[11pt]article
[utf8]inputenc [T1]fontenc [spanish]babel [a4paper,margin=2.5cm]geometry amsmath,amssymb,amsfonts,mathtools amsthm bm booktabs microtype enumitem graphicx hyperref titlesec array longtable multirow siunitx colorlinks=true,linkcolor=blue,citecolor=blue
remark Observación Nota

*

0pt8pt6pt *

0pt6pt4pt *

0pt4pt3pt

Revisión rigurosa (formato LaTeX) de: "Demostración de umbral de correlaciones"

Informe técnico estilo revisor Q1 (Nature) con fórmulas legibles

Fecha: October 8, 2025

Abstract

Este documento entrega una revisión exhaustiva y *autocontenida* de la **demostración del umbral de correlaciones** empleada para construir el grafo funcional de parámetros en la metodología *proxy por centralidades*. Se verifica la validez formal de los argumentos, se cuantifican vacíos y se aportan definiciones, lemas y recomendaciones para elevar el rigor a nivel Q1. Incluye: (i) criterio de conectividad (Erdős–Rényi), (ii) criterio de significancia estadística (Fisher z, comparaciones múltiples), (iii) métrica *información preservada* $I_G(\delta)$ y su conexión con la estructura del grafo, (iv) síntesis y elección de $\delta^\star$, (v) sensibilidad y robustez, y (vi) porcentajes de *eficiencia demostrativa* por sección.

Contents

1 Alcance y diagnóstico general

Objetivo. Justificar un umbral de correlación δ para definir aristas en un grafo G_δ de parámetros $\{\theta_1, \ldots, \theta_p\}$, donde se conecta (θ_i, θ_j) si $|\rho_{ij}| > \delta$, con ρ_{ij} la correlación entre perfiles de sensibilidad S_i, S_j.

Diagnóstico resumido. La demostración combina tres criterios: (i) conectividad de la red, (ii) filtrado de correlaciones espurias por ruido, (iii) preservación de información del modelo (*varianza* de la salida). El planteamiento es sólido, pero identificamos vacíos *subsanables*: falta citar y matizar el uso del umbral de conectividad de Erdős–Rényi; la regla $\delta > \rho_{\text{noise}}/2$ aparece sin justificación formal; $I_G(\delta)$ no está definido con precisión operativa. Este informe llena esos huecos.

2 Preliminares: notación y objetos

Sea un sistema dinámico (posiblemente no lineal en parámetros)

$$\dot{x}(t) = f\big(x(t), u(t), \theta\big), \qquad y(t) = g\big(x(t)\big), \tag{1}$$

con $x(t) \in R^n$, $u(t) \in R^m$, $\theta \in R^p$, $y(t) \in R^q$.

Sensibilidades. Definimos

$$J_\theta(t) = \frac{\partial x(t)}{\partial \theta} \in R^{n\times p}, \qquad S_j(t) = \frac{\partial x(t)}{\partial \theta_j} \in R^n, \tag{2}$$

y para la salida

$$J_y(t) = \frac{\partial y(t)}{\partial \theta} = \frac{\partial g}{\partial x}\big(x(t)\big)\, J_\theta(t). \tag{3}$$

Correlación L^2 temporal. En $[0,T]$: $S_j{}^2_{L^2} = \int_0^T \langle S_j(t),\, S_j(t)\rangle\ \mathrm{d}t$,

$$\rho_{ij} = \mathrm{Corr}(S_i, S_j) = \frac{\displaystyle\int_0^T \langle S_i(t),\, S_j(t)\rangle\ \mathrm{d}t}{S_{i\,L^2}\ \ S_{j\,L^2}}.$$

Grafo funcional con umbral δ.

$$G_\delta = (V, E_\delta), \quad V = \{\theta_1, \ldots, \theta_p\}, \qquad (\theta_i, \theta_j) \in E_\delta \iff |\rho_{ij}| > \delta. \tag{4}$$

Opcionalmente, definimos peso $w_{ij} = |\rho_{ij}|$ y distancia de arco $\ell_{ij} = 1 - w_{ij} + \varepsilon$ con $\varepsilon \in (0, 10^{-6}]$.

3 Criterio de conectividad (Erdős–Rényi)

Definición 1 (Densidad del grafo) *La densidad de G_δ es $\rho_G(\delta) = |E_\delta| p2$.*

Teorema 1 (Umbral de conectividad, forma práctica) *Para p nodos, el umbral asintótico de conectividad de un grafo aleatorio homogéneo es $\rho_{\text{crit}} \approx \ln pp$. Si $\rho_G(\delta) > \ln pp$, la probabilidad de conectividad es alta; por debajo, aparecen vértices aislados con alta probabilidad.*

[Aplicabilidad] Nuestro G_δ *no* es uniforme aleatorio; no obstante, tomamos ρ_{crit} como cota inferior conservadora y verificamos conectividad *empíricamente*. En modelos con estructura modular, densidades superiores a $\ln p/p$ pueden seguir sin garantizar 1 componente; por ello, exigimos además la verificación directa de conectividad (número de componentes $\phi_G(\delta) = 1$).

Ejemplo numérico (ilustrativo). Para $p = 22$ parámetros: $\rho_{\text{crit}} = \ln(22)/22 \approx 0.14$. Por tanto, es razonable requerir $\rho_G(\delta)0.14$ *y* comprobar que $\phi_G(\delta) = 1$.

Hueco detectado y solución. *Hueco*: Uso de $\ln p/p$ sin citar ni matizar su aplicabilidad. *Solución*: Añadir referencia estándar de grafos aleatorios y la observación anterior, dejando claro que es cota orientativa + verificación empírica.

4 Criterio de significancia estadística

Sea T el número de muestras temporales usadas para estimar ρ_{ij}. Usamos la aproximación de Fisher $z = atanh(\hat{\rho})$:

$$z \approx \mathcal{N}\left(atanh(\rho),\ \frac{1}{T-3}\right). \tag{5}$$

Un IC $100(1-\alpha)\%$ para $\rho = 0$ es

$$|\hat{\rho}| \leq \rho_{\text{noise}}(\alpha, T) = \tanh\left(\frac{z_{1-\alpha/2}}{\sqrt{T-3}}\right). \tag{6}$$

Para $T = 31$ y $\alpha = 0.05$, $z_{0.975} = 1.96$ y

$$\rho_{\text{noise}} \approx \tanh\left(\frac{1.96}{\sqrt{28}}\right) \approx 0.37. \tag{7}$$

Comparaciones múltiples

Con $M = p2$ pruebas, Bonferroni lleva a

$$\alpha' = \frac{\alpha}{M}, \qquad \rho_{\text{Bonf}} = \tanh\left(\frac{z_{1-\alpha'/2}}{\sqrt{T-3}}\right), \tag{8}$$

que para $p = 22$ y $T = 31$ produce $\rho_{\text{Bonf}} \approx 0.5$ (demasiado alto). Alternativamente, un control FDR (Benjamini–Hochberg) es menos conservador, pero requiere p-valores individuales de correlación.

Regla práctica y su justificación

Proponemos **filtrar efectos triviales** escogiendo

$$\delta 0.2, \tag{9}$$

apoyado en dos ideas: (i) umbral *cualitativo* de Cohen: $|\rho| < 0.2$ se considera efecto débil; (ii) compromiso entre *no inflar falsos positivos* y *no fragmentar* el grafo; $\delta = 0.2$ filtra ruido débil sin exigir significancia 95% por enlace.

[Cuantificación del riesgo de falsos positivos] Bajo $H_0 : \rho = 0$ y $T = 31$, $P(|\hat{\rho}| > 0.2) \approx 0.27$ por par; con $M = 231$ pares, si *todo* fuese ruido esperaríamos ≈ 62 enlaces falsos. En la práctica, la presencia de correlaciones reales reduce esta cifra efectiva y, lo crucial, el criterio de conectividad (Sec. 3) y la métrica $I_G(\delta)$ (Sec. 5) controlan que la red resultante *represente* la dinámica del modelo. Este balance debe explicitarse en el texto.

Hueco detectado y solución. *Hueco*: la regla $\delta > \rho_{\text{noise}}/2$ (p.ej. 0.18) aparece ad-hoc. *Solución*: Sustituirla por el argumento anterior (Cohen + riesgo) y/o presentar FDR como alternativa; justificar por qué $\delta \approx 0.2$ es un compromiso estable en el caso de estudio y cómo se replicaría en otros.

5 Métrica de información preservada $I_G(\delta)$

Definición 2 (Selección inducida por G_δ) *Para un umbral δ fijo, definimos un* procedimiento *de reducción $\mathcal{S}$:*

[leftmargin=1.2em]

- *Si G_δ es conexo ($\phi_G(\delta) = 1$), tomamos $\Theta_\delta = \{\theta_1, \ldots, \theta_p\}$.*
- *Si G_δ tiene $k > 1$ componentes, tomamos Θ_δ como los parámetros de la **componente gigante** (o, alternativamente, un representante por componente).*

Definición 3 (Información preservada por el grafo) *Definimos*

$$I_G(\delta) = \mathrm{Var}\big(y \,\big|\, \Theta_\delta\big), \tag{10}$$

esto es, la varianza de salida cuando solo los parámetros en Θ_δ varían *(los demás quedan fijados al nominal). Consideramos la fracción*

$$\mathcal{R}(\delta) = \frac{I_G(\delta)}{I_G(0)} \in [0,1], \tag{11}$$

y exigimos $\mathcal{R}(\delta) \geq 0.95$ como criterio de retención.

Proposición 1 (Conectividad y redundancia funcional) *Si G_δ es conexo, cada parámetro está conectado a los demás por cadenas de correlaciones fuertes, lo que sugiere* redundancia *parcial de efectos sobre y. Cuando G_δ se fragmenta ($\phi_G(\delta) = k > 1$), cada componente tiende a aportar influencia* no reproducible *por las otras, por lo que $\mathcal{R}(\delta)$ disminuye si se descartan componentes.*

La proposición alinea el comportamiento de $\mathcal{R}(\delta)$ con la estructura de G_δ. No requiere cuantificar exactamente la contribución de cada componente; basta observar la caída empírica de $\mathcal{R}$ cuando $\phi_G(\delta) > 1$.

Hueco detectado y solución. *Hueco*: $I_G(\delta)$ no estaba definido operacionalmente. *Solución*: incluir las definiciones anteriores y el procedimiento $\mathcal{S}$. Permite reproducibilidad y conecta formalmente grafo $\rightarrow$ métrica de varianza.

6 Síntesis: elección del umbral óptimo $\delta^\star$

Buscamos $\delta^\star$ tal que

$$\delta^\star\ satisface \quad \{\ (C1)\phi_G(\delta^\star) = 1\ y\ \rho_G(\delta^\star) > \ln(p)/p, (C2)\delta^\star 0.2\ (filtraefectosdébiles), (C3)\mathcal{R}(\delta^\star) \geq 0.95\ (r \tag{12}$$

Procedimiento recomendado

[leftmargin=1.2em]

1. Barrer δ en una rejilla (p.ej. 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40).
2. Para cada δ: calcular $|E_\delta|$, $\rho_G(\delta)$, $\phi_G(\delta)$, diámetro y $\mathcal{R}(\delta)$ mediante $I_G(\delta)$.
3. Seleccionar el *mínimo* δ que cumpla (C1)–(C3).

Ejemplo ilustrativo (valores típicos SCT)

δ	$\lvert E_\delta \rvert$	$\rho_G(\delta)$	$\phi_G(\delta)$	$\mathcal{R}(\delta)$
0.10	198	0.858	1	0.998
0.20	87	0.376	1	0.990
0.30	42	0.182	2	0.975
0.40	18	0.078	4	0.932

Con estos valores, $\delta^\star \approx 0.20$ satisface (C1)–(C3) y equilibra *parquedad* de aristas con preservación de información.

Sensibilidad cualitativa. Valores en $[0.15, 0.25]$ suelen comportarse de forma robusta: por debajo de 0.15 aumentan enlaces débiles sin ganancia material en $\mathcal{R}$; por encima de 0.25 tiende a crecer ϕ_G y cae $\mathcal{R}$.

7 Huecos/vacíos detectados y cómo llenarlos

[leftmargin=1.2em]

- **Conectividad sin cita ni matiz.** *Acción:* añadir referencia estándar de umbral $\ln p/p$ y remarcar que es una cota orientativa; siempre verificar $\phi_G(\delta) = 1$ en los datos.
- **Regla $\delta > \rho_{\text{noise}}/2$ ad-hoc.** *Acción:* reemplazar por la justificación (Cohen + riesgo) y/o reportar FDR; cuantificar brevemente el riesgo de falsos positivos a $\delta = 0.2$.

- **Definición ausente de $I_G(\delta)$.** *Acción:* introducir Θ_δ, procedimiento $\mathcal{S}$ y la fracción $\mathcal{R}(\delta)$; explicar su cálculo.
- **Generalidad.** *Acción:* sección breve sobre cómo varían ρ_{crit} y ρ_{noise} con p y T; instrucción para replicar el procedimiento en otros sistemas.

8 Robustez y análisis de sensibilidad

Dependencia en p (número de parámetros)

El umbral teórico de conectividad decrece como $\ln p/p$. Para p grande, (C1) es más fácil de cumplir; (C2) y (C3) dominan la selección de $\delta^\star$.

Dependencia en T (tamaño muestral)

$\rho_{\text{noise}}(\alpha, T)$ decrece con T. Para T alto, el *ruido* aceptable disminuye y podría preferirse $\delta^\star$ levemente menor que 0.2 si (C1) y (C3) se mantienen; para T bajo, ser más conservador con δ puede ser prudente para evitar enlaces espurios.

Estructura modular y heterogeneidad

Incluso con $\rho_G > \ln p/p$, grafos con comunidades pueden no ser conexos. (C1) exige verificar $\phi_G = 1$; si no, ajustar δ o incorporar aristas por *conocimiento a priori* (p.ej. excitación compartida) antes de aplicar centralidades.

9 Integración con "proxy por centralidades" y la demostración de centralidad

Apoyo a la metodología principal

La elección de $\delta^\star$ garantiza que el grafo base sea *conexo* y *limpio de ruido débil*, condiciones necesarias para que medidas de centralidad (especialmente betweenness) identifiquen *puentes estructurales* reales.

Complemento de la teoría de centralidad

La teoría de centralidad (lemas/corolarios en el otro manuscrito) asume un grafo que refleja redundancias funcionales. El presente umbral asegura que aristas representan correlaciones sustantivas, y que $I_G(\delta)$ permanece alto, de modo que la selección basada en centralidades conserva la dinámica principal del modelo.

10 Recomendaciones editoriales y de presentación

[leftmargin=1.2em]

1. Añadir referencias canónicas (Erdős–Rényi; Fisher z; Cohen).
2. Insertar la definición operativa de $I_G(\delta)$ y el procedimiento $\mathcal{S}$.
3. Reportar la tabla de barrido de δ y comentar cada criterio (C1)–(C3).
4. Incluir una breve sección de sensibilidad en p, T y estructura modular.

11 Puntuaciones de rigurosidad y eficiencia por sección

Sección	Rigor (%)	Eficiencia demostrativa (%)
Conectividad (Sec. 3)	**85**	**90**
Significancia (Sec. 4)	**75**	**80**
Información preservada (Sec. 5)	**80**	**85**
Síntesis y $\delta^\star$ (Sec. 6)	**80**	**85**
Global	**78–82**	**85**

12 Conclusiones

La demostración del umbral de correlaciones es conceptualmente sólida y, con las precisiones añadidas (definiciones, matices teóricos y cuantificación de riesgos), alcanza el rigor esperado de una revista Q1. El umbral $\delta^\star \approx 0.2$ equilibra conectividad, limpieza de ruido y retención de información, y habilita la aplicación fiable de centralidades para reducción de parámetros con mínima pérdida de desempeño.

Apéndice A: Intervalo de confianza de ρ (Fisher z)

Sea $\hat{\rho}$ la correlación muestral con T observaciones. Definimos $z = atanh(\hat{\rho}) = 12\ln\big(1+\hat{\rho}1-\hat{\rho}\big)$. Entonces, aproximadamente $z \sim \mathcal{N}\big(atanh(\rho), 1/(T-3)\big)$. Para $H_0 : \rho = 0$, un IC al $1-\alpha$ satisface

$$|z| \leq \frac{z_{1-\alpha/2}}{\sqrt{T-3}} \iff |\hat{\rho}| \leq \tanh\left(\frac{z_{1-\alpha/2}}{\sqrt{T-3}}\right). \qquad (13)$$

Apéndice B: Umbral de conectividad $\ln p/p$

En $G(p, \pi)$ con p nodos y probabilidad de arista π, el umbral asintótico de conectividad ocurre en $\pi \approx \ln p/p$. Para $\pi < (1-\epsilon)\ln p/p$ quedan vértices aislados a.s.; para $\pi > (1+\epsilon)\ln p/p$, el grafo es conexo a.s. En práctica, usamos $\rho_{\text{crit}} \approx \ln p/p$ como cota inferior.

Apéndice C: Pseudocódigo para seleccionar $\delta^\star$

```
INPUT: perfiles S_j(t), j=1..p; rejilla de deltas D; ventana temporal [0,T].
1. Calcular rho_ij = Corr(S_i, S_j) para todos i<j en [0,T].
2. Para cada delta in D:
   2.1 E_delta = {(i,j): |rho_ij| > delta}
   2.2 Construir G_delta=(V,E_delta); calcular densidad rho_G, #componentes phi_G, diámetro
   2.3 Construir Theta_delta según S; calcular R(delta) = I_G(delta)/I_G(0).
3. Elegir el mínimo delta que cumpla: phi_G=1, rho_G > ln(p)/p y R(delta) >= 0.95.
OUTPUT: delta_star.
```