

# Explicabilidad en algoritmos de búsqueda cuántica en hipercubo con valores de Shapley

## Explainability in Quantum Search Algorithms on Hypercube with Shapley Values

**María Cecilia Pezzini**



*Universidad Nacional de La Plata, Buenos Aires, Argentina.*

**Claudia Pons**



*Universidad Nacional de La Plata, Buenos Aires, Argentina*

*Comisión de Investigaciones Científicas de la Provincia de Buenos Aires*

*Universidad Abierta Interamericana, Buenos Aires, Argentina.*

**Luis Mariano Bibbó**



*Universidad Nacional de La Plata, Buenos Aires, Argentina.*

DOI <https://doi.org/10.59471/raia2025224>

**Enviado:** junio 2025. **Aceptado:** octubre 2025. **Publicado:** diciembre 2025

**Como citar:** Pezzini, M. C., Pons, C., & Bibbó, L. M. (2025). Explicabilidad en algoritmos de búsqueda cuántica en hipercubo con valores de Shapley. Revista Abierta De Informática Aplicada, 9(1). <https://doi.org/10.59471/raia2025224>

## Resumen

Este trabajo analiza la explicabilidad del algoritmo de búsqueda basado en caminatas cuánticas acuñadas sobre el hipercubo, integrando la metodología SMEF-E (Shapley–Matrix Explainability Framework – Energy). El enfoque combina teoría de juegos cooperativos con funciones de valor Hamiltonianas, con el fin de atribuir la contribución funcional y energética del oráculo, la moneda de Grover y el operador flip-flop durante la evolución del algoritmo. La descomposición mediante valores de Shapley permite interpretar de manera cuantitativa cómo se genera la interferencia constructiva y cómo se redistribuye la energía a medida que se alcanza la probabilidad de éxito óptima. Los resultados experimentales validan los modelos teóricos y aportan transparencia sobre los mecanismos internos que sustentan la ventaja cuántica en búsqueda espacial.

**PALABRAS CLAVES:** caminatas cuánticas; hipercubo; valores de Shapley; explicabilidad cuántica; Hamiltonianos; ventaja cuántica.

## Abstract

This work explores the explainability of the coined quantum walk search algorithm on the hypercube by incorporating the SMEF-E (Shapley–Matrix Explainability Framework – Energy)

methodology. In this approach, cooperative game theory is combined with Hamiltonian-based value functions to attribute, step by step, the energetic and functional impact of the oracle, the Grover coin, and the flip-flop shift. The resulting Shapley value decomposition offers an interpretable quantification of operator influence throughout the amplification process, revealing how constructive interference emerges and is redistributed as the algorithm approaches its optimal success probability. The experimental analysis confirms theoretical predictions while providing transparent insights into the mechanisms driving quantum advantage in spatial search.

**KEYWORDS:** quantum explainability; Shapley values; coined quantum walk; hypercube search; Hamiltonian-based interpretation; cooperative contribution.

## INTRODUCCIÓN

Los avances recientes en Computación Cuántica (QC) han demostrado una evolución sostenida en el desarrollo de algoritmos capaces de aprovechar los principios de superposición y entrelazamiento para resolver problemas de búsqueda, optimización y simulación con ventaja cuántica respecto a los métodos clásicos (Nielsen & Chuang, 2011; Portugal, 2018). Sin embargo, esta capacidad conlleva un desafío creciente: la opacidad de los procesos y la dificultad para interpretar los resultados obtenidos. La ejecución de un circuito cuántico implica la aplicación de operadores unitarios de alta dimensionalidad sobre estados en superposición, lo que dificulta determinar el rol individual de cada componente en el resultado final.

En este contexto surge la Computación Cuántica Explicable (xQC), una extensión conceptual de la Inteligencia Artificial Explicable (XAI) orientada al ámbito cuántico (Childs & Goldstone, 2004). El propósito de la xQC es aportar transparencia, trazabilidad y auditabilidad a los algoritmos cuánticos, permitiendo comprender cómo cada bloque u operador contribuye a la evolución y al resultado global del circuito.

De manera análoga a las técnicas clásicas de explicabilidad —como LIME o SHAP (Grover, 1996; Shenvi et al., 2003)— la xQC busca atribuir relevancia funcional y energética a los componentes de un sistema cuántico, de forma cuantitativa y coherente con la física subyacente. Distintas líneas de investigación recientes abordan esta problemática desde perspectivas complementarias. Entre ellas se destacan las explicaciones basadas en valores de Shapley aplicadas a redes cuánticas variacionales (SVQX), los enfoques de visualización como QuantumEyes, y las redes lógico-explicadas (Logic-Explained Quantum Networks), que incorporan principios de lógica formal en la inferencia cuántica (Shapley, 1953; Young, 1985).

No obstante, la mayoría de estos enfoques se centran en modelos híbridos cuántico-clásicos y no consideran la dimensión energética y cooperativa entre los operadores del circuito. Para abordar este vacío se propone un marco de análisis denominado *Shapley–Matrix Explainability Framework – Energy* (SMEF-E). El enfoque combina la teoría de juegos cooperativos con la formulación matricial de los circuitos cuánticos, utilizando valores de Shapley para

descomponer la contribución energética y funcional de los bloques constitutivos —oráculo, moneda y operador *shift*—.

**Organización del artículo.** La Sec. 1 presenta los paradigmas del análisis interpretativo; la Sec. 2 detalla los fundamentos teóricos; las Secs. 2.5 y 2.6 describen los operadores básicos; la Sec. 2.7 formaliza la evolución; la Sec. 3 introduce el juego y las funciones de valor Hamiltonianas; la Sec. 4 detalla SMEF-E y el juego por paso; y la Sec. 5 reporta validaciones y descomposiciones. Los símbolos y fórmulas se recopilan en el Apéndice.

## 1. Tendencias actuales en explicabilidad cuántica (xQC)

El estudio de la explicabilidad en computación cuántica (xQC) avanza sobre bases teóricas sólidas de la información y del algoritmo cuántico (Nielsen & Chuang, 2011; Portugal, 2018), y sobre paradigmas de búsqueda que sirven como banco de pruebas para el análisis interpretativo (por ejemplo, Grover y su formulación por caminatas cuánticas) (Childs & Goldstone, 2004; Grover, 1996; Shenvi et al., 2003). En este contexto, se observan tendencias convergentes orientadas a integrar la transparencia al propio ciclo de diseño y evaluación de algoritmos.

Primero, se consolida una línea de **atribución funcional** que traslada principios de la teoría de juegos cooperativos al análisis de circuitos: los *valores de Shapley* permiten descomponer, con criterios de equidad y eficiencia, la contribución de bloques u operadores a un objetivo (p. ej., probabilidad de éxito o variación energética) (Shapley, 1953; Young, 1985). Esta perspectiva aporta métricas cuantitativas y contrastables para auditar el rol de oráculos, monedas y operadores de desplazamiento en algoritmos de búsqueda y optimización.

Segundo, emergen esquemas de **explicabilidad estructural** que se apoyan en la formulación matricial y en subespacios simétricos para describir la dinámica global de circuitos sobre grafos altamente regulares (como el hipercubo), facilitando la lectura de fenómenos de interferencia y concentración de amplitud (Childs & Goldstone, 2004; Portugal, 2018). En estos modelos, la trazabilidad se beneficia de descomposiciones que relacionan operadores unitarios con variaciones observables a lo largo de las iteraciones.

Tercero, se fortalece la **evaluación comparativa** de explicaciones: sobre la base de métricas bien definidas (eficiencia y simetría en atribuciones; complejidad asintótica en búsqueda), se promueve medir la coherencia entre lo que el algoritmo declara “explicar” y lo que efectivamente optimiza (p. ej., éxito en  $O(\sqrt{N})$  para búsqueda no estructurada) (Childs & Goldstone, 2004; Grover, 1996). Este eje busca que la xQC no sea un posprocesado aislado, sino un componente verificable del diseño.

Se advierte una tendencia hacia **marcos integrados** que combinen atribución cooperativa, análisis espectral/geométrico y lectura operacional del circuito, con el objetivo de ofrecer

explicaciones consistentes con la física subyacente y útiles para decisiones de ingeniería (elección de operadores, tiempos de parada, y criterios de auditoría) (Nielsen & Chuang, 2011; Portugal, 2018). Bajo esta visión, la explicabilidad deja de ser un agregado ad hoc y pasa a formar parte del *pipeline* cuántico, desde el modelado hasta la validación empírica.

## 2. Fundamentos teóricos

Con el fin de contextualizar el marco de explicabilidad propuesto, en esta sección se presentan los conceptos que sustentan el algoritmo de búsqueda en el hipercubo. Se describen las propiedades del grafo y su relación con el espacio de Hamming, así como los operadores básicos que intervienen en la dinámica de la caminata cuántica: la moneda de Grover, el operador de desplazamiento *flip-flop* y el oráculo de fase.

### 2.1 Interpretabilidad y Explicabilidad en Computación Cuántica (xQC)

La computación cuántica promete resolver problemas intratables para los sistemas clásicos, como la simulación molecular o la optimización combinatoria. Sin embargo, enfrenta desafíos técnicos asociados a la coherencia cuántica, la corrección de errores y la comprensión de sus procesos internos. En este contexto, la capacidad de interpretar y explicar los modelos cuánticos resulta esencial para su adopción responsable en entornos críticos.

En este marco, la *Computación Cuántica Explicable* (xQC) busca dotar de transparencia a los algoritmos cuánticos, facilitando la identificación de la contribución de cada componente a la salida del sistema. A diferencia de la interpretabilidad —que se centra en la comprensión directa de la estructura interna de un modelo—, la explicabilidad emplea herramientas externas, métricas o visualizaciones para clarificar el comportamiento de sistemas complejos y probabilísticos. Debido a la naturaleza no determinista de la medición, existen regiones del espacio de estados que conforman una *banda de inexplicabilidad*, donde no es posible obtener explicaciones completas o deterministas.

### 2.2 Hipercubo y relación con el espacio de Hamming

El hipercubo  $Q_n$  es un grafo regular de grado  $n$ , con  $2^n$  vértices. Cada vértice puede representarse como un  $n$ -tupla binario:

$$V = (V_1, V_2, \dots, V_n), V_i \in \{0, 1\}.$$

Dos vértices están conectados si difieren en un único bit, es decir, si su distancia de Hamming es uno. Algunas características principales son:

Número de vértices:  $|V| = 2^n$ .

Adyacencia: dos vértices  $u, v \in V$  están conectados si  $d_H(u, v) = 1$ .

Regularidad: cada vértice tiene exactamente  $n$  vecinos, por lo que  $Q_n$  es un grafo regular de grado  $n$ .

Número de aristas:  $|E| = n 2^{n-1}$ .

## 2.3 Matriz de adyacencia y relación con Hamming

La conectividad del hipercubo puede formalizarse mediante su matriz de adyacencia  $A$ , definida por

$$A_{uv} = \begin{cases} 1, & \text{si } d_H(u, v) = 1, \\ 0, & \text{en otro caso.} \end{cases} \quad (1)$$

Es decir,  $A$  codifica directamente la condición de adyacencia basada en la distancia de Hamming.

En la base computacional, la acción de  $A$  sobre un vértice  $|v\rangle$  genera todos sus vecinos:

$$A |v\rangle = \sum_{a=1}^n |v \oplus e_a\rangle, \quad (2)$$

donde  $e_n$  es el vector unitario en la dimensión  $a$  y  $\oplus$  denota la suma bit a bit módulo 2. Así, cada aplicación de  $A$  produce los vértices que se diferencian de  $v$  en exactamente un bit.

De este modo, el hipercubo es isomorfo al grafo de Hamming  $H(n, 2)$ : la métrica  $d_H(u, v)$  mide el número de posiciones en las que difieren dos cadenas binarias y, en consecuencia, organiza el grafo en capas de Hamming. Esta estructura permite reducir el análisis a subespacios simétricos de baja dimensión, lo que facilita tanto el estudio espectral como el diseño de algoritmos de búsqueda eficientes (Portugal, 2018).

## 2.4 Propiedades relevantes del hipercubo para algoritmos de búsqueda cuántica

A continuación, se identifican las características más relevantes para el análisis y diseño de SKW (Portugal, 2018):

- **Alta simetría:** permite reducir problemas multidimensionales a espacios de baja dimensión.
- **Tiempos de arribo óptimos:** la caminata cuántica alcanza el vértice marcado en tiempo proporcional a  $\sqrt{2^n}$ , en concordancia con la complejidad de Grover (Childs & Goldstone, 2004).
- **Propagación controlada:** la amplitud se concentra progresivamente hacia la capa que contiene el vértice objetivo.

- **Adecuado para búsqueda en bases no estructuradas:** ofrece un entorno matemáticamente claro y físicamente realizable para implementar algoritmos de búsqueda cuánticos.

## 2.5 Moneda de Grover

En las caminatas cuánticas acuñadas, uno de los operadores más utilizados para definir la dinámica en el espacio de direcciones es la moneda de Grover. Su nombre proviene del algoritmo de búsqueda de Grover (Grover, 1996), donde se introduce un operador de reflexión respecto del estado uniforme. Este operador se adopta como moneda en caminatas cuánticas. Se define como:

$$G = 2 |D_c\rangle\langle D_c| - I_n, \quad (3)$$

donde  $I_n$  es la identidad en el espacio de la moneda (de dimensión  $n$ ) y  $|D_c\rangle$  es el estado uniforme en todas las direcciones. Sea  $\{|a\rangle\}^n$  la base de la moneda; en el hipercubo  $Q_n$  el grado coincide con  $n$ , por lo que

$$|D_c\rangle = \frac{1}{\sqrt{n}} \sum_{a=1}^n |a\rangle. \quad (4)$$

Para grafos regulares, esta definición es equivalente a la forma matricial (Portugal, 2018):

$$\begin{aligned} G_n &= \frac{2}{n} J_n - I_n, \\ (G_n)_{ij} &= \frac{2}{n} - \delta_{ij}, \end{aligned} \quad (5)$$

donde  $I_n$  es la matriz de unos de tamaño  $n \times n$  y  $\delta_{ij}$  es la delta de Kronecker. Esta forma hace explícito que  $G$  es una reflexión respecto del estado uniforme: las entradas fuera de la diagonal valen  $2/n$ , mientras que la diagonal vale  $2/n - 1$ . En el caso del hipercubo  $Q_n$ , la moneda de Grover se combina con el operador *shift flip-flop* para formar la evolución global (véase Sec. 2.7).

## 2.6 Operador de desplazamiento flip-flop

En las caminatas cuánticas acuñadas, el operador de desplazamiento *flip-flop*, denotado  $S$  (también llamado *flip-flop shift*), define la regla de movimiento entre vértices del grafo y garantiza una evolución unitaria y reversible. Su acción traslada el estado  $|a, v\rangle$  —donde  $v$  es un vértice del hipercubo y  $a$  la dirección (dimensión) asociada al bit que se conmuta— hacia el vértice adyacente en la dirección  $a$ , *preservando la etiqueta de moneda e invirtiendo la dirección* para el paso siguiente. Formalmente,

$$S |a, v\rangle = |a, v \oplus e_n\rangle, \quad (6)$$

donde  $e_n$  es el vector unitario en la dirección  $a$  y  $\oplus$  denota la suma bit a bit módulo 2.

En notación de producto tensorial,  $S$  actúa como un operador controlado por la moneda que actualiza exclusivamente el registro de posición:

$$S = \sum_{a=1}^n |a\rangle\langle a| \otimes \left( \sum_v |v \oplus e_a\rangle\langle v| \right) = \sum_{a=1}^n |a\rangle\langle a| \otimes S_a, \quad (7)$$

donde cada  $S_n$  conmuta el  $a$ -ésimo bit de la posición.

La convención *flip–flop* hace que  $S$  sea una involución unitaria:

$$S^\dagger = S, \quad S^2 = I_n \otimes I_N, \quad (8)$$

donde  $I_N$  es la identidad en el espacio de posiciones (dimensión  $N = 2^n$ ). Esta propiedad asegura reversibilidad y facilita la interferencia constructiva que concentra amplitud en el vértice objetivo (Portugal, 2018; Shenvi et al., 2003).

## 2.7 Evolución global (no marcada y marcada)

De acuerdo con las definiciones de la moneda de Grover (Sec. 2.5) y del operador de desplazamiento *flip–flop* (Sec. 2.6), el paso unitario no marcado de la caminata acuñada en el hipercubo se expresa como:

$$U = S (G \otimes I_N), \quad (9)$$

donde  $S$  es el operador de desplazamiento y  $G$  la moneda de Grover.

Para transformar esta dinámica en un algoritmo de búsqueda, se incorpora un oráculo que actúa sobre el vértice objetivo  $|t\rangle$ . Este oráculo de fase se define como:

$$R' = I - 2 |D_c\rangle\langle D_c| \otimes |t\rangle\langle t|, \quad (10)$$

y tiene el efecto de invertir la fase únicamente en el estado  $|D_c\rangle \otimes |t\rangle$ . El paso marcado del algoritmo resulta entonces:

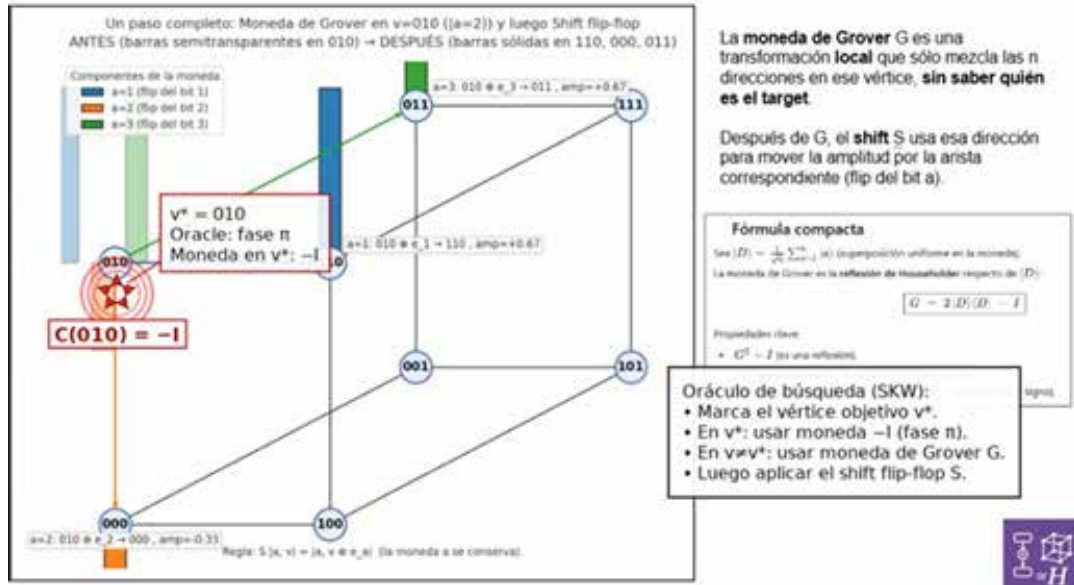
$$U' = UR' = S (G \otimes I_N) R'. \quad (11)$$

Tras  $t$  iteraciones, el estado del sistema es

$$|\psi(t)\rangle = (U')^t |\psi(0)\rangle. \quad (12)$$

Esta combinación garantiza una evolución unitaria y balanceada que, en contextos de búsqueda, logra concentrar amplitud en el vértice objetivo (ver complejidad temporal y expresiones analíticas en Sec. 2.8).

Para clarificar el funcionamiento del conjunto de operadores en el algoritmo de búsqueda en el hipercubo, se presenta un ejemplo visual (Fig. 1) que integra, en un mismo paso, al oráculo, la moneda de Grover y el operador *flip-flop*. En él se observa cómo el vértice objetivo recibe la fase del oráculo, la moneda de Grover actúa como  $-I$  en dicho vértice, y el operador *flip-flop* distribuye la amplitud hacia los vértices adyacentes (véase Sec. 2.6).



**Figura 1.** Ejemplo de un paso del algoritmo SKW en el hipercubo  $Q_n$ . El vértice marcado  $v^* = 010$  recibe la fase del oráculo ( $\pi$ ); la moneda de Grover aplica la reflexión  $I$  sobre el estado objetivo; y el operador *flip-flop* transfiere la amplitud hacia los vértices adyacentes. La secuencia completa muestra cómo la combinación oráculo–moneda–shift produce interferencia constructiva que concentra la amplitud en el objetivo. (Fuente: elaboración propia a partir de (Shenvi et al., 2003)).

## 2.8 Resumen del modelo SKW y predicciones analíticas

Partiendo del estado inicial uniforme  $|\psi(0)\rangle = |DC\rangle |DP\rangle$ , donde  $|DC\rangle$  se definió en la Sec. 2.5 y  $|DP\rangle = \frac{1}{\sqrt{N}} \sum_{v=0}^{N-1} |v\rangle$  es el estado uniforme en posiciones ( $N = 2^n$ ), la evolución de búsqueda se implementa aplicando los operadores definidos en las Secs. 2.5–2.7. En este marco, la probabilidad de éxito de encontrar el vértice objetivo  $|t\rangle$  tras  $t$  pasos se aproxima por (Portugal, 2018):

$$p_{\text{succ}}(t) = \frac{1}{c} \sin^2\left(\frac{2t}{\sqrt{cN}}\right), \quad (13)$$

Con número óptimo de pasos

$$t_{\text{opt}}^{\text{teo}} = \left\lfloor \frac{\pi}{4} \sqrt{cN} \right\rfloor, \quad (14)$$



Y pico esperado

$$p_{\text{máx}} \approx \frac{1}{c}. \quad (15)$$

En régimen asintótico se utiliza  $c \simeq 2$ . Para tamaños finitos, puede refinarse con

$$C = \frac{n}{2N} \sum_{k=1}^n \frac{1}{k} \binom{n}{k}, \quad c_n \simeq 2C, \quad (16)$$

sustituyendo  $c \leftarrow c_n$  en las Ecs. 13–14 para mejorar el ajuste en  $n$  moderados. El modelo SKW mantiene la complejidad temporal  $O(\sqrt{N})$ . Las Ecs. 13–16 proporcionan referencias analíticas directas para la validación de las simulaciones.

### 3. Valores de Shapley

Los valores de Shapley provienen de la teoría de juegos cooperativos y constituyen un mecanismo de reparto de contribuciones que satisface simultáneamente las propiedades de eficiencia, simetría, nulidad y aditividad (Shapley, 1953). Este mecanismo garantiza una asignación justa y única de contribuciones, evita sesgos derivados del orden de incorporación de los elementos Young, 1985, y ofrece una interpretación cuantitativa clara del aporte de cada componente. En contraste, métodos alternativos como gradientes o heurísticas locales no aseguran estas propiedades de equidad ni de unicidad (lundberg2017).

#### 3.1 Aplicación de Hamiltonianos como función de valor

Como *funciones de valor* del juego cooperativo, se emplean observables Hamiltonianos que permiten medir y descomponer las contribuciones de cada bloque del algoritmo de búsqueda cuántica (oráculo, moneda y *shift*) mediante valores de Shapley. Se consideran dos enfoques principales: el Hamiltoniano del problema y el Hamiltoniano energético.

##### 3.1.1 Hamiltoniano del problema

El Hamiltoniano del problema se define como:

$$H_{\text{prob}} = -|t\rangle \langle t|, \quad (17a)$$

$$E_{\text{prob}}(t) = \langle \psi(t) | H_{\text{prob}} | \psi(t) \rangle, \quad (17b)$$

$$P_{\text{succ}}(t) = |\langle t | \psi(t) \rangle|^2, \quad (17c)$$

$$\Delta E_{\text{prob}}(t) = -\Delta P_{\text{succ}}(t). \quad (17d)$$

Aquí,  $|t\rangle$  es el vértice objetivo y  $|\psi(t)\rangle$  el estado tras  $t$  iteraciones. En consecuencia, una disminución de energía equivale a un incremento de probabilidad, por lo que  $H_{\text{prob}}$  funciona como métrica directa del desempeño en términos de localización del objetivo (Portugal, 2018).

### 3.1.2 Hamiltoniano energético

Este observable incorpora tanto el objetivo marcado como la dinámica global de propagación en el grafo (Childs & Goldstone, 2004):

$$H_{\text{ener}} = -\gamma A - |t\rangle\langle t|, \quad (18)$$

donde  $A$  es la matriz de adyacencia del hipercubo y  $\gamma > 0$  regula la tasa de salto entre vértices. Para un estado  $|\psi(t)\rangle$ ,

$$E_{\text{ener}}(t) = \langle \psi(t) | H_{\text{ener}} | \psi(t) \rangle = -\gamma \langle \psi(t) | A | \psi(t) \rangle - |\langle t | \psi(t) \rangle|^2. \quad (19)$$

A diferencia de  $H_{\text{prob}}$ , aquí la energía no es simplemente el negativo de la probabilidad de éxito, sino que combina dos efectos complementarios:

- **Exploración:** propagación en la topología del grafo (término  $-\gamma A$ ).
- **Explotación:** sesgo hacia el objetivo (término  $-|t\rangle\langle t|$ , que reduce la energía al aumentar la amplitud en  $|t\rangle$ ).

## 3.2 Definición formal de los valores de Shapley

Sea el conjunto de bloques  $N = \{O, G, S\}$ ,

correspondiente al oráculo ( $O$ ), la moneda de Grover ( $G$ ) y el operador de desplazamiento flip-flop ( $S$ ). El circuito físico mantiene el orden canónico  $O \cdot G \cdot S$ ; la distinción con el orden de incorporación para Shapley se detalla en Sec. 4.4.

El cálculo de los valores de Shapley considera a los bloques como *jugadores* y evalúa sus contribuciones marginales bajo diferentes órdenes de incorporación. Para un subconjunto  $C \subseteq N$  la función de valor  $v(C)$  representa la variación del observable (probabilidad de éxito o energía) al aplicar únicamente los bloques de  $C$ . La contribución marginal de  $i \in N$  a la coalición  $C$  es

$$\Delta v(i, C) = v(C \cup \{i\}) - v(C). \quad (20)$$

El valor de Shapley de  $i$  promedia estas contribuciones sobre todas las permutaciones  $\pi$  de  $N$ :

$$\phi_i = \frac{1}{|N|!} \sum_{\pi \in \Pi(N)} [v(C_i^\pi \cup \{i\}) - v(C_i^\pi)], \quad (21)$$

□

donde  $C^\pi$  denota los bloques que preceden a  $i$  en la permutación  $\pi$ . Por construcción,

$$\sum_{i \in N} \phi_i = v(N). \quad (22)$$

### 3.3 Ejemplo ilustrativo

Para la permutación " $G \rightarrow O \rightarrow S$ ":

1. Se incorpora  $G$  primero, con  $O$  y  $S$  actuando como identidades:

$$[O = I] - [G = G] - [S = I].$$

2. Luego se añade  $O$ , ocupando su casilla canónica en la primera posición:

$$[O = O] - [G = G] - [S = I].$$

3. Finalmente se incorpora  $S$ , completando el conjunto:

$$[O = O] - [G = G] - [S = S].$$

Así, cambia el orden de incorporación en el cálculo, pero no el orden físico de ejecución del circuito.

### 3.4 Signo de las contribuciones

El signo de los valores de Shapley depende del observable elegido como función de valor:

- Con  $H_{\text{prob}} = -|t\rangle\langle t|$ , se cumple  $E_{\text{prob}}(t) = -P_{\text{succ}}(t)$ ; por tanto,  $\phi < 0$  indica un aumento en la probabilidad de éxito y  $\phi > 0$  una disminución.
- Con  $H_{\text{ener}}$ , un  $\phi > 0$  representa un descenso de energía (estado más alineado con el objetivo), mientras que  $\phi < 0$  refleja un ascenso energético.

## 4. Metodología de explicabilidad (SMEF-E)

Sobre la base de los operadores del algoritmo, la definición de valores de Shapley y los Hamiltonianos utilizados como funciones de valor (Sec. 3.1), se presenta el marco SMEF-E (Shapley–Matrix Explainability Framework–Energy). Este enfoque combina teoría de juegos cooperativos y dinámica cuántica para descomponer y atribuir, de manera justa y cuantitativa,

la contribución de cada bloque (véanse Secs. 2.5–2.7) tanto al desempeño global de búsqueda como a la redistribución energética a lo largo de la evolución.

#### 4.1 Flujo del método SMEF-E

1. **Selección del observable:** escoger  $H_{\text{prob}}$  o  $H_{\text{ener}}$  (definidos en Sec. 3.1) como función de valor para el juego cooperativo.
2. **Construcción de coaliciones y valor:** generar todos los subconjuntos  $C \subseteq \{O, G, S\}$  y evaluar  $v(C)$  según la definición formal de la Sec. 3.
3. **Cálculo de contribuciones:** asignar contribuciones marginales y valores de Shapley  $\phi_i(t)$  conforme a Sec. 3.
4. **Interpretación dinámica:** analizar  $\phi_i(t)$  en cada paso para explicar la cooperación de  $O$ ,  $G$  y  $S$  en la concentración de amplitud y la redistribución energética.

#### 4.2 Definición operacional del juego por paso

Sea  $\psi_t$  el estado al inicio del paso  $t$ . Para cada coalición  $C \subseteq \{O, G, S\}$  se aplican únicamente los operadores presentes en  $C$ , manteniendo identidad en los ausentes y respetando siempre el orden físico canónico  $O \rightarrow G \rightarrow S$ . Con este convenio, la función de valor del juego por paso se define como

$$v(C) = E(\psi_t) - E(\psi_t \xrightarrow{C}), \quad (23)$$

de modo que  $v(\emptyset) = 0$  y  $v(\{O, G, S\}) = E(\psi_t) - E(\psi_t \xrightarrow{OGS})$ . La asignación de Shapley se calcula promediando sobre las seis permutaciones (o, en  $|N|=3$ , mediante la forma cerrada), sin alterar el orden físico del circuito.

#### 4.3 Procedimiento operativo (por paso)

Para cada  $t$  en la vecindad del óptimo teórico (cf. Ec. (14) con Ec. (16)):

1. Evaluar la energía base  $E(\psi_t)$  (con  $H_{\text{prob}}$  o  $H_{\text{ener}}$ ).
2. Para las ocho coaliciones  $C \subseteq \{O, G, S\}$ , aplicar solo  $C$  (en orden  $O \rightarrow G \rightarrow S$ ) y calcular  $E$ .
3. Definir  $v(C) = E(\psi_t) - E(\psi_t \xrightarrow{C})$ . (24)
4. Calcular  $\phi_O(t)$ ,  $\phi_G(t)$ ,  $\phi_S(t)$  (Sec. 3).
5. Verificar la eficiencia por paso (Ec. (27)).
6. Acumular  $\Phi_p \leftarrow \Phi_p + \phi_p(t)$  y comprobar la eficiencia global (Ec. (29)).
7. Actualizar el estado físico:  $\psi_{t+1} = S G O \psi_t$ . (25)

#### 4.4 Orden físico vs. orden de incorporación

El circuito se ejecuta físicamente como  $O \rightarrow G \rightarrow S$  en todos los pasos (Secs. 2.5–2.7). En la asignación de Shapley solo varía el *orden de incorporación* para calcular contribuciones marginales, manteniéndose identidad en los operadores no presentes y sin alterar el orden físico del circuito. Véase también Sec. 3 para un ejemplo ilustrativo.

#### 4.5 Ventajas del método SMEF-E

El marco SMEF-E presenta ventajas frente a enfoques centrados solo en complejidad global o espectro Hamiltoniano (Childs & Goldstone, 2004; Nielsen & Chuang, 2011; Portugal, 2018):

- **Descomposición cuantitativa precisa.** Atribuye exactamente la variación de los observables (probabilidad o energía) a cada bloque.
- **Equidad en la asignación.** Distribución única e independiente del orden de aplicación (valores de Shapley).
- **Explicabilidad dinámica.** Permite seguir  $\phi_i(t)$  y detectar momentos críticos de cada bloque.
- **Detección de anomalías.** Patrones anómalos en las contribuciones pueden señalar oráculos mal definidos o calibraciones inadecuadas de la moneda.

### 5. Resultados de la simulación

En esta sección se presentan los experimentos realizados sobre el algoritmo de búsqueda en el hipercubo basado en caminata cuántica acuñada con moneda de Grover y operador flip-flop. Primero se valida que la implementación reproduce el comportamiento teórico esperado en términos de tiempo óptimo y probabilidad de éxito. Luego, se aplica el marco SMEF-E para analizar, mediante valores de Shapley, las contribuciones de los tres bloques del algoritmo durante la evolución.

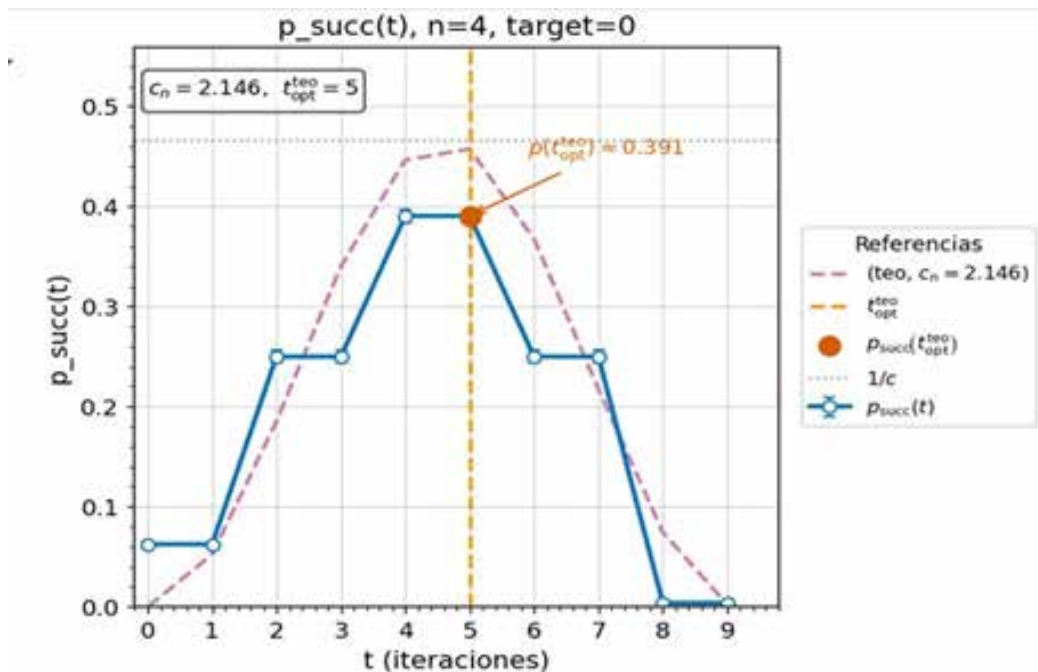
#### 5.1 Validación teórica

Conforme a (Portugal, 2018), el **número óptimo de pasos** es el dado en Ec. (14) y la **evolución de la probabilidad de éxito** sigue Ec. (13); en particular, el máximo teórico está caracterizado por Ec. (15).

Para tamaños finitos, se adopta la corrección  $c \rightarrow c_n \simeq 2C$  de Ec. (16).

Para todas las corridas se fijó  $n \in \{4, 6, 8\}$ , estado inicial uniforme y objetivo  $t = 0$ . Para visualizar la vecindad del máximo, se simuló hasta  $T = t^{\text{teo}} + \Delta$  (con  $\Delta$  pequeño) y se reportó  $p_{\text{succ}}(t)$  en cada paso junto a la referencia vertical en  $t^{\text{teo}}$  (Figs. 2–4). Los resultados de las

Figs. 2 a 4 se sintetizan en la Tabla 1, donde se muestran los parámetros teóricos relevantes para cada caso. En particular, se listan el valor de  $c_n$ , el número óptimo de iteraciones  $t^{\text{teo}}$  obtenido de Ec. (14), la probabilidad en dicho punto  $p(t^{\text{teo}})$  y la cota superior  $1/c_n$  que caracteriza el máximo teórico alcanzable según Ec. (15).



**Figura 2.** Probabilidad de éxito  $P_{\text{succ}}(t)$  para  $n = 4$ , estado inicial uniforme y objetivo  $t = 0$ . En azul se muestra la simulación; en línea punteada rosa, la predicción teórica con  $c_n = 2,146$ . La línea vertical naranja marca el óptimo teórico  $t^{\text{teo}}_{\text{opt}} = 5$ , y punto naranja indica  $P_{\text{succ}}(t^{\text{teo}}) \approx 0,391$ . La línea gris corresponde a la cota  $1/c_n$ .

**CUADRO 1. RESUMEN DE LOS VALORES ÓPTIMOS DE ITERACIONES Y PROBABILIDAD DE ÉXITO EN EL ALGORITMO SKW PARA HIPERCUBOS  $Q_N$  CON  $N = 4, 6, 8$ .**

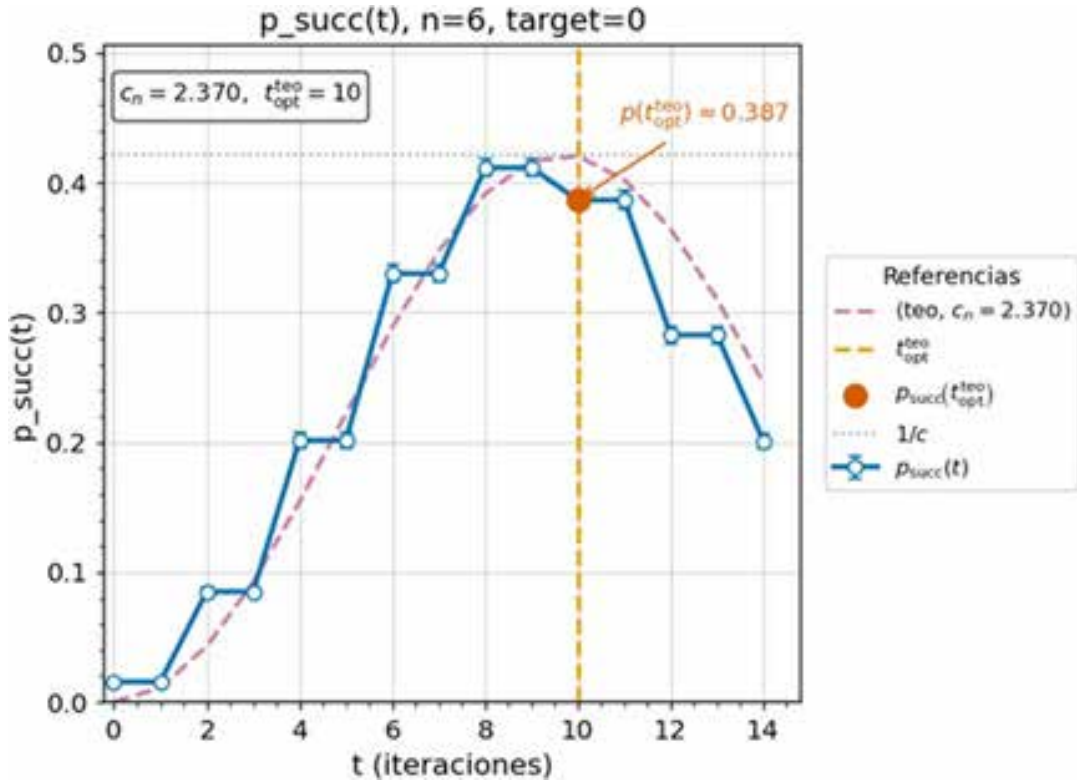
$n$	$c_n$	$t^{\text{teo}}_{\text{opt}}$	$p(t^{\text{teo}}_{\text{opt}})$	$1/c_n$
4	2.146	5	0.391	0.466
6	2.370	10	0.387	0.422
8	2.353	19	0.434	0.425

## 5.2 Validación con el Hamiltoniano del problema

Como control inicial, se verifica que el observable del problema (Sec. 3.1) reproduce la probabilidad de éxito en el modelo SKW: en las simulaciones se observa  $E_{\text{prob}}(t) \approx -P_{\text{succ}}(t)$  (Ec. (17d)) dentro del error numérico. Como referencia teórica, se incluyen la evolución de Ec. (13), el óptimo de Ec. (14) y la cota  $1/c$  asociada al máximo de Ec. (15), junto con los resultados simulados para  $n \in \{4, 6, 8\}$  (Figs. 5–7).

Cuando corresponde, se emplea la corrección de tamaño finito de Ec. (16),

sustituyendo  $c \rightarrow c_n$ . En los experimentos realizados, la relación  $E_{\text{prob}}(t) = -P_{\text{succ}}(t)$  se cumple con error numérico despreciable, lo que confirma que  $H_{\text{prob}}$  constituye un observable consistente y adecuado para evaluar el desempeño del algoritmo en términos de localización del vértice objetivo.



**Figura 3.** Probabilidad de éxito  $p_{\text{succ}}(t)$  para  $n = 6$ , estado inicial uniforme y objetivo  $t = 0$ . En azul se muestra la simulación; en línea punteada rosa, la predicción teórica con  $c_n = 2,370$ . La línea vertical naranja marca el óptimo teórico  $t_{\text{opt teo}} = 10$ , y el punto naranja indica  $p_{\text{succ}}(t_{\text{opt teo}}) \approx 0,387$ . La línea gris corresponde a la cota  $1/c_n$ .

### 5.3 Validación con el Hamiltoniano energético

Como control adicional, se considera  $H_{\text{ener}}$  (Sec. 3.1). En las simulaciones,  $E_{\text{ener}}(t)$  correlaciona positivamente con  $P_{\text{succ}}(t)$  y alcanza su máximo cerca de  $t^{\text{teo}}$ , en concordancia con las referencias teóricas de Sec. 2.8 (véanse pies de figura).

### 5.4 Marco propuesto

La descomposición por valores de Shapley del observable ( $H_{\text{prob}}$  o  $H_{\text{ener}}$ ) se realiza por bloque  $\{O, G, S\}$  en cada iteración  $t$ . En particular,

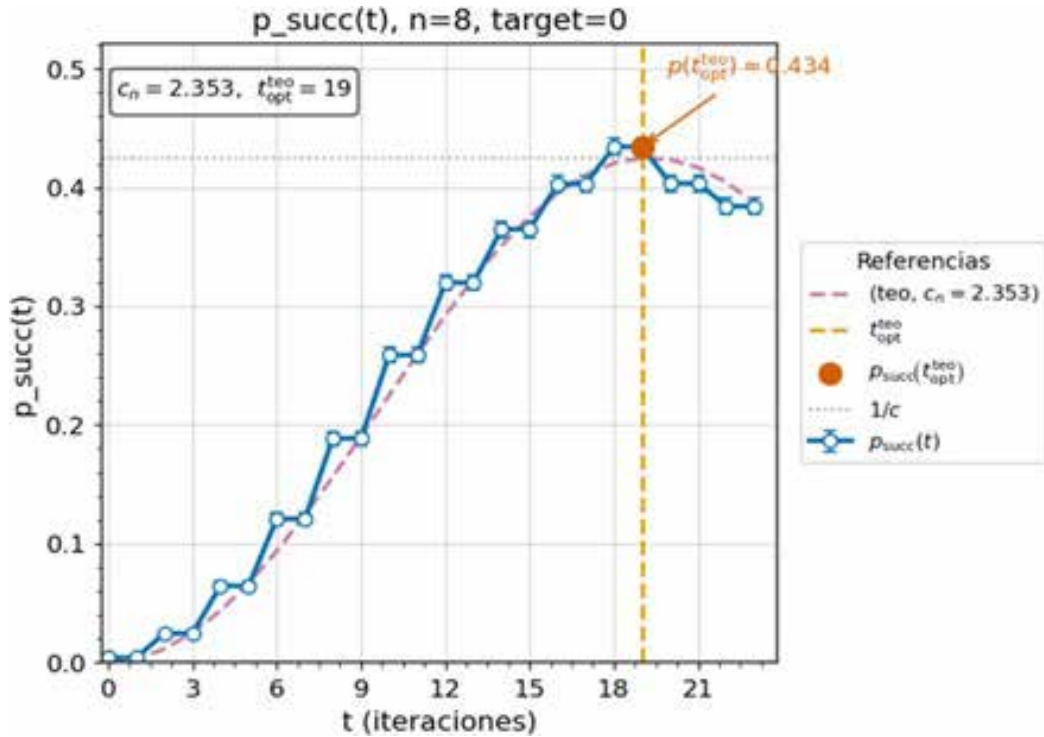
$$\Delta E_t = E(t) - E(t + 1), \quad (26)$$

se distribuye en  $\phi_O(t)$ ,  $\phi_G(t)$  y  $\phi_S(t)$ , verificando *eficiencia por iteración* y *eficiencia global* mediante:

$$\sum_{p \in \{O, G, S\}} \phi_p(t) = \Delta E_t, \quad (27)$$

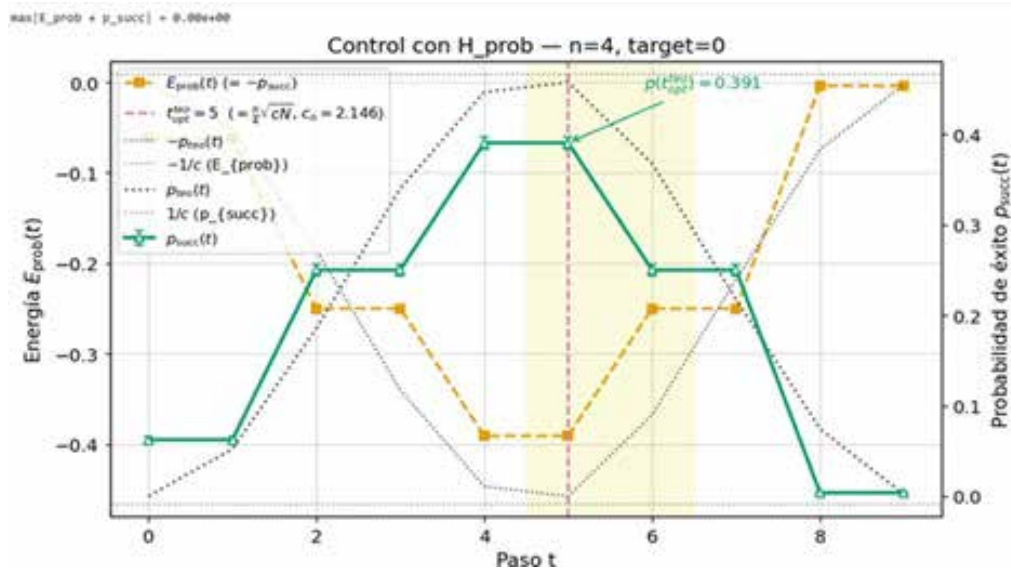
$$\Phi_p = \sum_{t=t_0}^{T-1} \phi_p(t), \quad (28)$$

$$\sum_{p \in \{O, G, S\}} \Phi_p = E(\psi_0) - E(\psi_T). \quad (29)$$



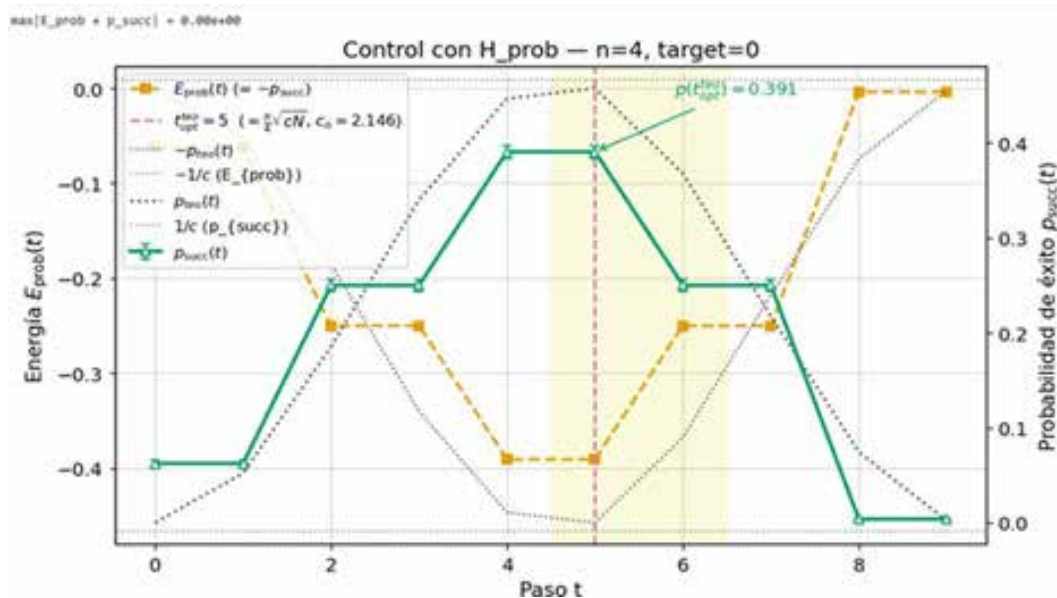
**Figura 4.** Probabilidad de éxito  $p_{succ}(t)$  para  $n = 8$ , estado inicial uniforme y objetivo  $t = 0$ . En azul se muestra la simulación; en línea punteada rosa, la predicción teórica con  $c_n = 2,353$ . La línea vertical naranja marca el óptimo teórico  $t_{opt\_teo} = 19$ , y el punto naranja indica  $p_{succ}(t_{opt\_teo}) \approx 0,434$ . La línea gris corresponde a la cota  $1/c_n$ .





**Figura 5.** Validación de  $H_{\text{prob}}$  para  $n = 4$  (estado inicial uniforme, objetivo  $t = 0$ ). Se grafican  $E_{\text{prob}} t$  y  $p_{\text{succ}} t$  con referencias teóricas (cf. Sec. 2.8).

El reparto de contribuciones se calcula preservando el orden físico de ejecución ( $O \cdot G \cdot S$ ) y la definición de los observables; véase la Sección 3. Las Fig. 11–14 muestran ejemplos representativos



**Figura 6.** Validación de  $H_{\text{prob}}$  para  $n = 6$  (configuración análoga a Fig. 5).

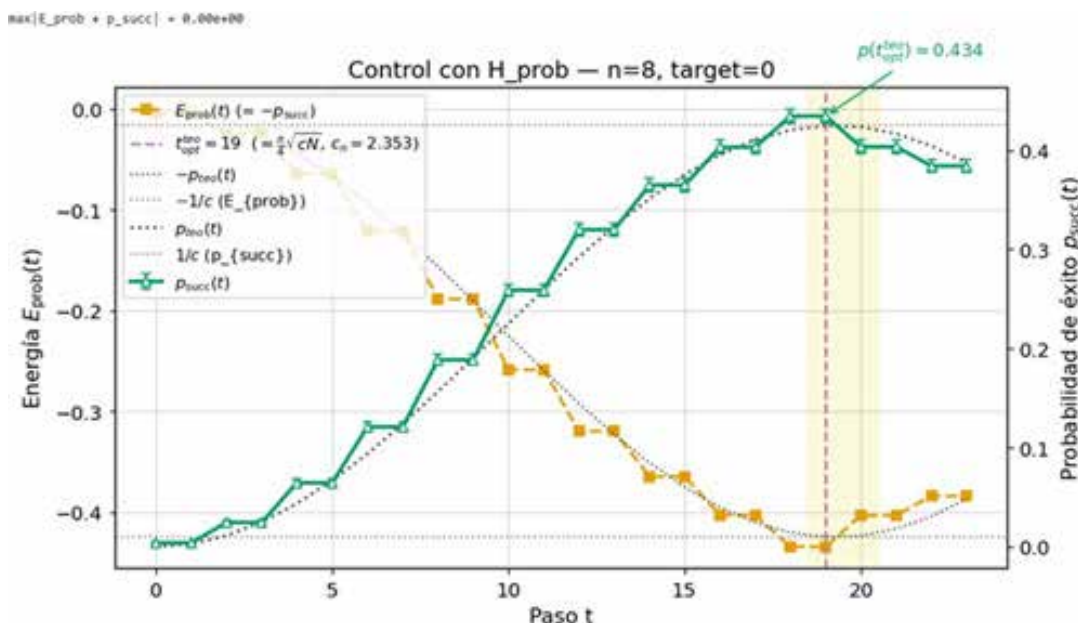


Figura 7. Validación de  $H_{prob}$  para  $n = 8$  (configuración análoga a Fig. 5).

(véanse los pies de figura para  $n$  y el objetivo). Estas verificaciones confirman que la descomposición de SMEF-E es coherente con la evolución del observable (conservación por paso y total), habilitando un análisis explicable de “qué bloque hace qué” cerca del óptimo temporal.

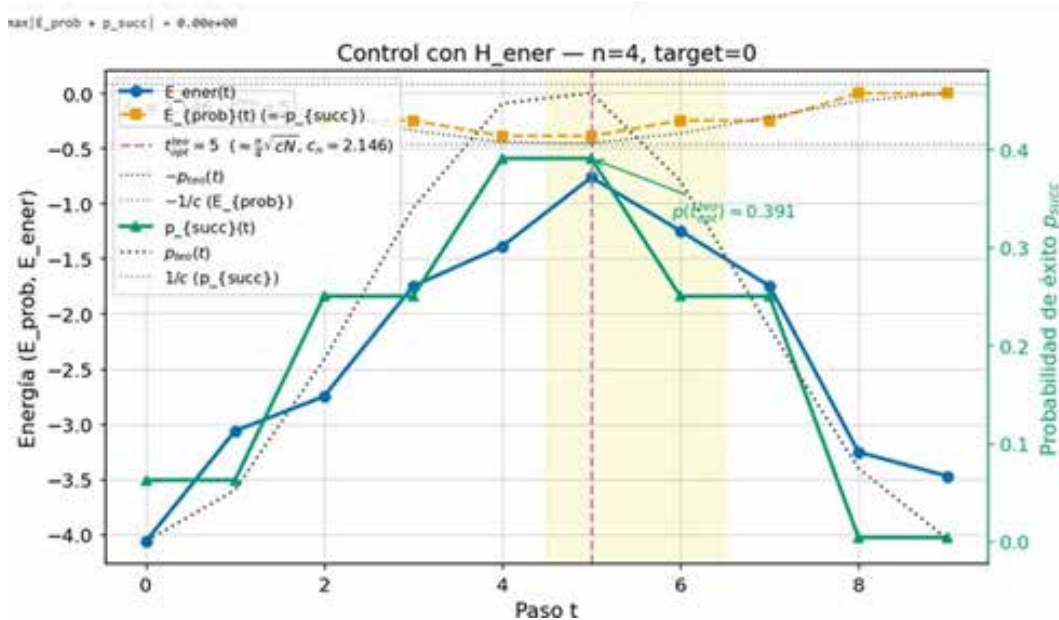


Figura 8. Validación con  $H_{ener}$  para  $n = 4$ . Se muestran  $E_{ener}(t)$ ,  $E_{prob}(t)$  y  $p_{succ}(t)$  con referencias teóricas

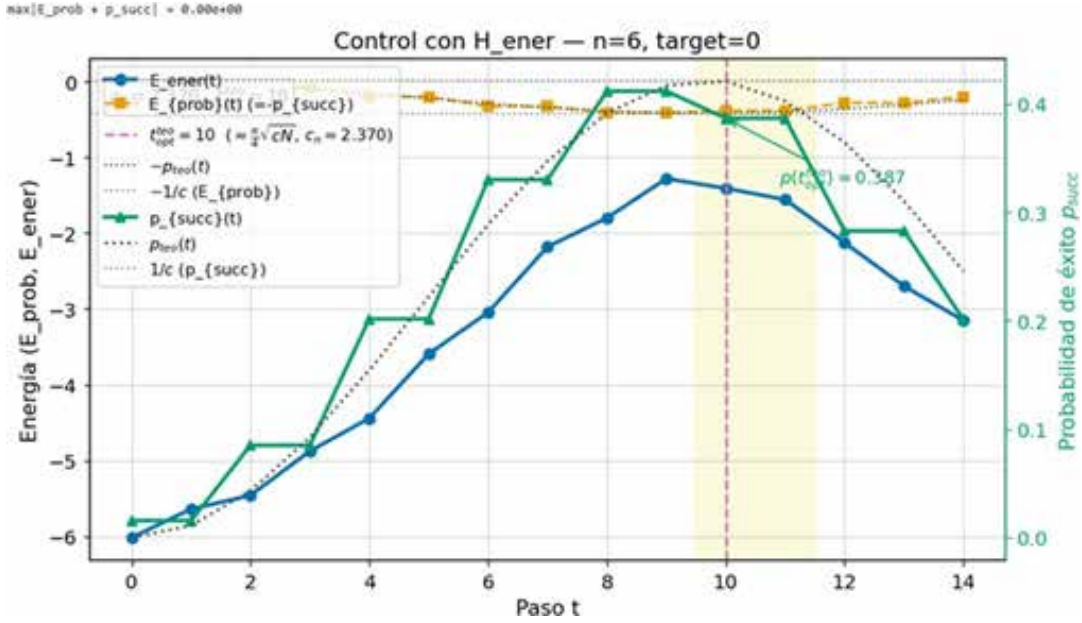


Figura 9. Validación con  $H_{ener}$  para  $n = 6$  (análoga a Fig. 8).

## 5.5 Configuración del experimento

Se empleó el marco SMEF-E para descomponer, paso a paso, la variación del observable mediante valores de Shapley sobre los tres bloques  $\{O, G, S\}$ . Configuración:

- **Tamaño y objetivo.** Ejemplos con  $n \in \{6, 8\}$ ; objetivo  $t = 0$ . (Específico en cada pie de figura.)
- **Observables.** Hamiltoniano energético  $H_{ener}$  y, como control,  $H_{prob}$ .
- **Horizonte temporal.** Zona “sube–pico–baja” alrededor de  $t^{teo} = \pi \sqrt{cN}$ , con  $N = 2^n$  y  $c$  asintótico  $\simeq 2$  o finito  $c_n \simeq 2C$  según (Portugal, 2018).
- **Operadores.** Oráculo tipo marcado  $R'$ , moneda de Grover local, y shift flip-flop estándar; esquema de distancia de Hamming lineal en  $H_{ener}$ .
- **Cálculo de Shapley.** En cada paso  $t$  se evalúa el observable para todas las coaliciones de  $\{O, G, S\}$  y se promedia sobre las 6 permutaciones, garantizando eficiencia.

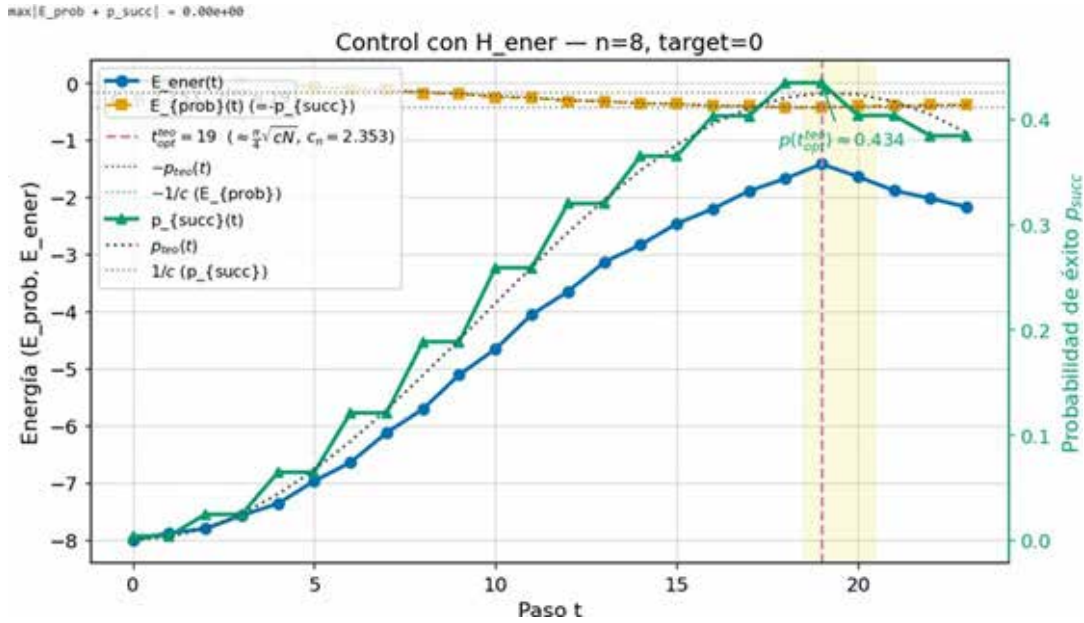


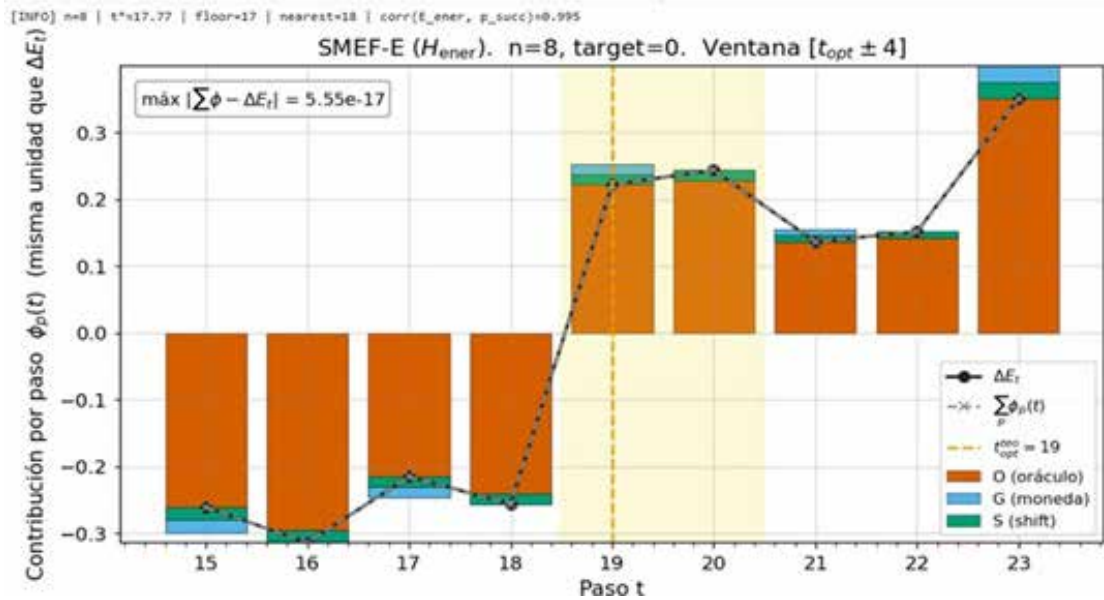
Figura 10. Validación con  $H_{\text{ener}}$  para  $n = 8$  (análoga a Fig. 8).

## 5.6 Resultados con $H_{\text{ener}}$

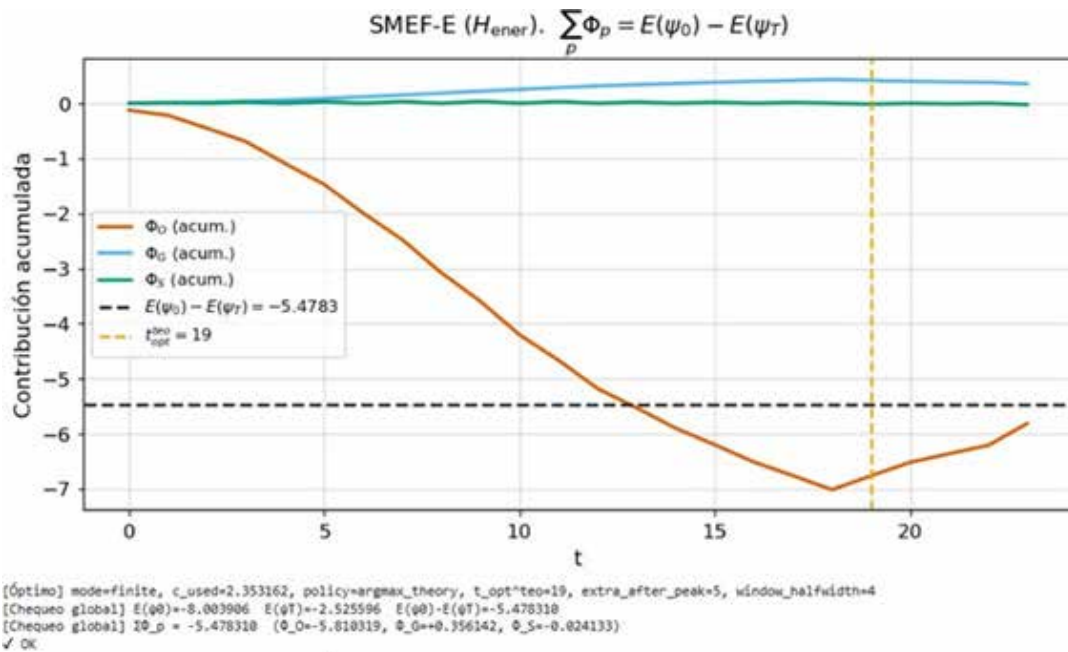
Las Fig. 11 y 12 muestran la coherencia de SMEF-E con  $H_{\text{ener}}$ . En la ventana sombreada, centrada en  $t^{\text{teo}}$  (línea discontinua; cf. Ec. (14)), las barras apiladas indican que el oráculo  $O$  domina la variación  $\Delta E_t$  definida en Ec. (26), mientras que  $G$  y  $S$  introducen correcciones de menor magnitud. La curva gris verifica la conservación por iteración de Ec. (27) (desacople máximo  $< 10^{-16}$  en el ejemplo). A nivel global, las contribuciones acumuladas  $\Phi_p$  en Fig. 12 satisfacen Ec. (29) dentro de la precisión numérica (reportada en el pie de figura).

## 5.7 Control con $H_{\text{prob}}$

Como control ortogonal, se repitió SMEF-E con  $H_{\text{prob}}$  (Fig. 13 y Fig. 14). Se confirma la conservación por paso y global. El patrón es complementario:  $G$  explica el aumento de  $P_{\text{succ}}$ ,  $O$  no contribuye directamente a esta métrica (es phase flip), y  $S$  aporta correcciones menores:  $\Delta E_t = -\Delta P_{\text{succ}}(t)$  (cf. Ec. (17d)).

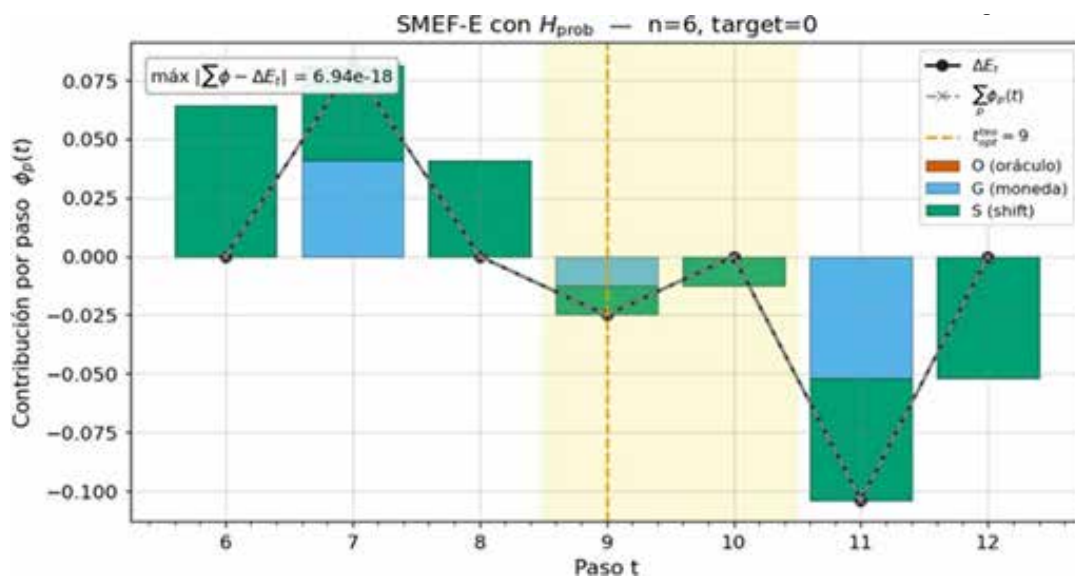


**Figura 11.** SMEF-E con  $H_{\text{ener}}$ : contribuciones por paso alrededor de  $t^{\text{iso}}$  (cf. Ec. (14)). Ejemplo:  $n = 8$ , objetivo  $t = 0$ . La línea negra es  $\Delta E_t$  (Ec. (26)); la gris,  $\sum_p \phi_p(t)$  (igualdad de Ec. (27)).

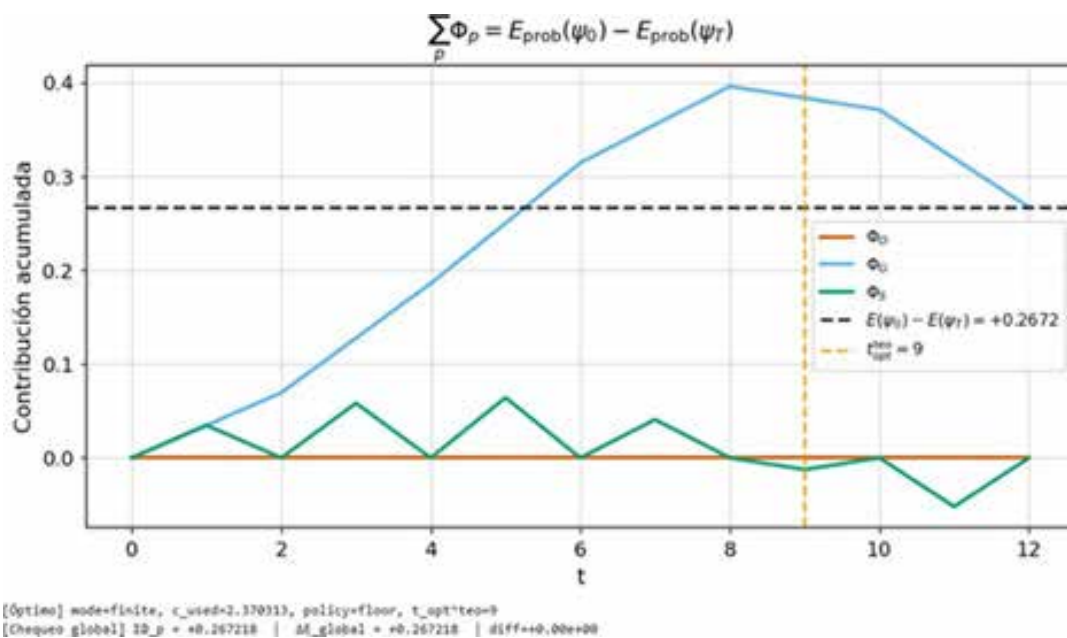


**Figura 12.** SMEF-E con  $H_{\text{ener}}$ : contribuciones acumuladas  $\Phi_p$  y verificación global  $\Phi_p = E\psi_0 - E\psi_T$  (Ec. (29)). Ejemplo consistente con la Fig. 11.





**Figura 13.** SMEF-E con  $H_{\text{prob}}$ : contribuciones por paso; ejemplo  $n = 6$ , objetivo  $t = 0$ . Se verifica  ${}_p \Phi_p(t) \approx \Delta E_t$  con  $\Delta E_t = -\Delta P_{\text{succ}}(t)$ .



**Figura 14.** SMEF-E con  $H_{\text{prob}}$ : contribuciones acumuladas  $\Phi_p$  y chequeo global  $\sum_p \Phi_p = E_{\text{prob}}(\psi_0) - E_{\text{prob}}(\psi_T)$ . En esta métrica domina G.

## CUADRO 2. Símbolos, definiciones y parámetros utilizados en el análisis

Símbolo / fórmula	Definición / significado
$n$	Número de qubits (dimensión del hipercubo / grado de $Q_n$ ).
$N = 2^n$	Número total de vértices del hipercubo.
$ V $	Cardinalidad del conjunto de vértices ( $2^n$ ).
$ E  = n 2^{n-1}$	Número de aristas del hipercubo.
$d_H(u, v)$	Distancia de Hamming: número de bits distintos entre $u$ y $v$ .
$A_{uv} = 1$ si $d_H(u, v) = 1$ ; 0 en otro caso	Matriz de adyacencia de $Q_n$ . Equivalente: $A v\rangle = \sum_{a=1}^n  v \oplus e_a\rangle$ .
$ D_C\rangle = \frac{1}{\sqrt{n}} \sum_{a=1}^n  a\rangle$	Estado uniforme en el registro de <i>moneda</i> .
$ D_P\rangle = \frac{1}{\sqrt{N}} \sum_{v=0}^{N-1}  v\rangle$	Estado uniforme en el registro de <i>posición</i> .
$I_n$	Identidad en el subespacio de la moneda (dimensión $n$ ).
$I_N$	Identidad en el espacio de vértices (dimensión $N$ ).
$G = 2 D_C\rangle\langle D_C  - I_n$	Moneda de Grover: reflexión respecto del estado uniforme de moneda.
$S a, v\rangle =  a, v \oplus e_a\rangle$	Operador shift <i>flip-flop</i> (actúa en posición preservando la etiqueta de moneda).
$U = S(G \otimes I_n)$	Operador global (paso no marcado) de la caminata cuántica acuñada.
$R' = I - 2( D_C\rangle\langle D_C  \otimes  t\rangle\langle t )$	Oráculo de fase para el vértice objetivo $ t\rangle$ .
$H_{\text{prob}} = - t\rangle\langle t $	Hamiltoniano del problema.
$E_{\text{prob}}(t) = \langle \psi(t)   H_{\text{prob}}   \psi(t) \rangle = -p_{\text{succ}}(t)$	Valor esperado de $H_{\text{prob}}$ y relación con $p_{\text{succ}}$ .
$p_{\text{succ}}(t) =  \langle t   \psi(t) \rangle ^2$	Probabilidad de éxito.
$\phi_p(t)$	Valor de Shapley por paso del bloque $p \in \{O, G, S\}$ .
$\Phi_p = \sum_t \phi_p(t)$	Contribución acumulada del bloque $p$ .

## 6. Conclusión

En este trabajo se abordó la explicabilidad en algoritmos de búsqueda cuántica en el hipercubo mediante el marco SMEF-E (Shapley-Matrix Explainability Framework – Energy). A partir de los fundamentos de las caminatas cuánticas y la equivalencia con el algoritmo de Grover (Nielsen & Chuang, 2011; Portugal, 2018; Shenvi et al., 2003), se demostró que los valores de Shapley permiten descomponer de manera justa y cuantitativa las contribuciones de los bloques del algoritmo —oráculo, moneda de Grover y shift flip–flop— tanto en la evolución de la probabilidad de éxito como en la dinámica energética.

Los experimentos realizados validaron que la implementación reproduce el comportamiento teórico esperado: la curva de probabilidad de éxito presenta el patrón ascendente–pico–descendente y el tiempo óptimo escala en concordancia con la complejidad de Grover. Asimismo, se comprobó que SMEF-E satisface la eficiencia local y global, confirmando que la suma de las contribuciones de Shapley coincide con la variación del observable analizado.

El análisis comparativo con los dos Hamiltonianos aporta una lectura complementaria de los roles de los bloques. Con el Hamiltoniano del problema ( $H_{\text{prob}}$ ), la moneda de Grover domina en la amplificación de la probabilidad de éxito, mientras que con el Hamiltoniano energético ( $H_{\text{ener}}$ ) es el oráculo quien concentra el aporte principal, al introducir la marca de fase que estructura la redistribución energética. El shift actúa de forma más moderada,

corrigiendo y sosteniendo la propagación (Burge & et al., 2024). Esta doble perspectiva confirma la coherencia del método y enriquece la interpretación de los mecanismos internos que sustentan la ventaja cuántica.

Además de su valor explicativo, SMEF-E podría emplearse como criterio de auditoría o diagnóstico de anomalías (véase Sec. 4.5), dado que la verificación de eficiencia por paso y global habilita chequeos de consistencia. En este sentido, el marco se proyecta no solo como herramienta de explicabilidad, sino también como criterio de consistencia interna aplicable en hardware cuántico ruidoso, donde la confiabilidad de la ejecución constituye un desafío.

SMEF-E integra teoría de juegos y dinámica cuántica, ofreciendo explicabilidad y transparencia en algoritmos de búsqueda sobre el hipercubo. La elección del hipercubo (SKW) como caso de estudio se justifica por sus propiedades estructurales y analíticas ya establecidas (Secs. 2–2.8), que habilitan la descomposición por bloques requerida por SMEF-E y su validación en el régimen  $O(\sqrt{N})$ , sin introducir supuestos adicionales.

Como líneas futuras, se propone extender el análisis a otros modelos de algoritmos y evaluar su aplicación en hardware real con ruido, con el fin de fortalecer la confiabilidad de la computación cuántica.

## Referencias

- Burge, D., & et al. (2024). Explainability for Quantum Circuits using Functional Attribution. *npj Quantum Information*.
- Childs, A. M., & Goldstone, J. (2004). Spatial search by quantum walk. *Physical Review A*, 70(2), 022314. <https://doi.org/10.1103/PhysRevA.70.022314>
- Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing (STOC '96)*, 212-219. <https://doi.org/10.1145/237814.237866>
- Nielsen, M. A., & Chuang, I. L. (2011). *Quantum Computation and Quantum Information: 10th Anniversary Edition (10th anniversary ed.)*. Cambridge University Press.
- Portugal, R. (2018). *Quantum Walks and Search Algorithms*. Springer. <https://doi.org/10.1007/978-3-319-97813-8>
- Shapley, L. S. (1953). A value for n-person games. En H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the Theory of Games* (pp. 307-317, Vol. 2). Princeton University Press.
- Shenvi, N., Kempe, J., & Whaley, K. B. (2003). Quantum random-walk search algorithm. *Physical Review A*, 67(5), 052307. <https://doi.org/10.1103/PhysRevA.67.052307>
- Young, H. P. (1985). Monotonic solutions of cooperative games. *International Journal of Game Theory*, 14(2), 65-72. <https://doi.org/10.1007/BF01770228>