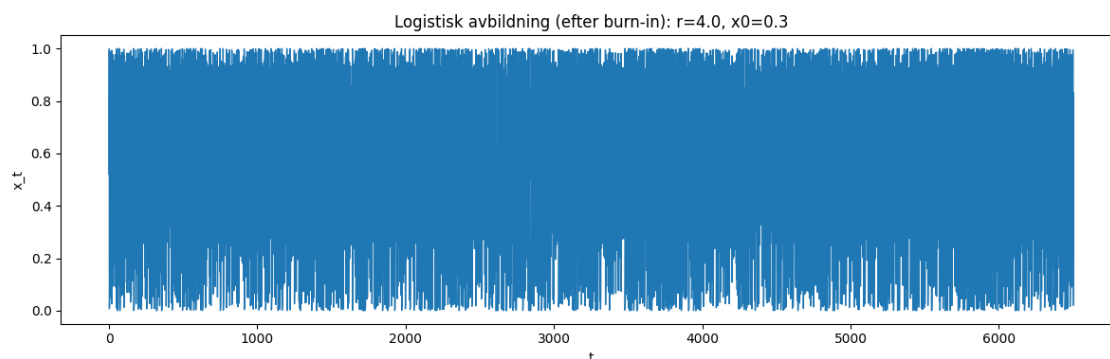
## 5.4 Jämförelse med stokastiska och kaotiska referensserier

För att bedöma huruvida SPY uppvisar egenskaper som liknar en stokastisk process eller deterministisk kaosdynamik jämförs dess Horizontal Visibility Graph (HVG) med två referensserier: i) vitt brus (i.i.d. normalfördelat), och ii) den logistiska avbildningen i kaotisk regim.

Den logistiska avbildningen definieras som

$$x_{t+1} = rx_t(1 - x_t),$$

och för  $r = 4.0$  uppvisar systemet fullt utvecklad deterministisk kaosdynamik. Efter en initial burn-in-fas erhålls den stationära kaotiska tidsserien som används som referens i analysen, vilken visas i Fig. 12.



Figur 12: Logistisk avbildning efter burn-in i kaotisk regim ( $r = 4.0$ ,  $x_0 = 0.3$ ). Serien används som deterministisk kaotisk referens i HVG-jämförelsen.

Enligt Sats 3.2 ska en HVG genererad från ett i.i.d. brus anta den teoretiska gradfördelningen

$$P(K = k) = \frac{1}{3} \left(\frac{2}{3}\right)^{k-2}, \quad k \geq 2, \quad (1)$$

vilket innebär en geometriskt avtagande sannolikhet. Denna form bekräftas av den simulerade HVG:n för vitt brus, som ligger nära teorikurvan i Figur 13a. SPY avviker däremot systematiskt: sannolikheten för låga gradtal ( $K = 2$ ) är lägre, medan sannolikheten för mellanhöga gradtal ( $K = 3-6$ ) är högre än för vitt brus. Detta innebär att fler noder i SPY har flera synliga grannar, vilket tyder på kortsiktig minnesstruktur och icke-i.i.d. dynamik i prisserien.

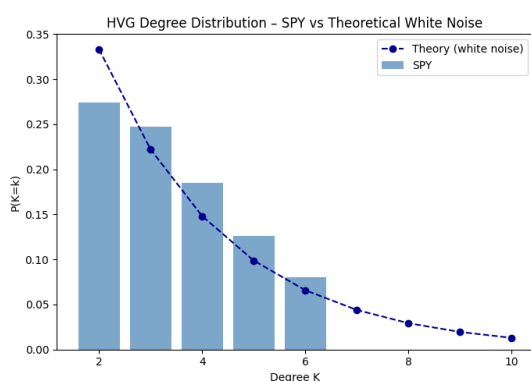
Utöver gradfördelningen ger även klustringskoefficienten viktig information. Som framgår av Tabell 1 är klustringen högst i den logistiska kartan ( $C \approx 0.67$ ), något lägre i vitt brus ( $C \approx 0.65$ ) och lägst i SPY ( $C \approx 0.59$ ). I HVG motsvarar hög klustring att tre datapunkter ofta är ömsesidigt synliga, vilket indikerar lokalt återkommande struktur i tidsserien. Den logistiska kartans höga klustring är därför förenlig med dess deterministiska dynamik, medan vitt brus får en något lägre nivå på grund av avsaknad av minne.

Att SPY uppvisar lägst klustringskoefficient innebär att prisserien bildar färre lokala trekantar i HVG än både en slumpserie och ett kaotiskt system. Detta kan tolkas som att prisrörelserna är mer varierade och mindre regelbundet återkommande än i kaotiska system, men att gradfördelningen ändå visar tydlig avvikelse från i.i.d. vitt brus. Tillsammans pekar därför (i) fler mellanhöga gradtal än vitt brus och (ii) lägre

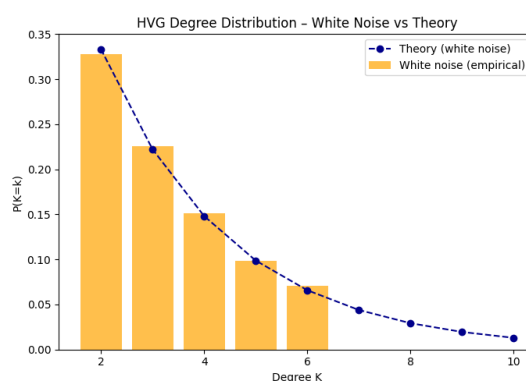
Tabell 1: HVG-mått för SPY, vitt brus och logistisk karta.

Serie	Medelgrad	$C$	$P(K=2)$	$P(K=3)$	$P(K=4)$	$P(K=5)$	$P(K=6)$
<b>SPY (empiri)</b>	3.842	0.5944	0.2743	0.2473	0.1849	0.1261	0.0799
<b>Vitt brus (empiri)</b>	3.994	0.6457	0.3279	0.2258	0.1510	0.0988	0.0707
<b>Vitt brus (teori)</b>	4	–	0.3333	0.2222	0.1481	0.0988	0.0658
<b>Logistisk karta</b>	3.993	0.6707	0.3347	0.3307	0.1007	0.0575	0.0375

*klustring än både brus och kaos* mot att SPY befinner sig mellan dessa ytterligheter: strukturen är mer organiserad än ren slump, men mer stokastisk och fragmenterad än deterministisk kaostrend.



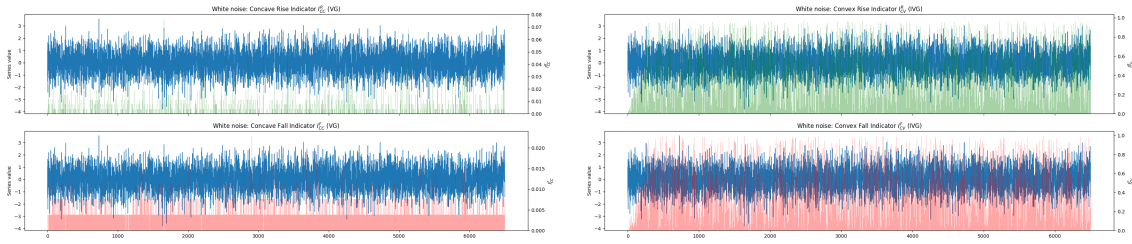
(a) SPY mot teoretisk HVG-fördelning. SPY visar lägre andel  $K=2$  och högre andel  $K=3-6$ .



(b) Vitt brus mot teoretisk HVG-fördelning. Den empiriska distributionen följer teorin väl.

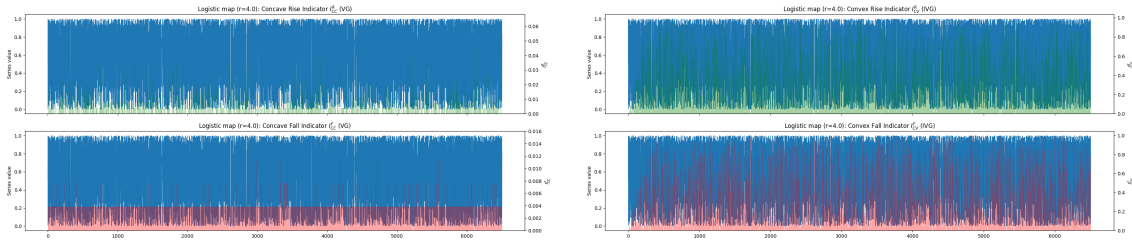
Figur 13: Jämförelse mellan empiriska HVG-gradfördelningar och teoretisk geometrisk fördelning för vitt brus. (a) SPY avviker från teorin, (b) vitt brus bekräftar den.

För att ytterligare belysa skillnaderna mellan stokastisk och deterministisk dynamik analyseras även lokala geometriska egenskaper hos tidsserierna via konkava och konvexa indikatorer baserade på Visibility Graphs (VG och IVG). Figurerna 14 och 15 visar motsvarande indikatorer för vitt brus respektive den logistiska avbildningen. Vitt brus uppvisar mer symmetriska och jämnt spridda indikatorer, i linje med avsaknad av deterministisk struktur, medan den logistiska avbildningen uppvisar tydligare variation och asymmetri, vilket speglar dess icke-linjära deterministiska dynamik.



(a) Vitt brus: konkava indikatorer (VG). (b) Vitt brus: konvexa indikatorer (IVG).

Figur 14: Konkava och konvexa indikatorer för vitt brus.



(a) Logistisk avbildning: konkava indikatorer (VG). (b) Logistisk avbildning: konvexa indikatorer (IVG).

Figur 15: Konkava och konvexa indikatorer för den logistiska avbildningen i kaotisk regim ( $r = 4.0$ ).

Sammanfattningsvis visar både gradfördelningen och klustringskoefficienten att SPY varken beter sig som ett rent vitt brus (gradfördelningen avviker från  $\frac{1}{3}(2/3)^{k-2}$ ) eller som en lågdimensionell deterministisk kaoskarta (klustringen är lägre än både vitt brus och logistikserien). Resultaten antyder i stället en blandad dynamik där SPY innehar stokastiska komponenter men även icke-linjär intern struktur, vilket stämmer överens med empiriska observationer från finansiella marknader där volatilitet och avkastningsprocesser ofta uppvisar autokorrelation och regimskillnader.

## 5.5 Diskussion av indikatorerna

Bilderna ovan visar att de fyra indikatorerna fångar olika aspekter av prisbanan i SPY. De konkava indikatorerna  $I_R^{CC}$  och  $I_F^{CC}$  uppträder som relativt spridda och lågintensiva signaler. Höga värden för  $I_R^{CC}$  tenderar att sammanfalla med längre uppgångsperioder, medan  $I_F^{CC}$  tar större värden under utdragna fall. Samtidigt är nivåerna ofta måttliga och fördelade över många tidpunkter. Indikatorerna sammanfattar därför främst formen på *det gångna året* snarare än att ge en skarp signal om ett förestående trendbrott.

Ur ett praktiskt perspektiv tyder detta på att de konkava indikatorerna är svåra att använda som ensam grund för att ”tajma” marknaden. De kan snarare tolkas som regressiva lägesmått: höga värden indikerar att prisrörelserna under året varit systematiskt utplanande i upp- eller nedgång, men det är inte uppenbart hur detta direkt översätts till en köp- eller säljsignal. Ett mer rimligt användningsområde är att använda dem som *komplement* till andra indikatorer, till exempel för att beskriva i vilken typ av trendmiljö en strategi verkar.

De konvexa indikatorerna  $I_R^{CV}$  och  $I_F^{CV}$  uppvisar ett tydligare mönster. I Figur 11 dominerar  $I_R^{CV}$  under långa uppgångsperioder, medan  $I_F^{CV}$  tar över under större krisepisoder som IT-kraschen, finanskrisen 2008–2009 och fallet i samband med Covid-19. Indikatorerna fungerar därmed som ett slags regimfilter: vid varje tidpunkt är det typiskt antingen den konvexa uppgångs- eller den konvexa nedgångsindikatorn som är dominerande.

Detta väcker frågan om de konvexa indikatorerna skulle kunna användas mer direkt i en handels- eller riskhanteringsregel, till exempel genom att minska exponeringen när  $I_F^{CV}$  överskrider en viss nivå. Resultaten säger dock inget om huruvida en sådan regel skulle vara lönsam eller stabil över tid; för att besvara den frågan krävs formella baktests och jämförelser med enklare tumregler baserade på till exempel volatilitet eller glidande medelvärden.

Sammantaget visar analyserna att visibility- och invisibility-baserade indikatorer kan beskriva den välkända asymmetrin i aktiemarknaden: uppgångar tenderar att vara långsamma och relativt konkava, medan nedgångar ofta är snabbare, mer konvexa och klustrade i tiden. I denna uppsats används indikatorerna huvudsakligen deskriptivt, men resultaten antyder att framför allt de konvexa indikatorerna kan vara ett lovande verktyg för att identifiera stressperioder och regimskiften i prisdata. En mer fullständig studie av prognosförmåga och praktisk användbarhet lämnas som ett naturligt ämne för framtida arbete.

## Referenser

Luque, B., Lacasa, L., Ballesteros, F., & Luque, J. (2010). *Horizontal visibility graphs: exact results for random time series*. arXiv:1002.4526.

Hämtad från <https://arxiv.org/pdf/1002.4526>.

Lacasa, L., & Flanagan, R. (2015). *Time reversibility from visibility graphs of non-stationary processes*. arXiv:1510.01318.

Hämtad från <https://arxiv.org/pdf/1510.01318>.

Liu, R., & Chen, Y. (2020). *Analysis of Stock Price Motion Asymmetry via Visibility-Graph Algorithm*. *Frontiers in Physics*.

Hämtad från <https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2020.539521/full>.

Ott, E. (2002). *Chaos in Dynamical Systems* (2nd ed.). Cambridge University Press.

Hämtad från [https://www.biodyn.ro/course/literatura/Nonlinear\\_Dynamics\\_and\\_Chaos\\_2018\\_Steven\\_H.\\_Strogatz.pdf](https://www.biodyn.ro/course/literatura/Nonlinear_Dynamics_and_Chaos_2018_Steven_H._Strogatz.pdf).

## Generative AI disclosure

I arbetet har generativ AI (ChatGPT) använts som ett stöd vid skrivprocessen. Modellen har hjälpt till att:

- formulera textavsnitt i en mer akademisk stil,
- omformulera och strukturera vissa delar av språket,
- konvertera text till  $\text{\LaTeX}$  samt förbättra typografi, och
- ge stöd vid optimering av Python-kod för beräkning av indikatorerna.

Allt innehåll har granskats, verifierats och bearbetats av författaren, som ensam ansvarar för uppsatsens slutliga utformning och resultat.

Motsägelsefulla tolkningar av konkava och konvexa indikatorer:

Den korrekta tolkningen är den som ges i ”4.4 Konkavitet, konvexitet och accelererande rörelser”.

Alltså ej den som ges i ”4.4 Konvexa indikatorer via Invisibility Graph”, eller den som ges i ”5.3 Konkava och konvexa årsindikatorer för SPY”.

Alltså:

Konvex + växande → accelererande uppgång.

Konvex + minskande → avtagande nedgång.

Konkav + växande → avtagande uppgång.

Konkav + minskande → accelererande nedgång

Följande del från ”5.5 Diskussion av indikatorer” kan också alltså bortses ifrån:

Sammantaget visar analyserna att visibility- och invisibility-baserade indikatorer kan beskriva den välkända asymmetrin i aktiemarknaden: uppgångar tenderar att vara långsamma och relativt konkava, medan nedgångar ofta är snabbare, mer konvexa och klustrade i tiden.