



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**POSGRADO EN CIENCIAS FÍSICAS**  
**INSTITUTO DE CIENCIAS FÍSICAS**

**MAPEO DE GUHR-KAELBER APLICADO A MATRICES DE  
CORRELACIÓN SINGULARES DE DOS MERCADOS  
FINANCIEROS**

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
**MAESTRO EN CIENCIAS (FÍSICA)**

PRESENTA:  
ELSA SUSANA OCHOA GONZÁLEZ

**TUTOR PRINCIPAL:**  
DR. THOMAS HENRY SELIGMAN SCHURCH  
INSTITUTO DE CIENCIAS FÍSICAS

**MIEMBROS DEL COMITÉ TUTOR:**  
DR. FRANCOIS ALAIN LEYVRAZ WALTZ  
INSTITUTO DE CIENCIAS FÍSICAS  
DR. CARLOS FRANCISCO PINEDA ZORRILLA  
INSTITUTO DE FÍSICA

MÉXICO D.F., NOVIEMBRE 2017  
CD.MX.



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*Dedicado a  
mi familia*



# Agradecimientos

Quiero agradecer a mis padres, Albino y Elsa, por el gran apoyo constante que me brindaron durante este proceso. Sin ellos no hubiera sido posible este logro.

A mi asesor Thomas Seligman por todo el apoyo, por siempre estar con la mejor disposición de ayudar, motivándome e inspirándome.

Al Dr. Anirban Chakraborti, con el que estuve trabajando al último en este proyecto.

A mis compañeros y amigos (José, Jorge, Lorenia, Melissa, Cecilia, Pasos, Adrian, Paulino, Roberto, Hector, Carlos y más), que además de hacer más amena mi estancia durante la maestría, siempre estuvieron accesibles para discutir cualquier duda.

A las Ochoa (Lupita, Beatriz, Ale y Carmen), por todo el apoyo y cariño.

A Alejandra Ochoa por ser una mujer que me inspira, por su tiempo y apoyo, que me ayudó a crecer académicamente.

A los profesores/investigadores por sus comentarios constructivos y tiempo invertido en este trabajo.

A CONACYT por la beca de maestría y Proyecto Fronteras 201, a PAPIIT-DGAPA-UNAM proyecto IG100516, y al Instituto de Ciencias Físicas de la UNAM, por darme la oportunidad de realizar este posgrado.



# Resumen

El estudio de las series de tiempo financieras es paradigmático para la econofísica y ha crecido significativamente debido a la disponibilidad de datos, y a la posibilidad de acceder y manejar estos datos muy rápidamente. Las matrices de correlación son una herramienta muy útil para el análisis de series de tiempo y se han introducido matrices aleatorias como hipótesis nula, donde los ensembles de Wishart son los apropiados.

En el presente trabajo se aplica el método de mapeo de Guhr-Kaelber o mapeo de potencias a matrices de correlación singulares, para analizar series de tiempo cortas con el objetivo de obtener un espectro de valores propios más amplio y así poder hacer un mejor análisis estadístico de ellos. Utilizando los datos del índice Standard & Poor's 500 (S&P500) a lo largo del período 1992-2014, se calculan matrices de correlación  $C$  de los rendimientos en una ventana de tiempo  $T = 44$ , creadas de los rendimientos de un día de bolsa al siguiente día de un conjunto de  $N = 293$  acciones. Entonces, se les aplica el mapeo de potencias, que consiste en elevar a una potencia  $q = 1 + \epsilon$ ,  $1 \gg \epsilon > 0$  al valor absoluto de los elementos de la matriz de correlación conservando el signo.

La no linealidad de este método rompe la degeneración de los valores propios cero que son la mayoría. El espectro resultante está bien separado del espectro distinto de cero original y lo vamos a nombrar *espectro emergente*. El espectro original sufre solo distorsiones mínimas si  $\epsilon$  es suficientemente pequeño. Se analiza la sensibilidad del espectro emergente a la correlación promedio de la matriz. Los resultados se comparan con las matrices de Wishart sin correlación o con correlación constante igual a la correlación promedio de la matriz observada. Encontramos que el espectro emergente es bastante sensible a la correlación promedio de la matriz, con un comportamiento particular en épocas de crisis en el mercado financiero o cercanas a ellas. También observamos que el número de valores propios negativos está determinado en gran medida por la correlación promedio de la matriz. Además el valor propio menor presenta un comportamiento característico en las diferentes crisis.



# Índice general

Resumen . . . . .	5
<b>1. Introducción</b>	<b>9</b>
<b>2. Conceptos fundamentales</b>	<b>13</b>
2.1. Conceptos financieros . . . . .	13
2.1.1. Sistema financiero . . . . .	13
2.1.2. Mercado primario vs Mercado secundario . . . . .	13
2.1.3. Índice financiero . . . . .	14
2.1.4. Series de tiempo financieras . . . . .	15
2.2. Matrices de correlación y matrices de correlación aleatorias . .	15
2.2.1. Matrices de correlación . . . . .	15
2.2.2. Ensemble de Wishart . . . . .	17
2.2.3. Ensemble de Wishart correlacionado . . . . .	18
2.3. Mapeo de potencias . . . . .	19
2.3.1. Metodología . . . . .	19
2.3.2. Mapeo de potencias para matrices de correlación sin- gulares . . . . .	19
<b>3. Resumen del método y aplicación</b>	<b>25</b>
<b>4. Resultados</b>	<b>27</b>
4.1. Distribución de valores propios . . . . .	27
4.1.1. Momentos del espectro emergente . . . . .	32
4.1.2. Valores propios extremos del espectro emergente y bulk	38
4.2. Número de valores negativos . . . . .	39
<b>5. Conclusiones</b>	<b>45</b>
<b>A. Tabla de compañías</b>	<b>47</b>
<b>Bibliografía</b>	<b>53</b>



# Capítulo 1

## Introducción

El análisis de matrices de correlación para series de tiempo es una herramienta muy útil en el estudio de sistemas complejos, como lo son los mercados financieros. En estos últimos se emplean con el fin de comprender la estructura y dinámica del mercado. Se ha encontrado en el estudio de los rendimientos varias características estadísticas comunes para diferentes activos, mercados y períodos de tiempo. En los sistemas financieros típicamente se usan matrices cuyos elementos son coeficientes de correlación de Pearson [16].

Uno de los análisis que se ha estudiado ampliamente es la aplicación del método de matrices aleatorias con o sin correlaciones como suposición nula en el estudio de las matrices de correlación financieras, esto con el propósito de entender las interrelaciones complejas entre los elementos de una matriz de correlación. Este análisis ha mostrado que las matrices que encontramos para mercados financieros presentan algunas propiedades universales de la teoría de matrices aleatorias [2].

Al analizar series de tiempo financieras nos enfrentamos con series que no son estacionarias, lo que las hace difícil de analizar. Una forma de lidiar con esto es tomar series de tiempo cortas, lo que permite asumir cuasi-estacionariedad sobre el intervalo de tiempo pero produce ruido en nuestro análisis. Esta dificultad ha inspirado el desarrollo de varios métodos para reducir el ruido [6, 20], uno de ellos es el método de *mapeo de Guhr-Kaelber* o *mapeo de potencias*<sup>1</sup>, que fue desarrollado por Guhr y Kaelber [7]. Dicho método consiste en elevar a una potencia  $q = 1 + \epsilon$ ,  $\epsilon > 0$  el valor absoluto de los elementos de matriz de correlación conservando el signo. Otro problema que tenemos con series de tiempo cortas, es que la matriz de correlación tiende a volverse singular, i.e. tenemos cada vez menos valores propios diferentes

---

<sup>1</sup>Del inglés power mapping.

de cero a medida que acortamos más las series de tiempo. Aunque el objetivo inicial de mapeo de potencias fue la reducción de ruido, en ref. [23] ha sido mostrada la efectividad de este método para romper la degeneración de valores propios iguales a cero, con potencias muy cercanas a 1, e.g  $q=1.005$ . Así se puede tener un mejor análisis estadístico de la distribución de valores propios. El espectro obtenido está bien separado del espectro original, si  $\epsilon$  es suficientemente pequeño, y ha sido llamado *espectro emergente* [23].

Anteriormente se ha visto que el espectro emergente para el modelo de Ising en dos dimensiones es sensible a las correlaciones que aparecen cerca de la temperatura crítica de este modelo, si usamos la dinámica de Metropolis [22]. Como se sabe que surgimientos fuertes del promedio de la matriz de correlación se relacionan con momentos dramáticos en los mercados financieros, esto sugiere que también el espectro emergente es susceptible a cambios que se asemejan a transiciones de fase en sistemas estacionarios.

Los datos usados en este trabajo se tomaron del índice financiero Standard & Poor's 500, nos referiremos a él como S&P500, por ser un referente al mercado de Estados Unidos, ya que contiene las 500 acciones más importantes de dos mercados financieros, que son: la bolsa NYSE (bolsa de valores de nueva york, por sus siglas en inglés) y Nasdaq (asociación nacional de distribuidores de valores cotización automatizada, por sus siglas en inglés). Tomamos de un conjunto de compañías distintas el precio de cierre de sus acciones a lo largo de un período de 23 años. Las compañías fueron seleccionadas por haber existido a lo largo del tiempo en condiciones económicas similares dentro del S&P500. Ante estas consideraciones vamos a estudiar 293 acciones del S&P500 para el período 1992-2014 usando el mapeo de potencias y enfocándonos principalmente en el espectro emergente, sus propiedades y su relación con el promedio de los elementos de la matriz de correlación.

En el capítulo 2 se introducen conceptos fundamentales para sistemas financieros, matrices de correlación y los modelos de matrices aleatorias, así como también se introduce al método de mapeo de potencias para matrices singulares.

En el capítulo 3 describimos nuestro método de análisis, el cual consiste en crear matrices de correlación  $C$  a partir de un conjunto de  $N$  series de tiempos de los *rendimientos* del mercado, obtenidos de la diferencia de precios de cierre diarios sucesivos de las acciones normalizados por el precio actual. Hacemos esto sobre pequeñas ventanas de tiempo (44 días de bolsa) a lo largo del período de 23 años de 1992-2014. Aplicamos entonces el método de mapeo de potencias con una potencia cerca de 1 e investigamos las propiedades del espectro emergente.

En el capítulo 4 se describen los resultados obtenidos. Encontramos que el espectro emergente es bastante sensible a la correlación promedio de la

matriz, con un comportamiento particular en situaciones críticas en el sistema financiero o cercanas a ellas.

Por último, en el capítulo 5 se muestran las conclusiones del trabajo desarrollado.



# Capítulo 2

## Conceptos fundamentales

### 2.1. Conceptos financieros

#### 2.1.1. Sistema financiero

Un sistema financiero es un sistema que permite el intercambio de fondos entre prestamistas, inversores y prestatarios. Se forma con *instituciones financieras* y *mercados financieros*. En los mercados financieros se compran y venden instrumentos financieros como acciones, bonos, opciones, futuros, etc. Las instituciones financieras son firmas como bancos comerciales, fondos de inversión, fondos de pensión, etc. que proveen de servicios financieros a sus consumidores e invierten sus fondos en instrumentos financieros [10]. Existen dos mecanismos por donde los fondos fluyen a través del sistema financiero, uno es el financiamiento directo, en el cual los fondos pasan directamente a los mercados financieros y el otro es el financiamiento indirecto, donde los fondos pasan primero por un intermediarios que comunica a los inversionistas con los solicitantes de fondos. El propósito de ambos mecanismos es comunicar de manera regular, eficiente y rentable entre ambas partes.

#### 2.1.2. Mercado primario vs Mercado secundario

Cuando se crea un instrumento financiero se vende por primera vez en el *mercado primario*. Por ejemplo, la venta de *acciones* recientemente creadas. Las acciones representan el valor de una de las fracciones iguales en que se divide el valor de los bienes de una compañía. En el *mercado secundario* se realizan la compra y venta de instrumentos financieros que se han comercializado previamente en un mercado primario. Las compañías no reciben nuevos fondos cuando se realiza una venta de algún instrumento en el mercado secundario, pero sí el individuo que lo vende. Los mercados secundarios hacen

mas rápida y fácil la venta de instrumentos financieros para adquirir dinero, a esto se le llama *liquidez*, ésta hace al instrumento mas deseable y por lo tanto ayuda a la compañía que lo crea a venderlo más fácilmente en el mercado primario. Las condiciones del mercado secundario son las más importantes al momento de crear nuevos instrumentos debido a que determinan el precio que se le da al instrumento por primera vez en los mercados primarios, ya que los inversionistas que compran instrumentos en el mercado primario no pagarán más del precio al que piensan que pueden vender en el mercado secundario [14].

Los mercados secundarios están organizados de dos formas. La primera es una bolsa (*exchange*), las bolsas son mercados centralizados y regulados, donde los instrumentos son comerciados de manera estandarizada y transparente. NYSE, Nasdaq y la Bolsa Mexicana de Valores (BMV) son ejemplos de bolsas. Generalmente las compañías grandes eligen éste tipo de intercambio para *cotizar* y comerciar sus instrumentos. La otra forma de organizarse es en mercados *extra-bursátiles* (*Over-The-Counter OTC*), que son redes descentralizadas de "dealers" (agentes mediadores), en donde los *brókers* y *dealers* interactúan directamente entre ellos ya sea por redes computacionales o por teléfono. Ésta forma de cotizar es menos transparente ya que, a diferencia de la bolsa, los precios de postura generalmente no son publicados; a menos que se trate de acciones que también están registradas en la bolsa.

### 2.1.3. Índice financiero

Un índice financiero es una medida del valor que representa el movimiento en conjunto de varios instrumentos financieros de distintas compañías que se consideran como representativas de un mercado.

Los índices financieros más populares son el Dow Jones, que mide el valor de 30 de las compañías más grandes de los Estados Unidos y el S&P500 que contiene 500 acciones consideradas como las más representativas de las bolsas NYSE y Nasdaq. Este índice es considerado comúnmente como referencia del mercado de Estados Unidos. En el caso de México, está el Índice de Precios y Cotizaciones (IPC), el cual mide el desempeño de 35 acciones que cotizan en la Bolsa Mexicana de Valores (BMV). Otro índice financiero es el Nikkei 225 (calculado por el periódico japonés Nihon Keizai Shinbun), que es el más popular del mercado japonés.

Cada índice se calcula de distinta manera pero por lo general son medidas ponderadas por el precio de las acciones o el valor de las compañías que las expiden.

### 2.1.4. Series de tiempo financieras

Las series de tiempo son cualquier secuencia de números o datos, aunque éstos no representen una función del tiempo. Ejemplos de series de tiempo son las alturas de las mareas oceánicas, el recuento de las manchas solares, una secuencia de genes sobre un cromosoma y el valor de cierre diario de alguna acción financiera.

Las series de tiempo financieras siguen la evolución de alguna cantidad de interés económico (p. ej. el precio de un instrumento financiero), durante un período de tiempo específico. Usualmente los registros se hacen a intervalos regulares, pero en mercados financieros se registra en cada operación efectuada. Esto puede ser rastreado a corto plazo, como el promedio del precio de las transacciones de un activo cada minuto (si en un minuto dado no hay transacciones se toma las del minuto anterior) en el transcurso de un día hábil; o a largo plazo, como el precio de un instrumento financiero al cierre de cada día hábil en un lapso de siete años [1].

La mayoría de los estudios financieros utiliza *rendimientos*, en lugar de precios de instrumentos financieros. Un rendimiento es la ganancia o pérdida de una inversión sobre un período de tiempo  $\Delta\tau$ . Dos de las principales razones para usar rendimientos son: primero, porque para el promedio de los inversores, el rendimiento es un resumen completo de la oportunidad de inversión; y segundo, las series de rendimientos son más fáciles de manejar que las series de precios porque tienen propiedades estadísticas más atractivas [19], ya que las series de precios presentan variaciones sistemáticas a largo plazo.

Definimos los rendimientos como:

$$R_{\tau}^j \equiv \frac{S_{\tau+\Delta\tau}^j - S_{\tau}^j}{S_{\tau}^j},$$

donde  $S_{\tau}^j$  es el precio de la acción de la  $j$ -ésima compañía en el tiempo  $\tau$  de un mercado formado por  $N$  compañías.

## 2.2. Matrices de correlación y matrices de correlación aleatorias

### 2.2.1. Matrices de correlación

Las matrices de correlación son una herramienta importante en el análisis de series de tiempo [11, 16, 17]. Las correlaciones miden el grado de dependencia lineal entre las variables estocásticas de diferentes series de tiempo.

Este método ha ganado atención en varias ramas de la ciencia como física, biología, econofísica, ciencias atmosféricas, etc.

Definimos la matriz de correlación  $C$  como aquella formada por los coeficientes de *correlación de Pearson*:

$$C_{jk} = \frac{\langle (R_\tau^j - \eta_j)(R_\tau^k - \eta_k) \rangle_T}{\sigma_j \sigma_k},$$

donde  $\langle \cdot \rangle_T$  representa el promedio sobre una ventana de tiempo de longitud  $T$ ,  $\eta_l$  y  $\sigma_l$  son el promedio y la desviación estándar del  $l$ -ésimo rendimiento. Podemos observar que la matriz  $C$  es real y simétrica, y que  $C_{jk}$  representa la correlación cruzada entre la  $j$ -ésima y la  $k$ -ésima variable.

Debido a que las diferentes acciones tienen distintas desviaciones estándar, es conveniente redefinir las series de tiempo normalizando los rendimientos:

$$A_\tau^j = \frac{R_\tau^j - \eta_j}{\sigma_j},$$

donde  $\eta_l$  y  $\sigma_l$  son el promedio y la desviación estándar de las series de tiempo de los rendimientos de la  $l$ -ésima acción y  $A^j$  es una serie de tiempo en términos de  $\tau$ . Entonces, la matriz de correlación puede escribirse como

$$C = \frac{1}{T} AA^t,$$

donde  $A^t$  es la transpuesta de la matriz  $A$ . Notemos que por la normalización se cumple  $0 \leq |C_{jk}| \leq 1$ , y los elementos de la diagonal son  $C_{jj} = 1$ .  $AA^t$  es una matriz semi-definida positiva, es decir: no tiene valores propios negativos. Además, como  $A$  tiene rango  $\min(N, T)$  la matriz  $C$  tiene el mismo rango. Para el caso que  $T \geq N$ , la matriz es en general definida positiva, aunque existen casos que es semi-definida positiva debido a dependencia lineal entre las series de tiempo, mientras que, para  $T < N$  tendremos al menos  $N - T + 1$  valores propios iguales a cero.<sup>1</sup> En este último caso  $C$  será una matriz singular positiva semi-definida.

Las correlaciones reales están definidas para  $T \rightarrow \infty$ , el cual es un concepto más teórico o más prácticamente para  $T$  muy grande. Cuando  $T$  es finita la matriz de correlación se vuelve ruidosa. Para un análisis estadístico necesitamos una hipótesis nula y las matrices de correlación de la teoría de matrices aleatorias son los candidatos más idóneos. Los modelos de Wishart

---

<sup>1</sup>El rango es el número de renglones o columnas linealmente independiente. Si se normalizan los renglones de  $A$ , entonces, su rango y el de la matriz  $C$  es  $T - 1$ . El número de valores propios diferentes de cero es el rango de la matriz.

caracterizan una hipótesis nula para las matrices de correlación. En los modelos Wishart todas las variables son estadísticamente independiente y las series de tiempo son definidas como series de tiempo de ruido blanco, este concepto se generaliza a matrices de Wishart correlacionadas. Los modelos de Wishart se describen en las siguientes subsecciones.

### 2.2.2. Ensemble de Wishart

Inicialmente la teoría de matrices aleatorias (sus siglas en inglés RMT), se remonta a los trabajos de Wishart [28] y Cartan [4] en 1930's. Pero no fue hasta 1950's que gracias al trabajo Wigner empieza un estudio intensivo de las propiedades de RMT en conexión con física nuclear [25–27]. RMT fue ganando popularidad en varias áreas de la física y las matemáticas, como la caracterización de sistemas caóticos, conductividad en metales desordenados, y para investigar propiedades espectrales de átomos y moléculas. [3, 8, 13].

Wishart ha introducido series de tiempo de ruido blanco como hipótesis nula para identificar las correlaciones cruzadas reales subyacentes. En RMT se estudian básicamente tres ensembles de Wishart invariantes, además de los ensembles correlacionados correspondientes. Para este trabajo sólo consideraremos los ensembles ortogonales sin correlación y correlacionados de Wishart.

A continuación se describe brevemente el ensemble ortogonal sin correlación de Wishart, y el ensemble ortogonal correlacionado de Wishart se explicará en la siguiente subsección.

Una matriz de Wishart [28] esta definida como  $C = \frac{1}{T}BB^t$  donde  $B$  es una matriz de  $N \times T$ ,  $B^t$  es la transpuesta de  $B$ , y las entradas de  $B$  son variables gaussianas independientes reales con promedio cero y varianza  $\sigma^2 = 1$ . En RMT el ensemble de matrices de Wishart se conoce como ensemble ortogonal de Wishart (WOE) [13]. En series de tiempo,  $C$  representa la matriz de covarianza calculada sobre las series de tiempo estocásticas de horizonte de tiempo  $T$  para  $N$  variables estadísticamente independientes. Esto quiere decir que en promedio  $C$  no tendrá correlaciones cruzadas [24], pero las matrices individuales si tendrán fluctuaciones, porque no normalizamos cada serie. Para  $T$  muy grande ese efecto será menos notable.

En la practica trabajamos con series de tiempo finitas por lo que es necesario normalizar (restándoles su promedio y dividiendo entre su desviación estándar) la series de tiempo para obtener una matrices de correlación.

Para este ensemble, la distribución de probabilidad conjunta (JPD, por sus siglas en inglés) de los valores propios de  $C$ ,  $\lambda_j$ , para  $1 \leq j \leq N$ , puede ser obtenida usando técnicas desarrolladas para ensembles gaussianos [13].

La JPD de los valores propios es:

$$P(\lambda_1, \dots, \lambda_N) \propto \prod_{j=1}^N w(\lambda_j) \prod_{j>k}^N |\lambda_j - \lambda_k|.$$

donde  $w(\lambda_j) = \lambda_j^{[N(\kappa-1)-1]/2} \exp(-N\kappa\lambda_j/2)$  y  $\kappa = T/N$ ,  $w(\lambda_j)$  es la función de peso de los polinomios de Laguerre asociados. El determinante de Vandermonde en la ecuación anterior proviene del jacobiano de la transformación del espacio del elemento de matriz al espacio valor propio - vector propio. para WOE todas las correlaciones espectrales se conocen en términos de polinomios de Laguerre.

Para grandes  $N$  y  $T$ , con  $\kappa$  finito, la densidad espectral está dada por la densidad de Marčenko Pastur [12]:

$$\rho(\lambda) = \kappa \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\lambda}. \quad (2.1)$$

Donde  $\lambda_{\pm} = (\kappa^{-1/2} \pm 1)^2$  son los puntos de los extremos de la densidad. Nótese que para  $\kappa < 1$ , la densidad  $\rho(\lambda)$  está normalizada a  $\kappa$  y no a 1, ya que la matriz  $C$  es singular. Así, tomando en cuenta los  $(N - T)$  ceros, para  $\kappa \leq 1$ , escribimos la densidad como:

$$\rho(\lambda) = \kappa \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\lambda} + (1 - \kappa)\delta(\lambda). \quad (2.2)$$

### 2.2.3. Ensemble de Wishart correlacionado

En el caso de las correlaciones cruzadas reales definimos el ensemble ortogonal de Wishart correlacionado (CWOE) [5,15]. Para tomar en cuenta las correlaciones entre las variables se considera  $C = \frac{1}{T}\xi^{1/2}BB^t\xi^{1/2}$ , donde  $\xi$  es una matriz positiva definida y simétrica que define las correlaciones [24].

Un tratamiento analítico de este ensemble es más difícil que en el caso no correlacionado. El caso de correlación constante, debido a su sencillez, tiene solución analítica. La matriz  $\xi$  representa la matriz de correlación constante por pares  $c$ . Los valores propios de esta matriz están dados por

$$\lambda_1 = Nc + 1 - c \quad \text{y} \quad \lambda_i = 1 - c, \quad (2.3)$$

para  $i \neq 1$ . Es decir, solo tiene un valor propio que sobresale. En este caso la densidad espectral toma la forma [21]

$$\rho(\lambda) = \rho'(\lambda) + \frac{1}{N}\delta\left(\lambda - \frac{(Nc + 1 - c)(Nc\kappa + 1 - c)}{Nc\kappa}\right), \quad (2.4)$$

$$\rho'(\lambda) = \kappa \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi(1-c)\lambda}, \quad (2.5)$$

donde  $\lambda_{\pm} = (1-c)(\kappa^{-1/2} \pm 1)^2$  y  $\kappa = T/N$ .

Podemos notar que la función delta de la densidad espectral, el valor propio más grande de la distribución esta en

$$\lambda_1 = \frac{(Nc + 1 - c)(Nc\kappa + 1 - c)}{Nc\kappa} = (Nc + 1 - c) \left(1 + \frac{1-c}{Tc}\right), \quad (2.6)$$

y éste difiere del valor propio más grande de  $\xi$  (Eq. 2.3) en la dependencia de  $(1-c)/Tc$ .

## 2.3. Mapeo de potencias

### 2.3.1. Metodología

Las matrices de correlación calculadas con serie de tiempo finitas contienen una cantidad considerable de ruido [9]. Por lo que el horizonte de tiempo que elijamos tiene que ser tal que permita tener poco error y al mismo tiempo sea un buen estimador de correlaciones futuras. En la practica no podemos evitar el error.

El mapeo de potencias es un método de contracción no lineal que fue desarrollado por Guhr y Kaelber [7]. El método toma cada elemento de la matriz de correlación y eleva su valor absoluto a una cierta potencia  $q = 1 + \epsilon$ , mientras que preserva el signo,

$$C_{kl}^{(q)} = \text{sign}(C_{kl}) |C_{kl}|^q. \quad (2.7)$$

Vale la pena resaltar que  $C^{(q)}$  no es lo mismo que  $C^q$ . Cuando  $q$  es mayor que uno, las entradas de la matriz serán contraídas, dado que sus valores absolutos son menores o iguales a uno. La idea de este método es que el ruido será suprimido más fuertemente que las correlaciones reales. Sin embargo, si la  $q$  es muy grande, las correlaciones reales serán también suprimidas.

### 2.3.2. Mapeo de potencias para matrices de correlación singulares

La aplicación del mapeo de potencias para matrices de correlación singulares se sigue del trabajo de Vinayak [23].

Las matrices de correlación singulares surgen cuando el número de series de tiempo por analizar excede su longitud. Usar mas series de tiempo no producirá más valores propios distintos de cero si la matriz se convierte singular.

Debido a la no-linealidad del mapeo de potencias al aplicarlo a las matrices de correlación singulares se rompe la degeneración de los valores propios iguales a cero. La parte del espectro que resulta al romper la degeneración la llamamos *espectro emergente*. El espectro emergente es sensible a las correlaciones y para potencias cerca de la identidad se muestra una buena separación entre el y el espectro que corresponde a los valores propios que eran originalmente diferentes de cero [23].

El mapeo de potencias usando  $q$  del orden de 1.5 a 2 es muy útil para la reducción de ruido y esto era su propósito original [7,18]. Desafortunadamente no se puede tener la separación de los espectros a estos valores de  $q$ . Por ende esta versión de la aplicación se presta sobre todo para estudiar los valores propios más grandes.

El espectro emergente nos da más información invariante a través de los valores propios. Aunque es difícil por la no linealidad, algunos resultados analíticos pueden ser calculados usando WOE y un caso especial de CWOE donde  $\xi_{jk} = \delta_{jk} + (1 - \delta_{jk})c$  y  $0 < |c| < 1$  [24]. Este modelo se refiere al modelo de matriz de correlación constante. Mencionaremos algunas propiedades importantes del espectro de  $C^{(q)}$  que se observan en una simulación numérica simple para WOE. Notamos que  $C^{(q)}$  es siempre una matriz real y simétrica, por lo que sus valores propios son reales, sin embargo, para  $q \neq 1$ , puede tener valores propios negativos especialmente cuando  $T$  es mucho más pequeña que  $N$ . La función densidad de los valores propios de  $C^{(q)}$  está separado en dos soportes como se puede ver en Fig.2.1 obtenida de la ref. [24], el primero corresponde al espectro emergente y es cercano a cero, mientras que el segundo es cercano al soporte definido por la densidad de Marčenko-Pastur debido a pequeñas correcciones al espectro original de valores propios diferente de cero. Conforme vamos aumentando la potencia estos dos soportes empiezan a traslaparse.

Para WOE estimaciones de los primeros dos momentos del espectro emergente han sido derivados en [23], en el cual los autores hacen un expansión alrededor de  $q = 1$  introduciendo el parámetro  $\alpha = (q - 1)$  y definiendo  $C^{(\alpha)} \equiv C^{(q)}$  como está definido en (Eq. 2.7). para  $\alpha$  pequeña,  $C^{(\alpha)}$  puede ser expandida como

$$\begin{aligned} C_{jk}^{(\alpha)} &= C_{jk}^{(0)} \exp\left[\frac{\alpha}{2} \ln(C_{jk}^{(0)})^2\right], \\ C_{jk}^{(\alpha)} &= C_{jk} + \frac{\alpha}{2} C_{jk} \ln(C_{jk}^2) [1 + O(\alpha)], \end{aligned} \tag{2.8}$$

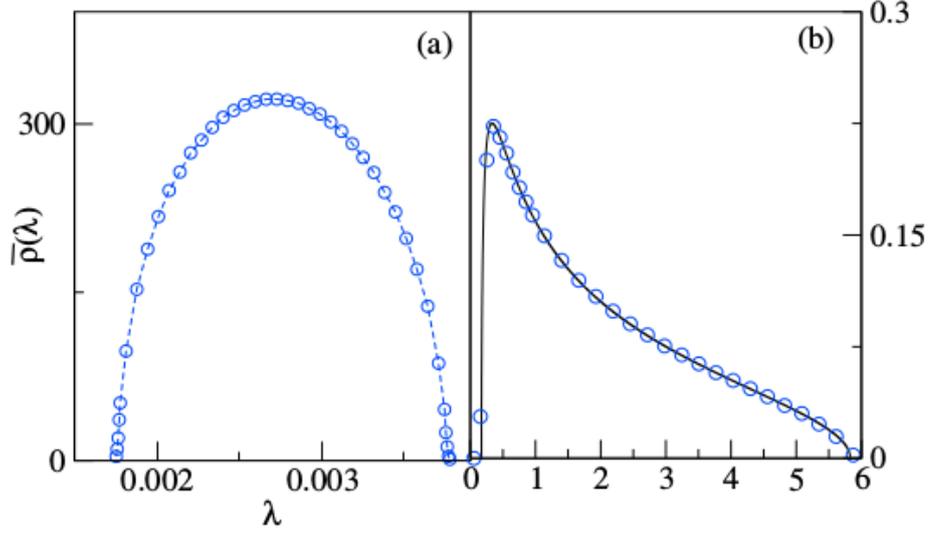


Figura 2.1: Densidad de los valores propios de  $C^{(a)}$  para el caso de WOE. En (a) se muestra la densidad del espectro emergente y en (b) se muestra la densidad de los ex valores propios diferentes de cero la cual es muy cercana a la densidad de Marčenko Pastur que se muestra con una linea solida. Figura tomada de [24].

en la ecuación anterior se usa  $C$  en vez de  $C^{(0)}$ . Ahora expandiendo los valores propios  $\lambda(\alpha)$  de  $C^{(\alpha)}$  como

$$\lambda_j(\alpha) = \lambda_j(0) + \alpha(\delta\lambda_j)[1 + O(\alpha)]. \quad (2.9)$$

donde los  $\lambda_j(0)$ 's son los valores propios de  $C$ , para  $j = 1, \dots, N$  y los  $\alpha(\delta\lambda_j)$ 's son las correcciones de orden que vienen del mapeo de potencias. Para un tiempo corto  $T < N$ ,  $\lambda_j(0) = 0$  para  $j \leq N - T + 1$ . Para  $\alpha$  pequeña asumimos que las estadísticas de los cambios relativos en los valores propios está dominada por el término lineal. Entonces, se derivan estimaciones para los momentos de los  $\alpha(\delta\lambda_j)$ 's en el regimen de respuesta lineal, conciderando  $\alpha > 0$ .

Para todas las correcciones de los valores propios, se define los momentos como

$$\overline{\delta m_n} \equiv \frac{\alpha^n}{N} \overline{\sum_{j=1}^N (\delta\lambda_j)^n}. \quad (2.10)$$

Notemos que en el régimen de respuesta lineal  $\overline{\delta m_n}$  también puede ser estimado como

$$\overline{\delta m_n} \simeq \alpha^n \quad \text{término en} \quad \frac{1}{N} [\overline{Tr(C^{(\alpha)})^n} - \overline{Tr(C)^n}]. \quad (2.11)$$

Definiendo  $\overline{\delta m_n^{(0)}}$  como los momentos del espectro emergente y  $\overline{\delta m_n^{(1)}}$  como los momentos de las correcciones de los valores propios diferentes de cero, entonces

$$\overline{\delta m_n} = \overline{\delta m_n^{(0)}} + \overline{\delta m_n^{(1)}}. \quad (2.12)$$

Nótese que  $\overline{\delta m_n^{(1)}} \equiv \overline{\delta m_n}$  para  $T \geq N$ . En Ref. [23] se obtiene en el régimen de respuesta lineal

$$\begin{aligned} \overline{\delta m_1} &\sim \frac{\alpha}{T}, \\ \overline{\delta m_2} &\sim \frac{\alpha^2}{4\kappa} ([\log(T) + c_1]^2 + c_2), \end{aligned} \quad (2.13)$$

donde  $c_1 = \gamma + \log(2) - 2 = -0,729637\dots$  y  $c_2 = \frac{\pi^2}{2} - 4 = 0,934802\dots$ ,  $\gamma$  es la constante de Euler, y

$$\begin{aligned} \overline{\delta m_1^{(1)}} &= \kappa \overline{\delta m_1} + s(1 - \kappa), \\ \overline{\delta m_2^{(1)}} &= \kappa \overline{\delta m_2} - \kappa \overline{\delta m_1}^2 + \frac{\overline{\delta m_1^{(1)}}^2}{\kappa}. \end{aligned} \quad (2.14)$$

donde  $s \sim -\frac{\alpha}{2} \sqrt{[\log(T) + c_1]^2 + c_2}$  es el parámetro de escala. En la derivación también se usó un parámetro de desplazamiento  $r = \overline{\delta m_1} - s(1 - \kappa)$ . Usando (Eq.2.12) la estimación de los momentos para el espectro emergente se vuelve trivial. Para valores grandes de  $T$  y  $\kappa \geq 1$ , tenemos

$$\begin{aligned} \overline{\delta m_1^{(0)}} &= -s(1 - \kappa), \\ \overline{\delta m_2^{(0)}} &= s^2(1 - \kappa). \end{aligned} \quad (2.15)$$

Para valores pequeños de  $\kappa$ , el error en la aproximación se vuelve grande.

Para matrices de correlación constante para el CWOE,  $\xi$  es una matriz densa sus valores propios  $\xi_j$ 's son simplemente dados por  $\xi_j = (1 - c)$  para  $1 \leq j \leq N$  y  $\xi_N = N_c + 1 - c$  y su densidad está dada por (2.4). La función delta aparece en el resultado anterior siempre y cuando  $C \geq (N\sqrt{\kappa})^{-1}$ . Usando esto encontramos

$$\overline{\delta m_1^{(1)}} = \kappa \overline{\delta m_1} + (1 - c)s(1 - \kappa). \quad (2.16)$$

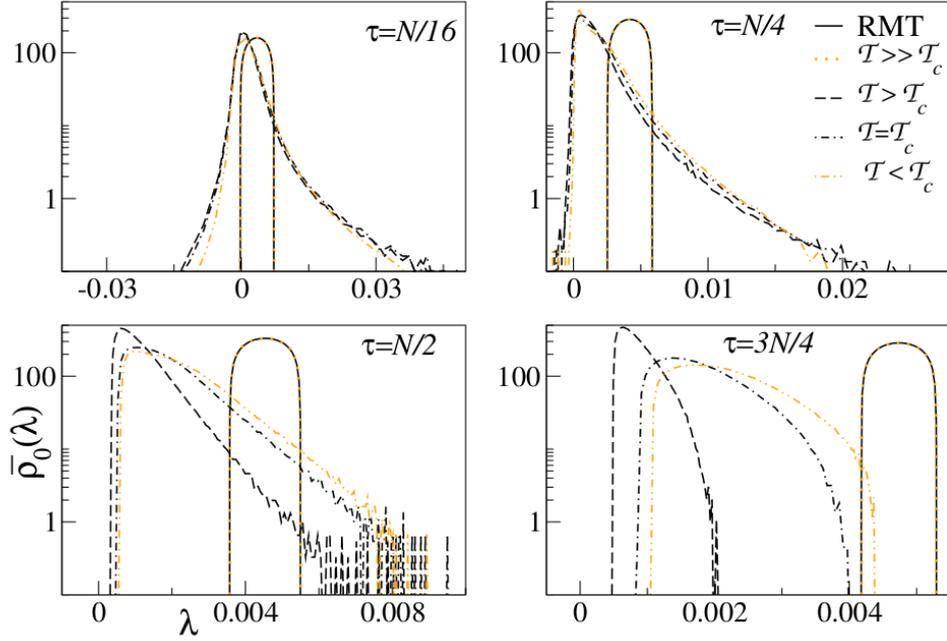


Figura 2.2: Densidad del espectro emergente  $\bar{\rho}_0(\lambda)$  para diferentes series de tiempo cortas a varias temperaturas, comparando con los resultados obtenidos para el modelo de Ising bidimensional en el punto crítico con los de RMT. Figura tomada de [22].

El resultado del segundo momento es el mismo que el obtenido en (2.14).

En este trabajo también nos enfocaremos en el estudio hecho en el artículo [22], donde aplican el método de mapeo de potencias a matrices de correlación singulares obtenidas de las simulaciones numéricas de un modelo Ising 2-D para encontrar una señal significativa en el comportamiento de las matrices de correlación en casos críticos y en comparación con resultados de alta temperatura que tienden a comportarse como los de RMT. En este trabajo se observa un comportamiento en el espectro emergente dependiente de la temperatura, como se muestra en la figura 2.2, obtenida de [22].



# Capítulo 3

## Resumen del método y aplicación

La mayoría de los análisis de las series de tiempo financieras asumen la estacionariedad, aunque ésta puede ser una suposición de validez variable. La validez del supuesto mejora conforme se utilizan series temporales más cortas.

En este trabajo usamos matrices altamente singulares, debido a que al construir la matriz de correlación utilizamos un número considerablemente más grande de series de tiempo que el número de pasos en el tiempo, lo que produce muy pocos valores propios diferentes de cero. Para romper la degeneración de los valores propios aplicamos el método de mapeo de potencias. Este método se explica a continuación.

Construimos las series de tiempo con el precio de cierre diario de 293 compañías del índice financiero Standard & Poor's 500 durante el período 1992-2014.<sup>1</sup> La lista de compañías aquí estudiadas fue otorgada por el M.C. José Morales la cual usó en su tesis de maestría y corresponde a las estudiadas en [16] salvo por algunas que se sacaron por no estar presentes hasta el 2014. De las series de tiempo obtenidas calculamos los rendimientos:

$$R_{\tau}^j \equiv \frac{S_{\tau+\Delta\tau}^j - S_{\tau}^j}{S_{\tau}^j},$$

donde  $S_{\tau}^j$  es el precio de cierre de la compañía  $j$ -ésima en el tiempo  $\tau$ , escogemos el intervalo  $\Delta\tau$  igual a 1 día.<sup>2</sup>

---

<sup>1</sup> Sólo se utilizan los días laborales.

<sup>2</sup>En [16] quitan los rendimientos de viernes a lunes porque es un salto de más de 24 horas, mientras que los días restantes de la semana son saltos de 24 horas. Varios calculos han mostrado que no hay diferencia cualitativa si incluimos el rendimiento de viernes-lunes. Por lo que nosotros tomaremos todos los datos y los trataremos de igual manera.

Sea  $A$  una matriz de dimensión  $N \times T$ , construida con las series de tiempo normalizadas de los rendimientos, es decir, el renglón  $A_j$  representa la serie de tiempo de los rendimientos de la acción  $j$ -ésima. Normalizamos la series de tiempo como:

$$A_{\tau}^j = \frac{R_{\tau}^j - \eta_j}{\sigma_j},$$

donde  $\eta_j$  y  $\sigma_j$  son el promedio y la desviación estándar de la  $j$ -ésima serie de tiempo.

Decimos que son series de tiempo cortas porque creamos las series tomando una ventana cada 44 días que son aproximadamente dos meses. Desplazamos la ventana a lo largo del período de tiempo 1992-2014, por lo que las matrices varían en el tiempo. Tomamos el tiempo inicial  $\tau$  de la ventana de tiempo como referencia de la matriz  $A$ .

Si se toma una matriz  $A$  para un tiempo dado, la matriz de correlación  $C$  se construye de la siguiente forma:

$$C = \frac{1}{T}AA^t,$$

donde  $A^t$  es la transpuesta de la matriz  $A$ .

Se aplica el método de mapeo de potencias a la matriz de correlación  $C$  de la siguiente forma:

$$C_{kl}^{(q)} = \text{sign}(C_{kl})|C_{kl}|^q. \quad (3.1)$$

Donde  $C_{kl}$  son los elementos de la matriz de correlación  $C$ .

# Capítulo 4

## Resultados

Como ya se mencionó en el capítulo anterior el conjunto de acciones que se analizaron corresponden a  $N = 293$  acciones del índice S&P500 durante el período de tiempo entre 1992 y 2014. Las compañías elegidas cotizaron sus acciones de manera ininterrumpida en el índice S&P500 a lo largo de este período y corresponden a las elegidas en [16] a excepción de algunas que se quitaron por no estar presentes hasta el 2014. Las matrices de correlación se construyen tomando una ventana de tiempo  $T=44$  de las series de tiempo de los rendimientos calculados de los precios de cierre diario de las acciones, de esta forma obtuvimos 131 matrices de correlación. A todas las matrices de correlación se les aplica el método de mapeo de potencias con una potencia  $q=1.05$ . Las series de tiempo fueron ordenadas por sectores industriales de acuerdo a la clasificación GICS (clasificación industrial global estándar, por sus siglas en inglés).<sup>1</sup> Los datos fueron obtenidos de Yahoo! Finance.

### 4.1. Distribución de valores propios

En las siguientes sección se estudiara como son las densidades espectrales. El espectro de valores propios está separado principalmente en dos soportes (fig.4.1), el primero es el espectro emergente que surge debido al rompimiento de la degeneración de los valores igual a cero después de aplicar el mapeo de potencias, éste cuenta con  $N-T+1$  valores propios, en nuestro caso son 250 valores propios. El segundo soporte es el formado por los valores propios correspondientes a los valores inicialmente diferente de cero denominado *bulk*.

---

<sup>1</sup> La clasificación GICS fue propuesta por Standard & Poor's y MSCI Barra en 1999, y ahora es reconocida internacionalmente. A las compañías se les asigna un sector dependiendo de cuál sea su principal actividad empresarial. Hay 10 sectores: Energía, Materiales, Industriales, Bienes de consumo discrecional, Cuidado de la salud, Finanzas, Tecnología de la información, Servicios de telecomunicaciones y Servicios públicos.

El bulk después de aplicar el mapeo de potencias con  $q=1.05$  no presenta grandes cambios con respecto a la distribución de valores que se obtiene antes de aplicar el mapeo de potencias. El valor propio más grande se separa notablemente de los demás valores propios, por lo que llamaremos bulk a la distribución formada por los valores propios correspondientes a los valores inicialmente diferente de cero sin el valor propio mayor. Por lo tanto el bulk cuenta con 42 valores propios. Cabe destacar que unos cuantos de los valores propios más grandes se separan un poco del resto, estos corresponden a la correlación dentro de los sectores industriales. También se puede ver esta separación en algunos de los valores propios más grandes del espectro emergente para algunas fechas del período, como en fig.4.1 (c).

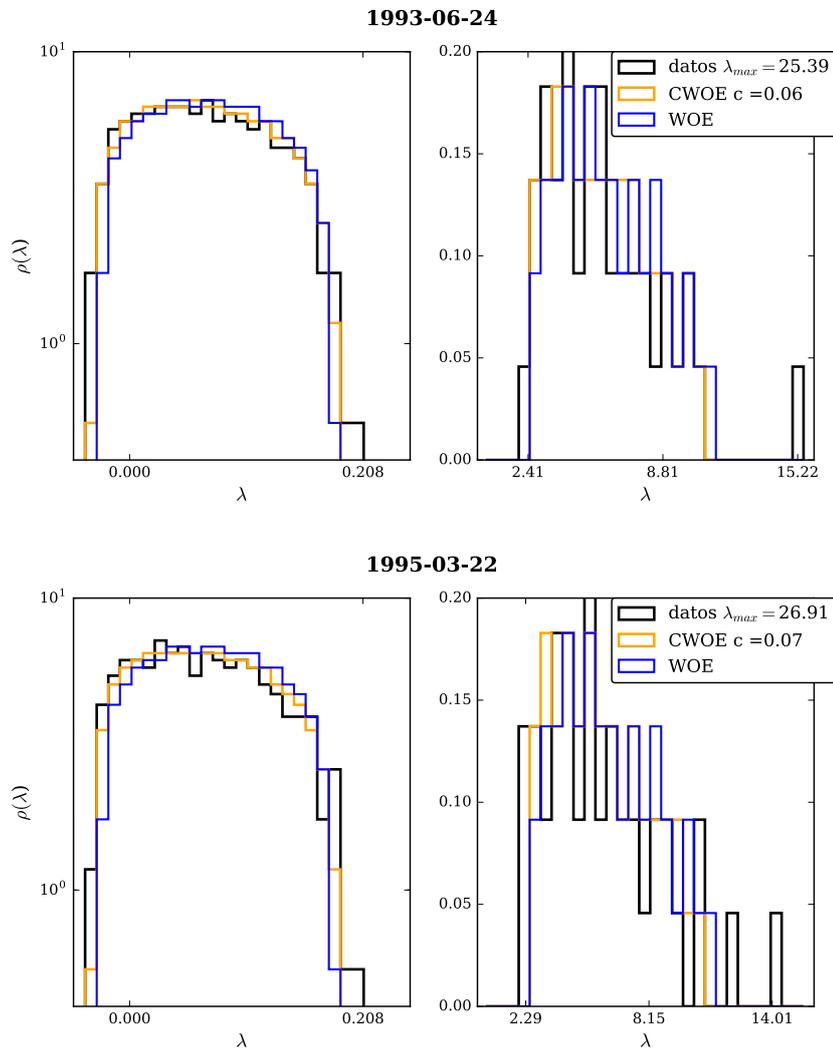


Figura 4.1: El espectro de valores propios aplicando el mapeo de potencias ( $q=1.05$ ) para dos fechas diferentes que corresponden a épocas sin crisis. En la imagen izquierda de cada figura se muestra el espectro emergente y del lado derecho vemos el bulk. Las figuras se muestran en diferente escala. En negro se muestra el espectro para los datos tomados del índice S&P500. En azul se muestra espectro de valores propios para WOE y en naranja para CWOE con correlación constante, usamos la correlación promedio de la matriz en cada época. El eje  $y$  está en escala logarítmica en las imágenes del lado izquierdo de cada figura.

En las figuras 4.1(a) y 4.1(b) podemos ver que, en los años en los que no hubo crisis, el espectro emergente y el bulk se asemejan más al caso aleatorio obtenido numéricamente para WOE, a excepción de unos cuantos valores que se separan del bulk. En cambio, cuando se acerca el período de crisis y durante las crisis el espectro es más asimétrico, alejándose del obtenido para WOE, 4.2(a) y 4.2(b). Por otra parte, podemos ver una separación mayor entre el bulk y el espectro emergente en épocas alejadas del período de crisis.

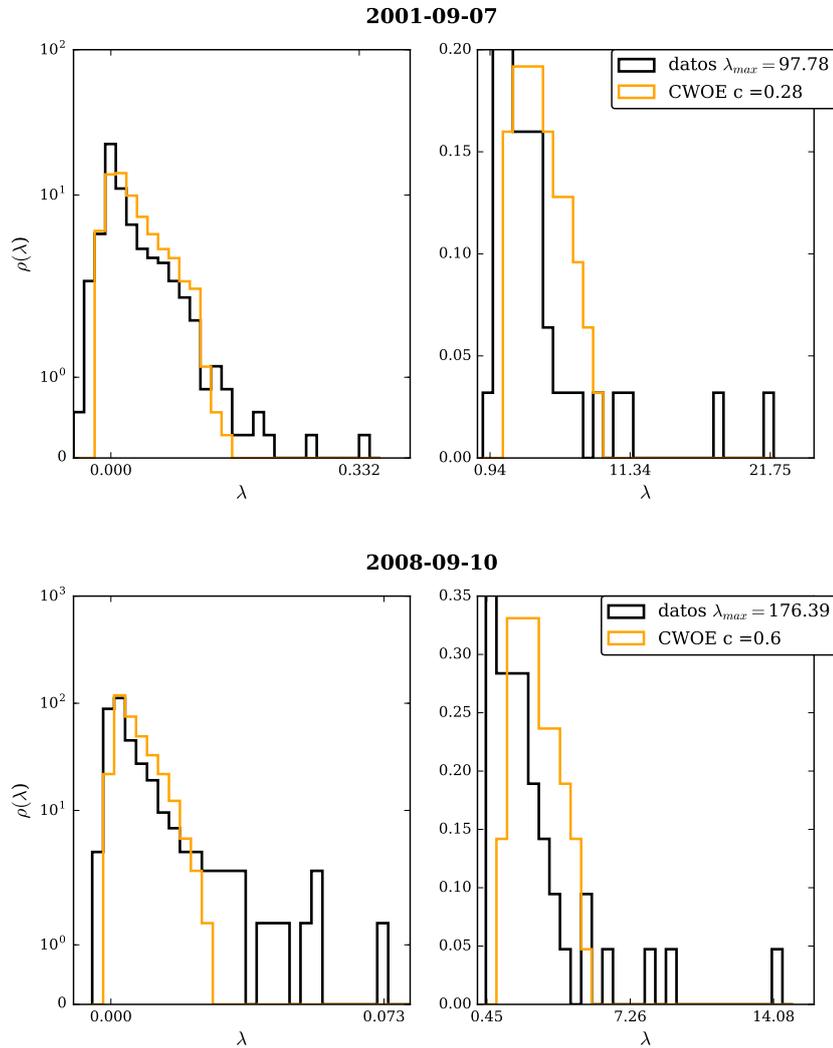


Figura 4.2: El espectro de valores propios aplicando el mapeo de potencias ( $q=1.05$ ) para dos fechas diferentes que corresponden a épocas con crisis. En la imagen izquierda de cada figura se muestra el espectro emergente y del lado derecho vemos el bulk. Las figuras se muestran en diferente escala. En negro se muestra el espectro para los datos tomados del índice S&P500. En naranja se muestra espectro de valores propios para CWOE con correlación constante, usamos la correlación promedio de la matriz en cada época. El eje  $y$  está en escala logarítmica en las imágenes del lado izquierdo de cada figura.

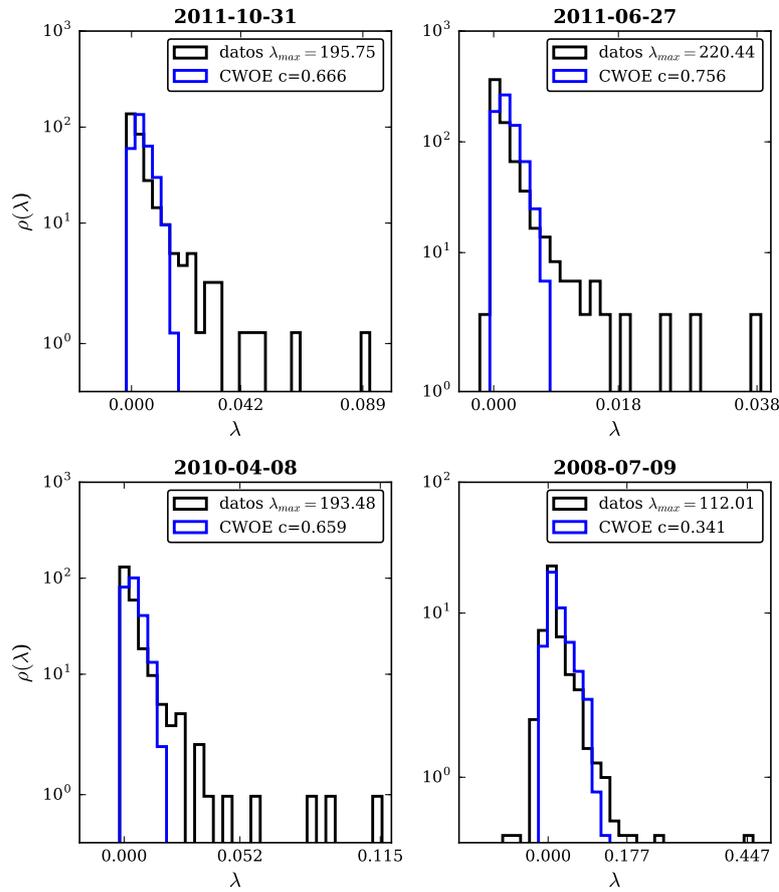
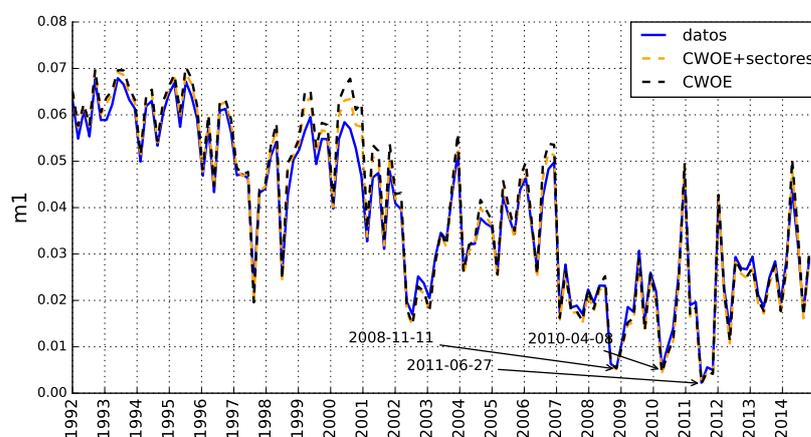


Figura 4.3: El espectro emergente para diferentes fechas de los datos tomados del índice S&P500 se muestra en color negro y en azul se muestra espectro emergente para CWOE con correlación constante. En los recuadro en la esquina superior derecha de las figuras se muestra el valor propio más grande del espectro completo y el valor de la correlación promedio correspondiente a cada período. Nótese que el eje  $y$  está en escala logarítmica

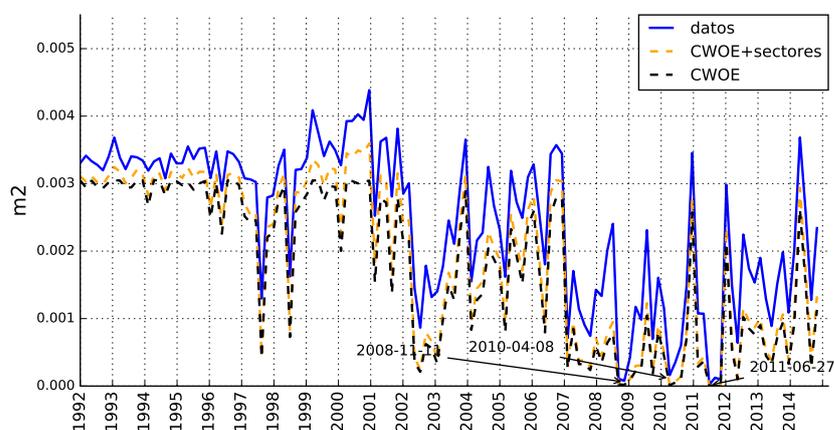
#### 4.1.1. Momentos del espectro emergente

Para analizar la forma de la distribución del espectro emergente estudiamos el primer momento (promedio), el segundo momento (anchura), la asimetría (skewness) y la curtosis de dicha distribución sobre todo el período estudiado. En la evolución temporal de los momentos, las fechas con valores más pequeños en el primer y segundo momento coinciden con las fechas de mayor curtosis y asimetría. A su vez, dichos valores corresponden a si-

tuaciones críticas del sistema financiero, donde la matriz de correlación se encuentra altamente correlacionada, como se ve en la figuras 4.4 y 4.5.



(a)



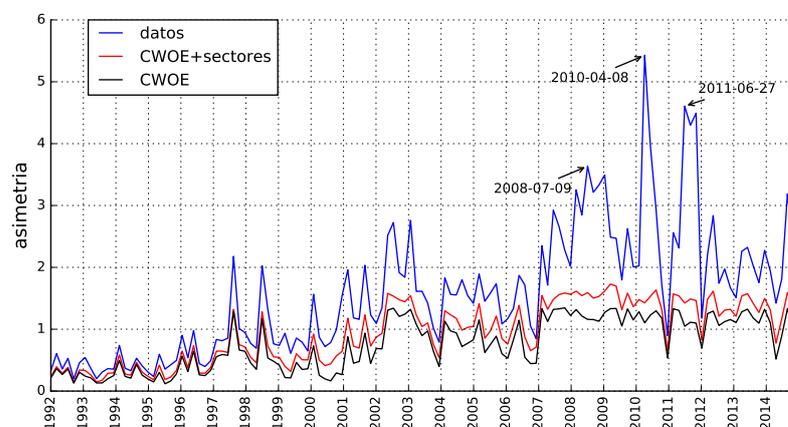
(b)

Figura 4.4: Se muestra la evolución temporal del primer y segundo momento del espectro emergente a lo largo del período 1992-2014. En (a) se muestra el primer momento del espectro emergente y en (b) se muestra el segundo momento del espectro emergente. En ambas figuras el resultado obtenido para los datos empíricos se muestra en azul y con líneas punteadas los obtenidos de forma numérica para CWOE. En amarillo se muestra el caso CWOE donde la correlación fija es una matriz de 10 bloques diagonal, cada bloque tiene el tamaño y la correlación promedio de su sector correspondiente en cada época a lo largo del período y en los elementos fuera de los bloques lleva la correlación promedio entre sectores. En negro se muestra los resultados obtenidos para CWOE con correlación constante, donde la correlación es la correlación promedio de la matriz en cada época a lo largo del período. Las anotaciones corresponden a los valores más bajos de los momentos y coinciden con épocas críticas del mercado financiero.

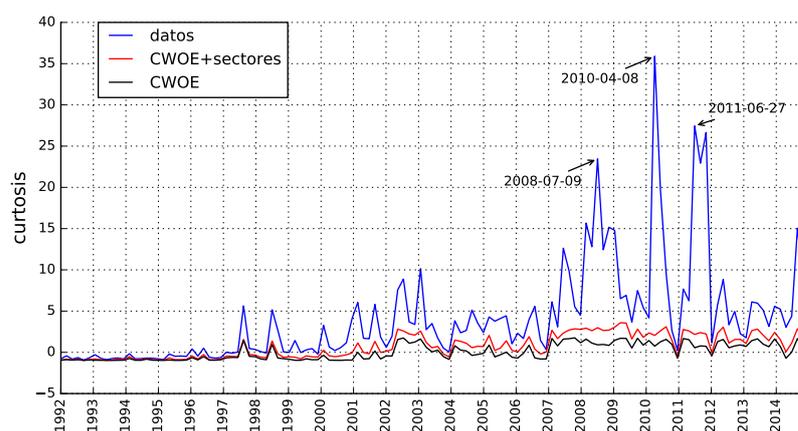
Se compararon los momentos de los datos empíricos con los obtenidos para el promedio de las distribuciones de dos casos de CWOE. El primer caso es con correlación constante, la correlación que se usó es el promedio de los elementos de la matriz de correlación para cada ventana de tiempo a lo largo del período estudiado. En el segundo caso, la correlación dada externamente es una matriz de bloques diagonal, construida con 10 bloques, para representar a los sectores industriales. Cada bloque tiene el tamaño y la correlación promedio de su sector correspondiente, en los elementos fuera de los bloques lleva la correlación promedio entre sectores. Se observa que para correlaciones planas el primer y segundo momento del espectro se comporta similar al de los datos reales, aunque su similitud es más notoria en el caso de CWOE con bloques (fig. 4.4). Para la asimetría y la curtosis se observa que el comportamiento es similar durante los primeros años, pero difiere durante el período del 2007 al 2012, el cual abarca quiebras financieras <sup>2</sup>

---

<sup>2</sup>Crisis de las hipotecas chatarra (Subprime) y crisis de deuda de la zona euro (euro debt crisis)



(a)



(b)

Figura 4.5: Se muestra la evolución temporal de la asimetría y la curtosis del espectro emergente a lo largo del período 1992-2014. En (a) se muestra la asimetría del espectro emergente y en (b) se muestra la curtosis del espectro emergente. En ambas figuras el resultado obtenido para los datos empíricos se muestra en azul y con línea negra y roja los obtenidos numéricamente para CWOE. En rojo se muestra el caso CWOE donde la correlación fija es una matriz de 10 bloques diagonal, cada bloque tiene el tamaño y la correlación promedio de su sector correspondiente en cada época a lo largo del período y en los elementos fuera de los bloques lleva la correlación promedio entre sectores. En negro se muestra los resultados obtenidos para CWOE con correlación constante, donde la correlación fija es la correlación promedio de la matriz a lo largo del período. Las anotaciones corresponden a los valores más altos de la asimetría y curtosis, los cuales coinciden con épocas críticas del mercado financiero.

En el espectro emergente para las épocas 2011-10-31, 2011-06-27, 2010-04-08 y 2008-07-09 se muestra en fig.4.3, la mayoría de valores propios permanece cerca de cero. Sin embargo, algunos de los valores más grandes sobresalen del resto, a diferencia de lo obtenido para los datos sintéticos. Para CWOE no se observan valores propios que se separen. Cabe mencionar que en todas las fechas el espectro emergente presenta valores negativos, como se puede ver en fig.4.8.

Los momentos del espectro emergente están relacionados con la correlación promedio de la matriz, como se ve en la figura 4.6. En general,  $\langle C_{ij} \rangle$  se muestra anticorrelacionada con la varianza, sin embargo, su anticorrelación es más notoria durante el período de crisis. Por otra parte, la correlación constante y la asimetría se muestran correlacionadas. No obstante, se observó que su comportamiento muestra una anticorrelación poco antes de la quiebra del banco Lehman Brothers (en septiembre del 2008) donde se registró una gran pérdida en el mercado financiero.

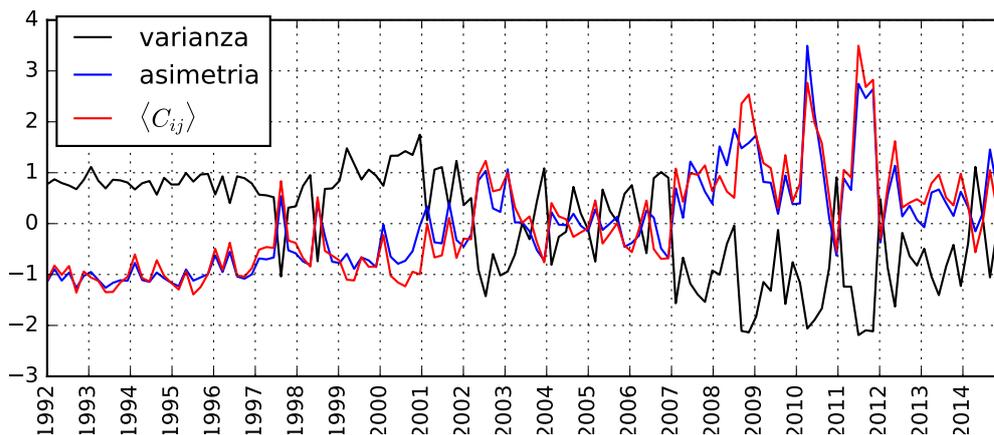
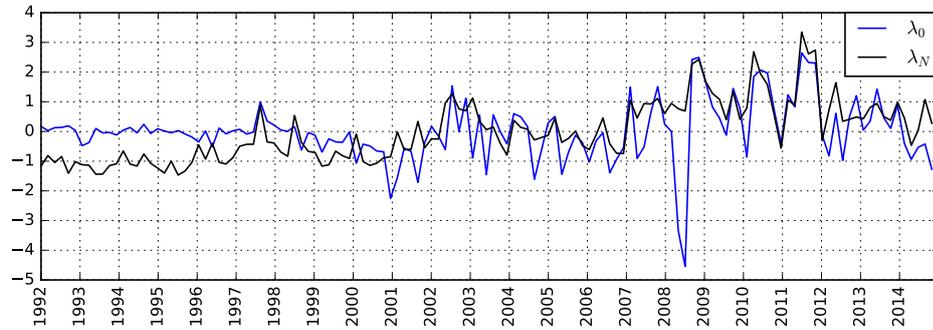
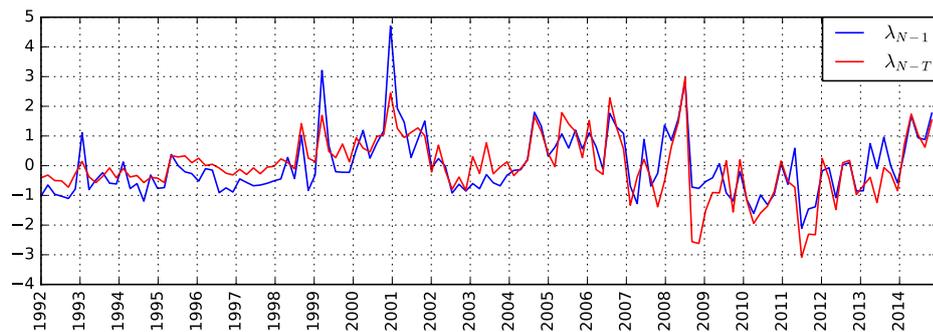


Figura 4.6: Comparamos la varianza y la asimetría con la correlación promedio de la matriz. Mostramos en negro la varianza del espectro emergente, en azul la asimetría del espectro emergente y en rojo la correlación promedio de la matriz  $\langle C_{ij} \rangle$  para cada ventana de tiempo a lo largo del período. En términos gruesos hay correlación con la asimetría y  $\langle C_{ij} \rangle$  y anticorrelación con la varianza. La anticorrelación entre  $\langle C_{ij} \rangle$  y la varianza es más notoria durante el período de crisis. La asimetría muestra un comportamiento anticorrelacionado junto antes de la quiebra del banco Lehman Brothers en septiembre del 2008. Los datos mostrados en esta figura fueron normalizados con promedio cero y varianza 1.

### 4.1.2. Valores propios extremos del espectro emergente y bulk



(a)



(b)

Figura 4.7: En (a) se compara el máximo y mínimo valor propio del espectro completo. En azul se muestra el valor propio menor y en negro el valor propio mayor. Podemos ver como el valor menor desde 2009 hasta 2014 correlaciona fuertemente con el valor mayor. Pero muestra una anticorrelación marcada durante 2000-2002. En (b) se compara el segundo mayor valor propio de todo el espectro, que se muestra en azul, y el valor propio mayor del espectro emergente, que se muestra en rojo. Los cuales tienen un comportamiento similar. Los valores están normalizados con promedio cero y varianza 1.

En la figura 4.7 (a) se comparan el valor propio mayor  $\lambda_N$  y el valor propio menor  $\lambda_0$  del espectro completo. Podemos ver como el valor menor se correlaciona y anticorrelaciona con el valor mayor. Se anticorrelacionan fuertemente durante el período 2000-2002, que fue cuando ocurrió la crisis por la burbuja punto com. Notamos una correlación significativa durante

2011-2012. En un estudio más detallado en progreso en una colaboración entre nuestro grupo y el del Dr. A. Chakraborti indican provisionalmente que para ventanas de tiempo más cortas con  $T=20$  el valor propio menor del espectro emergente se correlaciona o anticorrelaciona en períodos de crisis.

El valor mayor del espectro emergente está altamente correlacionado con el segundo valor mayor del espectro completo, como se ve en la figura 4.7 (b).

## 4.2. Número de valores negativos

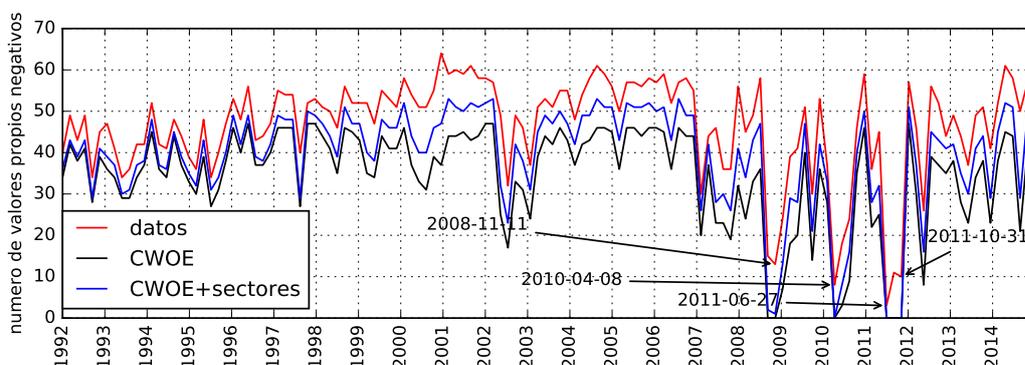


Figura 4.8: En esta figura se muestra el número de valores propios negativos presentes en el espectro emergente durante el período 1992-2014. En color rojo se muestra el resultado para los datos empíricos. En azul se muestra el caso CWOE donde la correlación fija es una matriz de 10 bloques diagonal, cada bloque tiene el tamaño y la correlación promedio de su sector correspondiente, en los elementos fuera de los bloques lleva la correlación promedio entre sectores para cada época a lo largo del período. En negro se muestra los resultados obtenidos para CWOE con correlación constante donde la correlación es la correlación promedio de la matriz en cada época a lo largo del período.

La presencia de valores propios negativos es notoria en el espectro emergente. En la fig. 4.8 se muestra el número de valores propios negativos para cada época. Vemos más valores negativos en los períodos alejados de la época de crisis y en períodos de crisis encontramos menor cantidad de valores negativos. Las épocas con menos número de valores negativos coinciden con las épocas que sobresalen y que se muestran con anotaciones en fig.4.4 y fig.4.5.

El número de valores negativos muestra una alta anticorrelación con la correlación promedio en gran parte del período estudiado (a partir de la mitad del 2002 y hasta el 2014), en especialmente durante las crisis. Por otra parte, el número de valores negativos y la correlación promedio se muestran fuertemente correlacionados en los primeros años (antes de 1997), los cuales corresponden a épocas lejanas a las crisis. Durante 1997-2002 se observa un cambio notorio de correlación fuerte a anticorrelación, la primera anticorrelación se nota en el 1997 y a partir del 2002 la anticorrelación se hace muy marcada. Ésto se puede ver como el dramático cambio que sufrieron los mercados en este tiempo. Como se muestra en la figura 4.9.

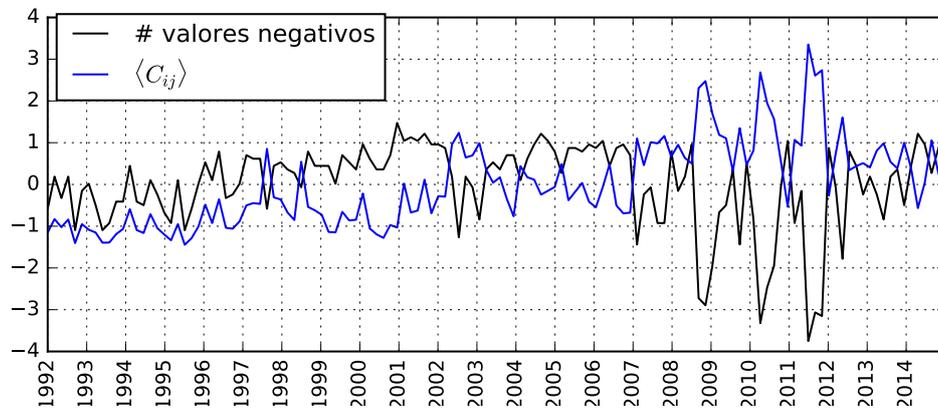


Figura 4.9: Comparando el número de valores propios negativos en el espectro emergente con la correlación promedio de la matriz en cada época del período 1992-2014. En color negro la evolución temporal del número de valores propios y en azul se muestra el valor de la correlación promedio de la matriz a lo largo del tiempo. Los datos mostrados en esta figura fueron normalizados con varianza 1 y promedio cero. Durante 1992-1997 notamos un fuerte correlación entre el número total de valores negativos y la correlación promedio de la matriz. A partir de la mitad del 2002 hasta 2014 vemos una fuerte anticorrelación. Entre estas fechas el comportamiento oscila. Esto se puede ver como el dramático cambio que sufrieron los mercados en este tiempo.

La relación de la correlación promedio  $\langle C_{ij} \rangle$  y el número de valores negativos, así como la relación de  $\langle C_{ij} \rangle$  y los primero dos momentos del espectro emergente y la asimetría se observan en la fig.4.10, donde se comparan con lo obtenido para CWOE donde la correlación constante es la correlación pro-

medio de la matriz de datos empíricos para cada época a lo largo del período estudiado y *CWOE* con bloques, donde la correlación fija es una matriz de 10 bloques diagonal, cada bloque tiene el tamaño y la correlación promedio de su sector correspondiente tomada de la matriz de datos empíricos para cada época a lo largo del período estudiado, en los elementos fuera de los bloques lleva la correlación promedio entre sectores.

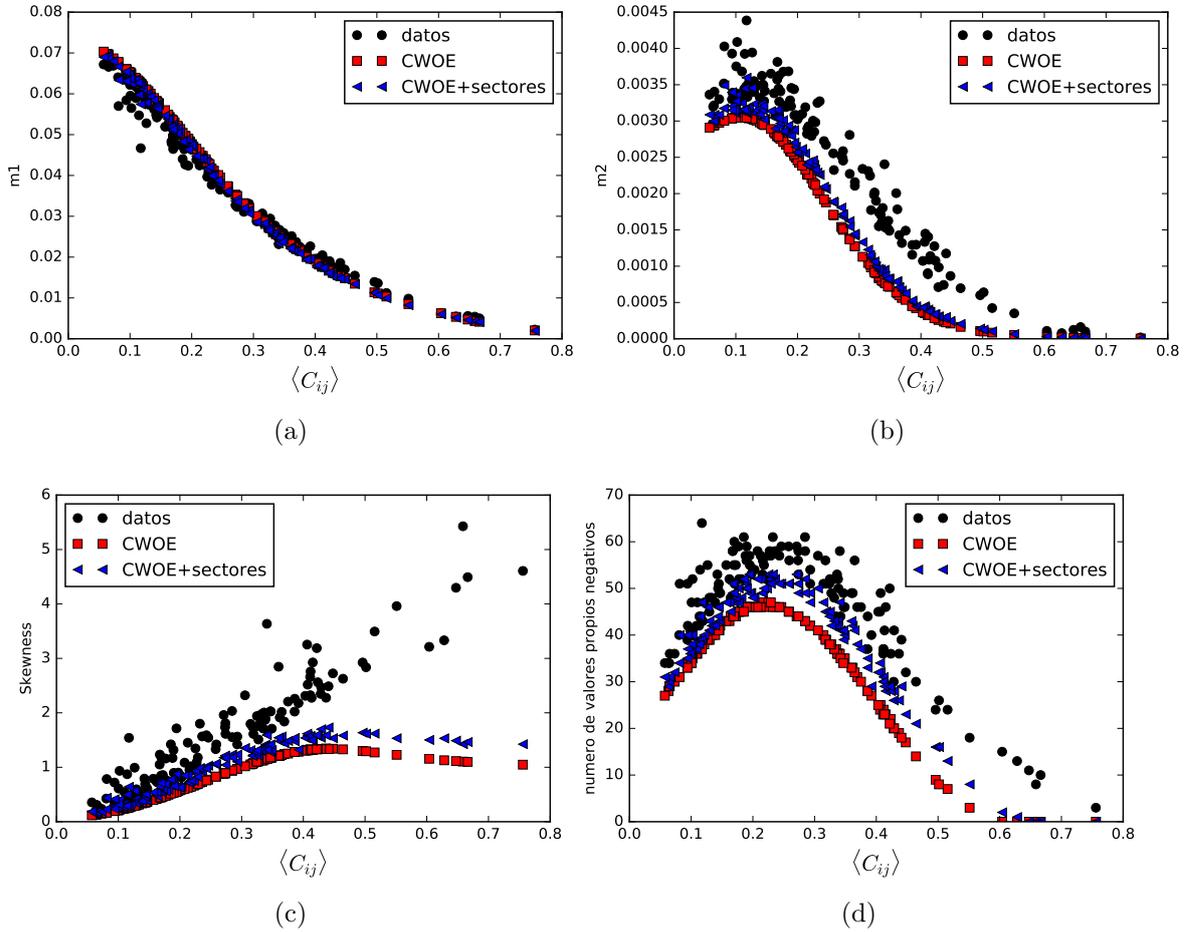


Figura 4.10: Se muestra los primero dos momentos, la asimetría y el número de valores propios negativos en función de la correlación promedio de la matriz. En (a) se muestra el primer momento del espectro emergente contra la correlación promedio, en (b) el segundo momento del espectro emergente contra la correlación promedio y en (c) la asimetría del espectro emergente contra la correlación promedio. En (d) se muestra el número de valores propios negativos contra la correlación promedio de la matriz. Los resultados obtenidos de los datos empíricos se muestran con puntos negros, con cuadros rojos los resultados para CWOE donde la correlación constante es la correlación promedio de la matriz para cada época a lo largo del período estudiado y con triángulos azules los resultados para CWOE con bloques, donde la correlación fija es una matriz de 10 bloques diagonal, cada bloque tiene el tamaño y la correlación promedio de su sector correspondiente tomada de la matriz de datos empíricos en cada época, en los elementos fuera de los bloques lleva la correlación promedio entre sectores.

Se observa que a medida que va aumentando la correlación, el promedio del espectro emergente se va acercando a cero. El segundo momento decrece muy rápido conforme va aumentando la correlación. Para asimetría los casos CWOE difieren significativamente a partir de la correlación 0.4. En 4.10(d) se puede ver como el número de valores cero aumenta con la correlación pero a partir de la correlación alrededor de 0.23 empieza a decrecer al aumentar la correlación. También se observa, que para los casos CWOE con correlaciones grandes no presentan valores negativos, en 4.10(d) se muestran CWOE con bloques, en el cual su relación entre el número de valores negativos con la correlación se asemeja más a los datos empíricos. Eso implica que la clasificación usada para los sectores se refleja en la matriz de correlación de forma significativa. Lo que nos dice que el espectro emergente no sólo es sensible a la correlación promedio de la matriz si no también a la estructura de la matriz.



# Capítulo 5

## Conclusiones

En este trabajo se aplica a datos del mercado financiero de Nueva York (S&P500) un nuevo método [21] para analizar series de tiempo cortas que permite obtener más valores propios y así hacer un mejor análisis estadístico de ellos.

El espectro de valores propios se encuentra distribuido en dos soportes. Los casos alejados de épocas críticas coincide con el comportamiento obtenido para RMT, conforme se va acercando al período de crisis, éste difiere del resultado obtenido para WOE, por lo que en dichas épocas la distribución de CWOE es más apropiada. En general, se presenta una separación mayor entre los dos soportes durante las épocas alejadas de crisis.

Se determinó que el espectro emergente resulta ser muy sensible a las correlaciones. Esto coincide con los resultados obtenidos para el modelo de Ising en dos dimensiones en el trabajo [22]. Existe una relación clara entre los primeros dos momentos del espectro y la correlación promedio de la matriz. De igual forma, se observaron valores negativos en el espectro emergente, los cuales también presentan sensibilidad a la correlación promedio de la matriz. Sin embargo, a diferencia de dicho estudio en [22], fue notoria la cantidad de valores negativos presentes para los casos más alejados de crisis. Además el espectro emergente también muestra sensibilidad a la estructura de la matriz, y en particular a los sectores del mercado conocidos, como vemos de un modelo esquemático de correlaciones con estos bloques.

El valor propio mayor del espectro emergente está más relacionado con el segundo mayor del espectro completo. El valor menor y el valor mayor del espectro completo tienen un comportamiento semejante, a excepción de algunos períodos críticos en el mercado, donde se puede ver que se correlacionan o anticorrelacionan fuertemente.

Se logró caracterizar mejor las épocas de crisis con el espectro emergente, es decir, los momentos de la distribución presentaron comportamiento carac-

terístico durante momentos dramáticos en el mercado financiero. Se observó, una mayor correlación entre la asimetría (skewness) y la correlación promedio durante épocas críticas en el sistema, pero una anticorrelación notoria con la varianza. Incluso, se presenta una fuerte anticorrelación entre el número de valores negativos y la correlación promedio durante mediados del 2002 hasta 2014 y una pronunciada correlación entre ellos durante 1992-1997, debido a que entre 1997 y 2002 la estructura de los mercados cambia radicalmente.

Existe un comportamiento notorio en la asimetría y la varianza durante las fechas donde se registraron grandes pérdidas en los mercados financieros. Los valores de la asimetría sobresalen del resto en fechas de crisis, se observó que su comportamiento muestra una anticorrelación poco antes de la quiebra del banco Lehman Brothers donde se registró una gran pérdida en el mercado financiero.

Resta señalar, que a partir de estos resultados, queda la pregunta abierta ¿Qué tanta información previa a las crisis podemos obtener usando el método de mapeo de potencias?

Un estudio más detallado y con resultados de las últimas semanas es un artículo por publicar, con los siguientes autores: Anirban Chakraborti, Hirdesh Pharasi, Susana Ochoa, Kiran Sharma, Rakesh Chatterjee y Thomas H. Seligman. En este artículo se muestra fuerte correlación entre el valor propio menor del espectro emergente y la correlación promedio fuera de tiempos de crisis, pero durante las crisis se observa a veces correlación y a veces anticorrelación.

# Apéndice A

## Tabla de compañías

<b>Símbolo</b>	<b>Compañía</b>	<b>Símbolo</b>	<b>Compañía</b>
HES	Hess Corporation	BHI	Baker Hughes Inc.
EQT	EQT Corporation	HAL	Halliburton Co.
XOM	Exxon Mobil Corp.	CVX	Chevron Corp.
SWN	Southwestern Energy	SLB	Schlumberger Ltd
APA	Apache Corporation	APC	Anadarko Petroleum Corp.
MUR	Murphy Oil	NBL	Noble Energy Inc
RDC	Rowan Companies Plc	OXY	Occidental Petroleum
TSO		VLO	Valero Energy
ESV	Ensc o plc	COG	Cabot Oil and Gas
MRO	Marathon Oil Corp.	COP	ConocoPhillips
EOG	EOG Resources	NBR	Nabors Industries
WMB	Williams Cos.	ARG	Airgas Inc
IP	International Paper	VMC	Vulcan Materials
APD	Air Products and Chemicals Inc.	DOW	
SIAL	Sigma-Aldrich Corp	DD	DuPont
BMS	Bemis Company	OI	Owens-Illinois
MWV		ECL	Ecolab Inc.
SEE	Sealed Air	X	United States Steel Corporation
PPG	PPG Industries	AA	American Airlines Group
BLL	Ball Corp	WY	Weyerhaeuser Corp.
IFF	Intl Flavors and Fragrances	NEM	Newmont Mining Corporation
NUE	Nucor Corp.	FMC	FMC Corporation
HON	Honeywell Int'l Inc.	PH	Parker-Hannifin
RTN	Raytheon Co.	CAT	Caterpillar Inc.
DOV	Dover Corp.	EFX	Equifax Inc.

<b>Símbolo</b>	<b>Compañía</b>	<b>Símbolo</b>	<b>Compañía</b>
FLS	Flowserve Corporation	FDX	FedEx Corporation
EMR	Emerson Electric Company	AVY	Avery Dennison Corp.
CSX	CSX Corp.	CTAS	Cintas Corporation
PBI	Pitney Bowes Inc	PCAR	PACCAR Inc.
TXT	Textron Inc.	CMI	Cummins Inc.
DE	Deere and Co.	LUV	Southwest Airlines
LMT	Lockheed Martin Corp.	BA	Boeing Company
R	Ryder Systems Inc.	JEC	Jacobs Engineering Group
ITW	Illinois Tool Works	NSC	Norfolk Southern Corp.
UNP	Union Pacific	NOC	Northrop Grumman Corp.
PCP	Precision Castparts Corp.	FAST	Fastenal Co
UTX	United Technologies	EXPD	Expeditors International
DNB	Dun and Bradstreet	GE	General Electric
GWW	Grainger (W.W.) Inc.	DHR	Danaher Corp.
MAS	Masco Corp.	GD	General Dynamics
ETN	Eaton Corporation	ROK	Rockwell Automation Inc.
RRD	R.R. Donnelley	ODP	Office Depot Inc.
IGT	International Game Technology	BBY	Best Buy Co. Inc.
LEG	Leggett and Platt	FDO	Family Dollar Stores Inc.
SHW	Sherwin-Williams	NWL	Newell Brands
MCD	McDonald's Corp.	OMC	Omnicom Group
TJX	TJX Companies Inc.	GCI	
MDP		CMCSA	Comcast Corp.
MAT	Mattel Inc.	DV	DeVry, Inc.
WHR	Whirlpool Corp.	SWK	Stanley Black and Decker
SPLS	Staples Inc.	PHM	Pulte Homes Inc.
F	Ford Motor	LEN	Lennar Corp.
DIS	The Walt Disney Company	JCP	J.C. Penney
CCL	Carnival Corp.	JWN	Nordstrom
HAR	Harman Int'l Industries	NKE	Nike
HOT	Starwood Hotels and Resorts Worldwide Inc	SNA	Snap-On Inc.
NYT	The New York Times Co.	IPG	Interpublic Group
POM	Pepco Holdings Inc	ED	Consolidated Edison
AES	AES Corp	AEP	American Electric Power
PNW	Pinnacle West Capital	DTE	DTE Energy Co.
STR	Questar Corp.	D	Dominion Energy

<b>Símbolo</b>	<b>Compañía</b>	<b>Símbolo</b>	<b>Compañía</b>
GPS	Gap Inc	HRB	Block HandR
HAS	Hasbro Inc	AZO	AutoZone Inc
BIG	Big Lots Inc.	GPC	Genuine Parts
GT	Goodyear Tire and Rubber	JCI	Johnson Controls International
VFC	V.F. Corp.	TIF	Tiffany and Co.
LOW	Lowe's Cos.	HOG	Harley-Davidson
KBH		TGT	Target Corp.
HD	Home Depot	PEP	PepsiCo Inc.
BF-B	Brown-Forman Corp.	HSY	The Hershey Company
CVS	CVS Health	KR	Kroger Co.
KO	Coca-Cola Company	CPB	Campbell Soup
PG	Procter and Gamble	HRL	Hormel Foods Corp.
SVU	Supervalu Inc	MKC	McCormick and Co.
CAG	Conagra Brands	KMB	Kimberly-Clark
GIS	General Mills	MO	Altria Group Inc
ADM	Archer-Daniels-Midland Co	K	Kellogg Co.
TAP	Molson Coors Brewing Company	CCE	Coca-Cola Enterprises
AVP	Avon Products	COST	Costco Wholesale Corp.
WMT	Wal-Mart Stores	TSN	Tyson Foods
CL	Colgate-Palmolive	CLX	The Clorox Company
SYY	Sysco Corp.	HUM	Humana Inc.
PFE	Pfizer Inc.	UNH	United Health Group Inc.
MYL	Mylan N.V.	CELG	Celgene Corp.
CI	CIGNA Corp.	MDT	Medtronic plc
MRK	Merck and Co.	TMO	Thermo Fisher Scientific
PKI	PerkinElmer	BCR	Bard (C.R.) Inc.
BIIB	Biogen Inc.	LH	Laboratory Corp. of America Holding
AET	Aetna Inc	JNJ	Johnson and Johnson
CVH	Coventry Health Care	BDX	Becton Dickinson
CAH	Cardinal Health Inc.	LLY	Lilly (Eli) and Co.
SYK	Stryker Corp.	XRAY	Dentsply Sirona
STJ	St. Jude Medical	AGN	Allergan Inc
ABT	Abbott Laboratories	VAR	Varian Medical Systems
AMGN	Amgen Inc	THC	Tenet Healthcare
BMY	Bristol-Myers Squibb	BAX	Baxter International Inc.
FHN	First Horizon	PBCT	People's United Financial
TRV	The Travelers Companies Inc.	LM	Legg Mason
AFL	AFLAC Inc	KEY	KeyCorp
AIG	American International Group, Inc.	WM	Waste Management Inc.
STT	State Street Corp.	USB	U.S. Bancorp
MBI		STI	SunTrust Banks
MTB	MandT Bank Corp.	BAC	Bank of America Corp
JPM	JPMorgan Chase and Co.	L	Loews Corp.

<b>Símbolo</b>	<b>Compañía</b>	<b>Símbolo</b>	<b>Compañía</b>
SLM	SLM Corporation	MMC	Marsh and McLennan
HBAN	Huntington Bancshares	PGR	Progressive Corp.
ZION	Zions Bancorp	PNC	PNC Financial Services
CINF	Cincinnati Financial	BBT	BBandT Corporation
LUK	Leucadia National Corp.	HCP	HCP Inc.
PCL	Plum Creek Timber	CMA	Comerica Inc.
C	Citigroup Inc.	WFC	Wells Fargo
PSA	Public Storage	BEN	Franklin Resources
KIM	Kimco Realty	NTRS	Northern Trust Corp.
TMK	Torchmark Corp.	FITB	Fifth Third Bancorp
SCHW	Charles Schwab Corporation	TROW	T. Rowe Price Group
RF	Regions Financial Corp.	AXP	American Express Co
VNO	Vornado Realty Trust	UNM	Unum Group
BK	The Bank of New York Mellon Corp.	CB	CBOE Holdings
XL	XL Capital	LNC	Lincoln National
HST	Host Hotels and Resorts	WDC	Western Digital
ADSK	Autodesk Inc	TSS	Total System Services
AMAT	Applied Materials Inc	INTC	Intel Corp.
AAPL	Apple Inc.	GLW	Corning Inc.
HPQ	HP Inc.	APH	Amphenol Corp
CSCO	Cisco Systems	AMD	Advanced Micro Devices
XLNX	Xilinx Inc	HRS	Harris Corporation
ORCL	Oracle Corp.	TXN	Texas Instruments
ADBE	Adobe Systems Inc	IBM	International Business Machines
FISV	Fiserv Inc	EMC	EMC Corporation
ADP	Automatic Data Processing	XRX	Xerox Corp
ALTR	Altera Corp	TER	Teradyne
SYMC	Symantec Corp.	LLTC	Linear Technology
PAYX	Paychex Inc.	MSFT	Microsoft Corp.
QCOM	QUALCOMM Inc.	CSC	Computer Sciences Corp
ADI	Analog Devices, Inc.	KLAC	KLA-Tencor Corp.
CA	CA, Inc.	MU	Micron TechnologyV
S	Sprint Nextel Corp.	FTR	Frontier Communications
T	ATandT Inc	VZ	Verizon Communications
CTL	CenturyLink Inc	SO	Southern Co.
GAS	AGL Resources	DUK	Duke Energy
EXC	Exelon Corp.	EIX	Edison Int'l
TEG	Integrus Energy Group Inc.	WEC	Wec Energy Group Inc
ETR	Entergy Corp.	SCG	SCANA Corp
CNP	CenterPoint Energy	XEL	Xcel Energy Inc
TE	TECO Energy	PCG	PGandE Corp.
PPL	PPL Corp.	CMS	CMS Energy
NI	NiSource Inc.		





# Bibliografía

- [1] time series. [www.investopedia.com/terms/t/timeseries.asp](http://www.investopedia.com/terms/t/timeseries.asp).
- [2] Jean-Philippe Bouchaud. An introduction to statistical finance. *Physica A: Statistical Mechanics and its Applications*, 313(1):238 – 251, 2002. Fundamental Problems in Statistical Physics.
- [3] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong. Random-matrix physics: spectrum and strength fluctuations. *Rev. Mod. Phys.*, 53:385–479, Jul 1981.
- [4] Elie Cartan. Sur les domaines bornés homogènes de l’espace des variables complexes. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 11(1):116–162, Dec 1935.
- [5] G. Enderlein. Wilks, s. s.: Mathematical statistics. j. wiley and sons, new york–london 1962; 644 s., 98 s. *Biometrische Zeitschrift*, 6(3):214–215, 1964.
- [6] L. Giada and M. Marsili. Data clustering and noise undressing of correlation matrices. *pre*, 63(6):061101, June 2001.
- [7] T. Guhr and B. Kalber. A new method to estimate the noise in financial correlation matrices. *Journal of Physics A Mathematical General*, 36:3009–3032, March 2003.
- [8] Thomas Guhr, Axel Mallerâ Groeling, and Hans A. Weidenmueller. Random-matrix theories in quantum physics: common concepts. *Physics Reports*, 299(4â6):189 – 425, 1998.
- [9] Laurent Laloux, Pierre Cizeau, Jean-Philippe Bouchaud, and Marc Poters. Noise dressing of financial correlation matrices. *Phys. Rev. Lett.*, 83:1467–1470, Aug 1999.
- [10] Jeff Madura. *Financial Markets and Institutions*. Cengage learning, 2014.

- [11] R.N. Mantegna and H.E. Stanley. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, 1999.
- [12] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457, 1967.
- [13] Madan Lal Mehta. *Random Matrices; 3rd ed.* Pure and applied mathematics series. Elsevier, San Diego, CA, 2004.
- [14] Frederic S. Mishkin and Stanley Eakins. *Financial Markets and Institutions (7th Edition) (The Prentice Hall Series in Finance)*. Prentice Hall, 2011.
- [15] Robb John Muirhead. *Aspects of multivariate statistical theory*. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. John Wiley and Sons, New York, 2008.
- [16] M. C. Münnix, T. Shimada, R. Schäfer, F. Leyvraz, T. H. Seligman, T. Guhr, and H. E. Stanley. Identifying states of a financial market. *Scientific Reports*, 2:644, Sep 2012.
- [17] Vasiliki Plerou, Parameswaran Gopikrishnan, Bernd Rosenow, Luis A. Nunes Amaral, Thomas Guhr, and H. Eugene Stanley. Random matrix approach to cross correlations in financial data. *Phys. Rev. E*, 65:066126, Jun 2002.
- [18] Rudi Schäfer, Nils Fredrik Nilsson, and Thomas Guhr. Power mapping with dynamical adjustment for improved portfolio optimization. *Quantitative Finance*, 10(1):107–119, 2010.
- [19] Ruey S. Tsay. *Analysis of Financial Time Series*. Wiley, 2002.
- [20] M. Tumminello, F. Lillo, and R.N. Mantegna. Correlation, hierarchies, and networks in financial markets. *ArXiv e-prints*, September 2008.
- [21] Vinayak and Akhilesh Pandey. Correlated wishart ensembles and chaotic time series. *Phys. Rev. E*, 81:036202, Mar 2010.
- [22] Vinayak, T. Prosen, B. Buca, and T.H. Seligman. Spectral analysis of finite-time correlation matrices near equilibrium phase transitions. *EPL (Europhysics Letters)*, 108:20006, October 2014.

- [23] Vinayak, Rudi Schäfer, and Thomas H. Seligman. Emerging spectra of singular correlation matrices under small power-map deformations. *Phys. Rev. E*, 88:032115, Sep 2013.
- [24] S. Vinayak, Thomas H. Time series, correlation matrices and random matrix models. In *American Institute of Physics Conference Series*, volume 1575 of *American Institute of Physics Conference Series*, pages 196–217, January 2014.
- [25] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62(3):548–564, 1955.
- [26] Eugene P. Wigner. Characteristics vectors of bordered matrices with infinite dimensions ii. *Annals of Mathematics*, 65(2):203–207, 1957.
- [27] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327, 1958.
- [28] John Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, 20A(1/2):32–52, 1928.