

# Análisis no Markoviano del precio de acciones mediante aprendizaje de Sato-Crutchfield

Ricardo Y. Díaz-Bonifaz  
Facultad de Ciencias, UNAM, México.  
Junio 2022

En este trabajo se propone un juego denominado *apostador de bolsa* que pretende medir la capacidad de predecir los cambios en el precio de acciones. Se propone un jugador cuya estrategia evoluciona de forma no Markoviana al introducir de forma paramétrica el proceso de aprendizaje de Sato-Crutchfield. Utilizando datos de los precios de las acciones de Disney y Google se determinaron parámetros que optimizan la ganancia total promedio y reducen el riesgo de pérdida, obteniendo ganancias superiores las que se obtienen con una aproximación Markoviana. Se observó que la combinación de memoria y aleatoriedad que optimiza las ganancias varía para diferentes acciones.

## I. INTRODUCCIÓN

La bolsa de valores corresponde a uno de los sistemas complejos más diversamente estudiados en la actualidad, sumando esfuerzos de profesionales como economistas, matemáticos y físicos. Una buena descripción probabilística de los cambios en los precios de las acciones puede traducirse en nuevas estrategias de inversión que maximicen la ganancia y reduzcan el riesgo de pérdida. Se ha observado que los precios de las acciones en la bolsa son susceptibles a cambios abruptos, esporádicos y difíciles de predecir. Por lo anterior, los precios de las acciones son asociados con distribuciones de colas pesadas como la distribución de Pareto, así como con procesos no Markovianos con memoria de largo alcance [1, 2]. Entre los esfuerzos por hacer predicciones en la bolsa de valores con objetivos de inversión se encuentra el análisis en términos de vuelos de Levy, el análisis de similaridad en redes y la detección de comunidades [3–5].

Una aproximación comúnmente usada en economía es la que proporciona la teoría de juegos. En un juego se tienen dos o más jugadores, cada uno de ellos con un conjunto (catálogo) de acciones (jugadas). En una ronda del juego, cada jugador selecciona una jugada de su catálogo. Cada jugador tiene asociada una matriz de pagos que indica la recompensa a obtener dada su elección y la de todos los demás jugadores. Cuando en un juego se tienen dos jugadores con la misma matriz de pagos se dice que se tiene un juego simétrico [6].

Para entender mejor un juego con dos jugadores, consideremos la Fig. 1. En ella, se observan dos jugadores (A y B), cada uno con dos diferentes posibles jugadas (no necesariamente iguales). Consideraremos la matriz de pagos del jugador A dada por la ecuación 1

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \quad (1)$$

La matriz en la ecuación 1 indica que si en una ronda el jugador A selecciona la jugada  $i$ , mientras que el jugador B selecciona la jugada  $j$ , entonces el pago del jugador A será  $a_{ij}$ .

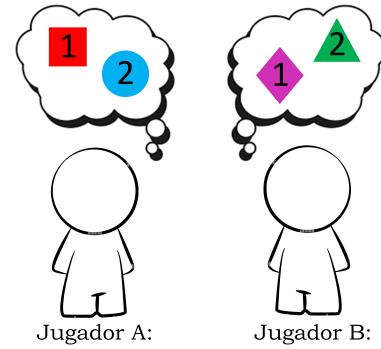


Figura 1. Caricatura de un juego asimétrico con dos jugadores.

Con el objetivo de maximizar sus ganancias, cada jugador puede asignar probabilidades a cada una de las acciones que puede llevar a cabo en una ronda. A este conjunto de probabilidades se le denomina una *estrategia mixta* [7]. Como conjunto, una estrategia mixta puede escribirse de la forma

$$X^A = \left\{ x_i^A \middle| 0 < x_i^A < 1, \sum_{i=1}^N x_i^A = 1 \right\}, \quad (2)$$

donde  $x_i^A$  representa la probabilidad de que el jugador A seleccione la jugada  $i$ . En teoría de juegos, se tiene que cuando todos los jugadores tienen conocimiento preciso sobre las acciones de sus contrincantes y actúan racionalmente, se alcanza un estado estacionario en el espacio de estrategias conocido como *equilibrio de Nash* [7]. Bajo estas consideraciones, ningún jugador puede aumentar sus ganancias modificando sus acciones individualmente sin considerar las acciones de sus contrincantes.

Una forma en la que se ha estudiado la evolución de las estrategias es mediante procesos de aprendizaje. En particular, es posible modelar la dinámica de estrategias al considerar poblaciones de ideas [6]. Esta aproximación consiste en considerar que cada jugador tiene un número finito de ideas, cada una de ellas correspondiente a alguna de las acciones que puede realizar en el juego. Si se

tiene un jugador cuya población de ideas es de tamaño  $N_A$ , con  $M_A$  posibles acciones, entonces cada una de las  $M_A$  posibles acciones estará representada  $n_i$  veces en la población. Luego, se tiene que

$$N_A = \sum_{i=1}^{M_A} n_i. \quad (3)$$

En estos términos, una estrategia está dada en términos de la representación de una acción en la población, es decir, la probabilidad  $x_i^A$  de que el jugador seleccione la acción  $i$  estará dada como

$$x_i^A = \frac{n_i}{N_A}. \quad (4)$$

La evolución de la estrategia puede obtenerse a partir de un proceso de replicación. Este proceso consiste en reemplazar una de las acciones por otra con una probabilidad que dependerá de la matriz de pagos y de la elección de su contrincante en una ronda dada. En poblaciones finitas de ideas es posible observar fenómenos de fijación (donde la estrategia llega a un punto estacionario) y extinción de ideas [6].

En el límite de poblaciones infinitas ( $N \rightarrow \infty$ ), el proceso de replicación de ideas puede expresarse en las ecuaciones de evolución de Sato-Crutchfield [6]. Este aprendizaje está basado en el aprendizaje por reforzamiento y tiene además una componente de memoria cuyo alcance puede ajustarse de forma paramétrica. La estrategia se calcula a partir de la expresión

$$x_i^A(t) = \frac{e^{\Gamma Q_i^A(t)}}{\sum_{j=1}^M e^{\Gamma Q_j^A(t)}}. \quad (5)$$

En la ecuación 5  $x_i^A(t)$  representa la probabilidad de que el jugador  $A$  seleccione la acción  $i$  al tiempo  $t$ . El parámetro  $\Gamma \in [0, \infty]$  corresponde a una *sensibilidad de aprendizaje*. El término  $Q_i^A$  se conoce como la *atracción* hacia la acción  $i$ . Estas atracciones se calculan en cada ronda como

$$Q_i^A(t+1) = (1 - \alpha)Q_i^A(t) + \sum_{j=1}^{M_B} a_{ij}x_j^B, \quad (6)$$

donde  $\alpha \in [0, 1]$  corresponde a un parámetro de memoria. Para  $\alpha = 0$  se tiene memoria de todas las rondas anteriores, mientras que para  $\alpha = 1$  no se tiene memoria. El segundo término de la ecuación 6 corresponde a la ganancia esperada al elegir la acción  $i$  dadas las posibles jugadas del jugador  $B$ . En la suma,  $M_B$  es el número de jugadas de las que dispone el jugador  $B$  y  $a_{ij}$  son los elementos de la matriz de pagos del jugador  $A$ . Una observación interesante es que la expresión 5 tiene la forma

de la probabilidad de un ensamble canónico en donde  $\Gamma$  sería proporcional al inverso de la temperatura.

En este trabajo se propone un juego para medir la capacidad de predecir cuándo el precio de las acciones de una empresa va a subir o bajar. En la sección II se define el juego, la implementación del algoritmo de Sato-Crutchfield (SC), así como diferentes estrategias del juego que compiten con el algoritmo SC. La sección IV contiene los resultados obtenidos tras el análisis de los precios de las acciones de Disney y Google. La sección V contiene una discusión de los resultados y de posibles acciones para mejorar los resultados obtenidos.

## II. JUEGO DEL APOSTADOR DE BOLSA

Con el objetivo de estudiar la capacidad de predecir alzas o caídas en los precios de diferentes acciones, se propone un juego formado por dos jugadores. El jugador A, al cual nos referiremos como *el apostador*, puede realizar cualquiera de las siguientes jugadas

### Jugadas del apostador:

1. **Compra-venta:** El apostador compra una acción al inicio del día al precio de apertura y la vende al final del día al precio de cierre.
2. **Venta-compra:** El apostador vende una acción al inicio del día, recibiendo el precio de apertura, y la vuelve a comprar al final del día al precio de cierre.

El segundo jugador, al que nos referiremos como *la bolsa*, tiene a su vez dos posibles jugadas:

### Jugadas de la bolsa:

1. Subir el precio de una acción.
2. Bajar el precio de una acción.

El monto que la acción sube o baja no se considera relevante para la definición del juego y puede considerarse como aleatorio.

Cada día corresponde a una ronda del juego, en la que el apostador decide realizar compra-venta o venta-compra y al final del día recibe su recompensa dependiendo de la jugada que haya realizado la bolsa (subir o bajar). La diferencia de precio de una acción a entre el inicio y el final de un día se expresa como

$$\Delta(t) = P_{cr}(t) - P_{ap}(t), \quad (7)$$

donde  $t$  indica el día,  $P_{cr}$  y  $P_{ap}$  son respectivamente los precios de apertura y cierre de la acción en ese día. Luego, la matriz de pagos puede definirse en términos de  $\Delta$  como

$$\mathbf{A}(t) = |\Delta(t)| \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (8)$$

La matriz de la ecuación 8 indica que la ganancia del apostador al realizar la jugada 1 (compra-venta) es igual al cambio en el precio de la acción, teniendo ganancias positivas si la acción subió y negativas si esta bajó. Análogamente, si el apostador elige la jugada 2 (venta-compra) su ganancia será igual a el cambio de precio con signo opuesto, de esta forma se tienen ganancias positivas si el precio de la acción bajo y negativas si esta subió.

La matriz de pagos del jugador  $B$  se desconoce dado que no es claro en qué se beneficia o afecta la bolsa de valores de las diferentes jugadas del apostador. Así mismo, asumimos que las decisiones del apostador tienen un efecto despreciable en la estrategia de la bolsa. Esto adquiere sentido si se considera que la compra o venta de una única acción no es de interés al establecer los precios de las acciones. Lo anterior se traduce en que la estrategia del jugador  $B$ , dada por las probabilidades  $x_i^B$ , es independiente de las decisiones del apostador.

El análisis del juego se realiza utilizando datos reales del precio de acciones de dos compañías, que representan al jugador  $B$ . Las decisiones del apostador serán obtenidas de simulaciones Monte Carlo (MC). Con el objetivo de hacer un análisis comparativo se consideran 4 tipos de apostadores en el juego: el uniforme, el aleatorio, el Markoviano y el SC. Cada uno de ellos se diferencia por la forma en la que elige su estrategia. A continuación se elabora sobre cada uno de ellos.

### A. Apostador uniforme

Este inversionista asigna probabilidad  $x_i^A = 1$  a alguna de las dos jugadas durante todo el juego. Al seguir esta estrategia por todo un periodo de tiempo se tiene que las ganancias totales tras elegir siempre la jugada 1 serán las mismas que al elegir siempre la jugada 2, pero con signo opuesto. Luego, esta estrategia garantiza que se obtengan ganancias no negativas si se elige la jugada correcta. En el resto del presente trabajo se considerará la ganancia total del apostador uniforme como la ganancia positiva independientemente de a cuál de las dos jugadas corresponda.

### B. Apostador aleatorio

En este caso se tiene una estrategia mixta constante que corresponde a asignar la misma probabilidad a ambas jugadas, es decir,

$$x_i^A = \frac{1}{2}, \quad (9)$$

para  $i = 1, 2$ . Las ganancias totales promedio de este apostador se utilizarán como referencia para contrastar la aleatoriedad de las ganancias del resto de los apostadores.

### C. Apostador Markoviano

La estrategia del apostador Markoviano está basada en la probabilidad de que el precio de la acción suba al día siguiente de que el precio haya subido  $T(\uparrow | \uparrow)$  o bajado  $T(\uparrow | \downarrow)$ . Una forma de calcular estas probabilidades es en términos de los contadores  $T_{ij}$  que indican el número de veces que el precio de la acción subió ( $j = \uparrow$ ) o bajó ( $j = \downarrow$ ) al día siguiente de haber subido ( $i = \uparrow$ ) o bajado ( $i = \downarrow$ ). Luego, las probabilidades de transición se obtienen como

$$T(\uparrow | \uparrow) = \frac{T_{\uparrow\uparrow}}{T_{\uparrow\uparrow} + T_{\uparrow\downarrow}} \quad (10)$$

$$T(\uparrow | \downarrow) = \frac{T_{\downarrow\uparrow}}{T_{\downarrow\downarrow} + T_{\downarrow\uparrow}}. \quad (11)$$

El apostador Markoviano utiliza las probabilidades de transición dadas por 10 y 11 para definir su estrategia. Este apostador elegirá vender-comprar con probabilidad  $x_1^A = T(\uparrow | \uparrow)$  al día siguiente de que el precio de la acción haya subido y con probabilidad  $x_1^A = T(\uparrow | \downarrow)$  al día siguiente de que ésta haya bajado.

Para inicializar el juego, no se realiza ninguna apuesta con los datos de los primeros 100 días. Éstos datos se utilizan para calcular las probabilidades de transición iniciales. Posteriormente, éstas se irán actualizando en cada ronda, resultando en una estrategia que evoluciona en el tiempo.

### D. Apostador SC

Definimos al apostador SC como aquel que modifica su estrategia mixta en cada ronda basado en las ecuaciones 5 y 6. Considerando que las decisiones de la bolsa corresponden a datos reales que no pueden modificarse, la estrategia  $\{x_i^B\}$  no puede evolucionar mediante el proceso de aprendizaje. No obstante, para realizar el cálculo de las atracciones  $Q_i^A$  es necesario conocer la estrategia del jugador  $B$ . Para solucionar esta situación se considerará la estrategia percibida por el jugador  $A$ . Una forma de obtener esta estrategia percibida es calculándola a partir de las probabilidades de transición Markovianas definidas en las ecuaciones 10 y 11. En estos términos, la probabilidad de que el precio de una acción suba estará dada por  $x_1^B = T(\uparrow | \uparrow)$  si el precio de la acción subió el día anterior, o bien, por  $x_1^B = T(\uparrow | \downarrow)$  si bajó. Al igual que con el apostador Markoviano, los datos de los primeros 100 días se utilizan para inicializar la estrategia  $\{x_i^B\}$ , misma que irá evolucionando en cada ronda siguiendo las ecuaciones 10 y 11.

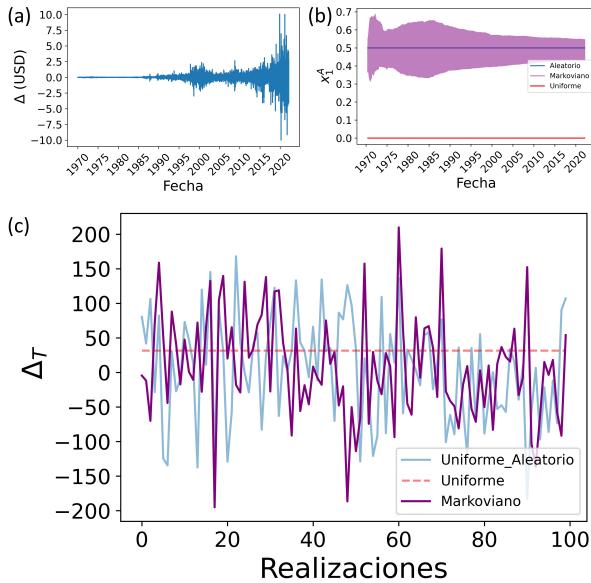
Otro aspecto a considerar es la elección de los parámetros  $\alpha$  y  $\Gamma$ , los cuales se tomarán como parámetros con el

objetivo de reproducir los resultados de los otros apostadores y con fines de optimización de ganancias.

### III. RESULTADOS

El primer análisis del juego se realizó utilizando el registro diario (de lunes a viernes) de los precios de las acciones de Disney desde enero de 1970 y hasta febrero de 2022. Para cada día, se calculó la diferencia de precios  $\Delta$ , obteniendo la gráfica 2a. En ella podemos observar que en las mayores variaciones se tienen para períodos recientes. En la figura 2b se puede apreciar el valor de la probabilidad  $x_i^A$  de realizar una acción de compra-venta. Se observa que la estrategia del apostador Markoviano presenta variaciones que van disminuyendo con el paso del tiempo. Esto es esperado ya que al avanzar el tiempo las actualizaciones en la probabilidad de transición se hacen cada vez menos relevantes.

Tras realizar 100 realizaciones de la simulación MC se calculó la ganancia total  $\Delta_T$  en todo el periodo a partir de la suma de las ganancias diarias de cada apostador. En la figura 2c se muestra el valor de  $\Delta_T$  en cada realización para cada los apostadores uniforme, aleatorio y Markoviano.



Como un caso límite con el cual comparar, se realizó la simulación MC del apostador SC para  $\Gamma = 0$ . De la ecuación 5 se tiene que en ese caso la estrategia mixta no evoluciona, sino que asigna probabilidad  $x_i^A = 0.5$  a ambas acciones. En la figura 3 podemos observar que la es-

trategia es la igual a la del apostador uniforme, mientras que la ganancia total también se comporta de manera similar a la del apostador Markoviano.

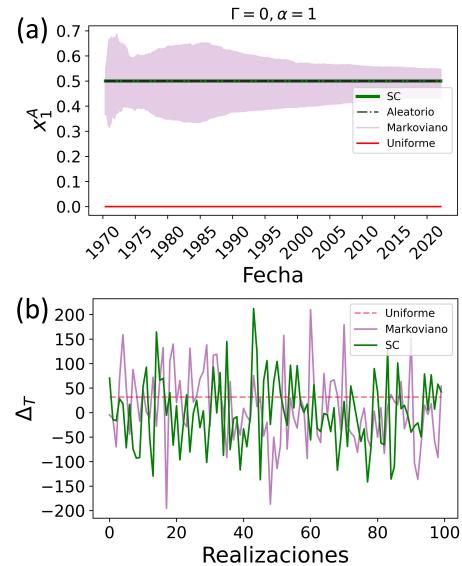


Figura 3. (a) Simulación MC del espacio de estrategias de los 4 apostadores en términos de la probabilidad  $x_1^A$ . (b) Ganancia total de 3 de los apostadores en 100 realizaciones de la simulación MC

Una vez explorado el caso límite, queda la pregunta sobre qué parámetros son óptimos para maximizar  $\Delta_T$ . Esto se exploró realizando simulaciones para diferentes combinaciones de  $\Gamma$  y  $\alpha$  como se observa en las Figs. 4a y 4b, donde se muestran los valores del promedio  $\langle \Delta_T \rangle$  y su desviación estándar  $\sigma_\Delta$  para diferentes combinaciones de  $\alpha$  y  $\Gamma$ . Cada pixel corresponde al promedio de 50 realizaciones. Podemos observar además que a mayor valor de  $\Gamma$  se tiene una menor  $\sigma_\Delta$ . Este es un resultado esperado dado que la aleatoriedad en la decisión disminuye al aumentar  $\Gamma$ . Por otro lado, no se tiene un comportamiento monótono de la ganancia con respecto de ninguno de los dos parámetros. En cambio, se tienen diferentes combinaciones de  $\Gamma$  y  $\alpha$  que maximizan la  $\Delta_T$ . Nótese además que las diferentes combinaciones de  $\langle \Delta_T \rangle$  y  $\sigma_\Delta$  se traducen en diferentes perfiles de apuesta. Por ejemplo, elegir una combinación con  $\langle \Delta_T \rangle$  y  $\sigma_\Delta$  intermedios implica la posibilidad de obtener ganancias con riesgo moderado. En cambio, uno podría elegir una combinación con  $\langle \Delta_T \rangle$  y  $\sigma_\Delta$  más elevados que hacen posible obtener una mayor ganancia pero con el riesgo de obtener una ganancia menor a la combinación de ganancia moderada.

En la figura 4d podemos apreciar  $\Delta_T$  en 100 realizaciones para una elección de  $\langle \Delta_T \rangle$  alto y  $\sigma_\Delta$  moderado. Se observa que las menores ganancias de esta combinación son mayores a las ganancias del apostador Markoviano. Por lo anterior se concluye que el apostador SC con parámetros adecuados puede obtener ganancias su-

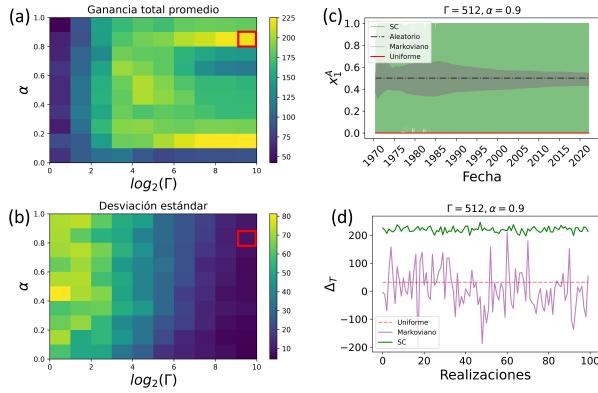


Figura 4. (a) y (b) Ganancia total promedio y desviación estándar tras 50 realizaciones para diferentes combinaciones de  $\alpha$  y  $\Gamma$ . (c) Simulación de la estrategia ( $x_1^A$ ) del apostador SC para los parámetros optimizados marcados en rojo en (a) y (b). (d) Ganancia total en 100 realizaciones para los diferentes apostadores con parámetros optimizados para el apostador SC.

periores a las de sus competidores. Obsérvese en la Fig. 4c que las variaciones en la estrategia del apostador SC son tales que en cada ronda se asigna una probabilidad muy cercana a uno a alguna de las dos jugadas, reduciendo el papel de la aleatoredad. Por otro lado, el valor de  $\alpha$  es tal que la memoria es de corto alcance.

Los parámetros que optimizan las ganancias y reducen el riesgo para acciones de Disney son en general diferentes a los que optimizan las ganancias del apostador para alguna otra empresa. Para mostrar esto se analizaron los precios diarios de las acciones de Google de marzo de 2014 a febrero de 2022. Se analizó  $\Delta_T$  y  $\sigma_\Delta$  de la misma forma en que se hizo con las acciones de Disney tomando las mismas combinaciones de  $\Gamma$  y  $\alpha$ . Los resultados obtenidos se observan en las gráficas 5a y 5b.

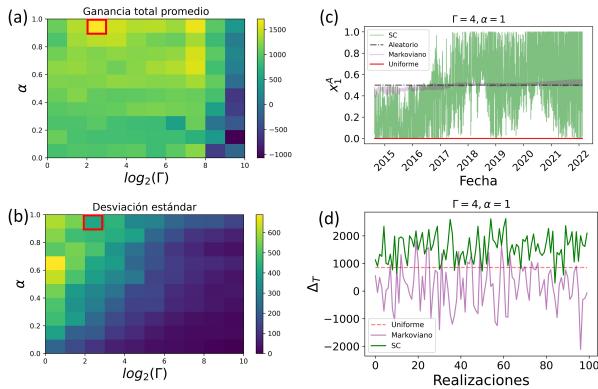


Figura 5. (a), (b), (c) y (d) análogos a la Fig. 4. En esta ocasión considerando el precio de las acciones de Google.

Podemos apreciar en las Figs. 5a y 5b que los parámetros que optimizan las ganancias del apostador SC son diferentes a los que optimizan las acciones de Disney. En este caso, un posible valor óptimo es  $\Gamma = 4$  para  $\alpha = 1$ . Por un lado,  $\alpha = 1$  indica que se trata de un límite Markoviano (sin memoria), mientras que el valor relativamente bajo de  $\Gamma$  sugiere una participación relevante de la aleatoredad en contraste con la estrategia óptima para Disney (4c) que tenía un carácter más determinista. La aleatoredad que optimiza a Google se puede observar en la Fig. 5c, dado que la estrategia consiste en asignar frecuentemente probabilidades alejadas de 1 a cualquiera de las dos jugadas.

#### IV. CONCLUSIONES

El juego del apostador de bolsa permite medir la capacidad de una estrategia de anticipar cambios en el precio de acciones. Una mayor ganancia en este juego implica que se acertó en el cambio de precio de una acción en más ocasiones o en días con cambios más importantes. Entre las estrategias estudiadas, se observó que el jugador que utiliza el proceso de aprendizaje de Sato-Crutchfield es capaz de obtener la mayor ganancia total si se utilizan los parámetros correctos. A la vez, las diferentes combinaciones de  $\Gamma$  y  $\alpha$  dan origen a diferentes perfiles de apuesta con diferentes ganancias promedio y riesgos. Se observó además que diferentes acciones se benefician de diferentes combinaciones de memoria y aleatoredad.

Por lo anterior, el algoritmo SC puede ser usado en el futuro para definir portafolios de inversión. Si bien el juego del apostador de bolsa permite obtener resultados positivos con este algoritmo, su desempeño en escenarios de inversión más realistas con montos de inversión variables e interés compuesto están aún por estudiarse.

[1] H. F. Coronel-Brizio and A. R. Hernández-Montoya, On fitting the Pareto-Levy distribution to stock market index

data: Selecting a suitable cutoff value, Physica A: Sta-

- tistical Mechanics and its Applications **354**, 437 (2005), arXiv:0411161 [cond-mat].
- [2] F. R. Brown, D. Pravica, and M. Bier, Self-similarity and non-Markovian behavior in traded stock volumes, European Physical Journal B **88**, 1 (2015).
  - [3] F. Kia, G. Jeney, and J. Levendovszky, Loss-minimal Algorithmic Trading Based on Levy Processes, TEM Journal **3**, 210 (2014).
  - [4] X. Yan, H. Yang, Z. Yu, and S. Zhang, A Network View of Portfolio Optimization Using Fundamental Information, Frontiers in Physics **9**, 1 (2021).
  - [5] L. Zhao, C. Wang, G.-J. Wang, H. E. Stanley, and L. Chen, Community detection and portfolio optimization, (2021), arXiv:2112.13383.
  - [6] R. Nicole, P. Sollich, and T. Galla, Stochastic evolution in populations of ideas, Scientific Reports **7**, 1 (2017), arXiv:1609.04443.
  - [7] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*, 1st ed. (The MIT Press, Cambridge, Massachusetts, 1994).