Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ

Carlos Mario Braga, Aurelio J. Horneros, Manuel A. Serrano, and Eduardo Fernández-Medina

Universidad de Castilla-La Mancha, Ciudad Real, España
CarlosMario.Braga1@alu.uclm.es, aurelio.horneros@uclm.es, Manuel.Serrano@uclm.es, Eduardo.FdezMedina@uclm.es

Resumen Los dispositivos cuánticos ruidosos de escala intermedia (Noisy Intermediate-Scale Quantum, NISQ) imponen restricciones estructurales al entrenamiento de algoritmos variacionales cuánticos (VQAs). En ausencia de corrección completa de errores, cada puerta introduce una probabilidad no nula de fallo que se acumula con la profundidad del circuito, mientras que la estimación del gradiente mediante muestreo finito añade variabilidad estadística adicional. Como consecuencia, la convergencia depende de un equilibrio delicado entre coherencia física y varianza de estimación. En este contexto, la profundidad del ansatz (arquitectura variacional parametrizada) no solo determina su expresividad, sino también su viabilidad operativa bajo ruido. Sin una caracterización explícita de la interacción entre acumulación de error, número de mediciones (*shots*) y estrategias de control, el incremento de recursos experimentales puede no traducirse en mejoras efectivas de rendimiento e incluso resultar contraproducente. En este trabajo analizamos empíricamente cómo la profundidad afecta a la señal práctica del gradiente bajo ruido depolarizante. Observamos un patrón compatible con una atenuación exponencial efectiva de la contribución coherente del gradiente, caracterizable mediante una tasa por capa, ϵ_{eff}, cuya magnitud depende de la complejidad estructural del circuito y refleja principalmente factores físicos y arquitectónicos más que el presupuesto de medición. En conjunto, la profundidad útil del ansatz puede interpretarse como una propiedad emergente del régimen señal–ruido efectivo, gobernada por un parámetro estructural medible más que por un umbral crítico universal. Este marco ofrece una base cuantitativa para orientar decisiones de diseño y control en VQAs bajo condiciones NISQ.

Keywords: optimización variacional cuántica, VQAs, NISQ, atenuación del gradiente, profundidad efectiva

1. Introducción

En los últimos años, los algoritmos variacionales cuánticos (Variational Quantum Algorithms, VQAs) se han consolidado como uno de los enfoques más prometedores para explotar dispositivos cuánticos en el denominado régimen NISQ

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

(Noisy Intermediate-Scale Quantum) [10,1]. El término NISQ hace referencia a procesadores cuánticos de escala intermedia que operan sin corrección completa de errores y están afectados por ruido significativo en cada operación [10]. En este contexto, los VQAs resultan especialmente adecuados, ya que combinan circuitos cuánticos parametrizados con optimización clásica en un esquema híbrido que permite entrenar modelos en presencia de ruido físico significativo sin requerir corrección completa de errores [2,14].

Este paradigma híbrido ha permitido implementar y evaluar tareas concretas como clasificación supervisada, optimización combinatoria y simulación de sistemas cuánticos en dispositivos actuales [5,4,3].

Sin embargo, el entrenamiento de estos modelos presenta dificultades bien documentadas. Entre los problemas más conocidos se encuentra la aparición de *barren plateaus* (mesetas estériles en el paisaje de optimización), en los que la varianza del gradiente puede decrecer exponencialmente con el número de qubits o con la profundidad del circuito, dificultando la optimización [6,2].

La literatura ha mostrado además que la aparición y severidad de estos fenómenos depende críticamente de la arquitectura del ansatz y de la estructura de correlaciones del circuito, pudiendo ciertas configuraciones mitigar o exacerbar la concentración exponencial del gradiente [8,9]. Asimismo, se ha evidenciado que la presencia de ruido físico puede inducir una atenuación exponencial del gradiente incluso en circuitos de profundidad moderada, dando lugar a los denominados *noise-induced barren plateaus* [15].

Además de esta degradación estructural, la estimación estocástica del gradiente mediante muestreo finito introduce una fuente adicional de variabilidad cuya magnitud depende del número de mediciones (*shots*) y del método de optimización empleado [12,13]. En presencia de ruido físico significativo, esta variabilidad estadística puede dominar la dinámica del entrenamiento y dificultar la extracción de señal coherente del gradiente.

En el régimen NISQ, cada puerta cuántica introduce una probabilidad no nula de error, y la ausencia de corrección completa implica que estas imperfecciones se acumulan con la profundidad del circuito [10,1]. La composición sucesiva de canales ruidosos degrada progresivamente la coherencia del estado cuántico y la fidelidad global de la preparación [11]. En un algoritmo variacional, el valor de la función de coste se obtiene como el valor esperado de un observable sobre el estado preparado por el ansatz. El gradiente de entrenamiento depende de cómo varían estas expectativas respecto a los parámetros del circuito. Cuando el ruido reduce la coherencia del estado y lo aproxima a un estado altamente mezclado, las diferencias entre configuraciones parametrizadas se atenúan, lo que disminuye la información útil disponible para la optimización. Esto refuerza la pregunta acerca de hasta qué punto la profundidad del ansatz limita estructuralmente la viabilidad del entrenamiento.

Por otro lado, el número de mediciones por circuito determina la varianza del estimador del gradiente. Bajo supuestos estándar de estimación estadística, la desviación típica del estimador de valores esperados escala como $\mathcal{O}(1/\sqrt{N_{\text{shots}}})$ [7,12], de modo que incrementar el presupuesto de medición reduce la incertidum-

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

bre asociada al muestreo. Sin embargo, cuando la señal coherente del gradiente se encuentra fuertemente atenuada por el ruido físico, esta mejora estadística puede no traducirse en una recuperación efectiva de información estructural relevante [15,1], quedando la dinámica del entrenamiento gobernada por el equilibrio entre señal física degradada y varianza de estimación.

Aunque la literatura ha analizado extensamente la degradación del gradiente con la profundidad [6,15] y, por separado, el impacto de la varianza de estimación asociada al muestreo finito [12], resulta menos claro cómo ambos mecanismos interactúan conjuntamente en presencia de ruido físico para determinar la profundidad práctica alcanzable por un ansatz en un régimen NISQ.

Si bien trabajos previos han evidenciado que el ruido puede inducir atenuaciones exponenciales del gradiente [15] y que el incremento de *shots* mejora la precisión estadística del estimador, no se ha formulado de manera explícita un marco cuantitativo operativo que conecte la acumulación de error por capa, el presupuesto de medición y la viabilidad efectiva del entrenamiento en términos de una profundidad útil emergente.

Motivados por la evidencia teórica y empírica de que la composición sucesiva de canales ruidosos puede inducir una contracción aproximadamente exponencial de las componentes coherentes del estado cuántico [15], proponemos modelar la magnitud coherente efectiva del gradiente como una función aproximadamente exponencial de la profundidad, caracterizada por una tasa de atenuación por capa, ϵ_{eff}. Esta tasa se introduce como un parámetro operativo que resume el efecto combinado del ruido por puerta y de la complejidad estructural del circuito. Bajo este modelo efectivo, la profundidad útil del ansatz puede interpretarse como una propiedad emergente del régimen señal–ruido y de la arquitectura considerada, más que como un umbral crítico fijo independiente del contexto físico.

Las principales contribuciones de este trabajo son

- Se propone un modelo efectivo que describe la degradación estructural del entrenamiento en VQAs bajo ruido NISQ mediante una tasa de atenuación por capa, ϵ_{eff}, que integra efectos físicos y arquitectónicos.
- Se introduce un procedimiento empírico para estimar ϵ_{eff} en simulación bajo distintos niveles de ruido y configuraciones de qubits, permitiendo caracterizar cuantitativamente el régimen señal–ruido efectivo.
- Se formaliza la noción de profundidad útil como una cantidad emergente dependiente de ϵ_{eff} y del presupuesto de medición, proporcionando una interpretación estructural frente a la idea de un umbral crítico universal.
- Se analizan las implicaciones de este marco en el diseño de estrategias de control externo, evaluando su comportamiento relativo bajo distintos regímenes de atenuación estructural.

El resto del artículo se organiza como sigue. La Sección 2 introduce el marco teórico y formaliza la hipótesis de modelado. En la Sección 3 se describe la configuración experimental y el procedimiento estadístico adoptado. La Sección 4

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

presenta los resultados empíricos y su análisis. La Sección 5 discute las implicaciones estructurales del modelo propuesto. Finalmente, la Sección 6 recoge las conclusiones y líneas futuras de investigación.

2. Marco teórico e hipótesis

Consideramos un circuito variacional compuesto por L capas de puertas unitarias parametrizadas. En régimen NISQ, cada operación cuántica está afectada por ruido físico, de modo que la evolución ideal del estado ρ no se implementa exactamente en hardware real. En lugar de ello, el estado sufre una degradación progresiva que reduce las coherencias cuánticas responsables de la interferencia. En ausencia de corrección completa de errores [10], esta degradación se acumula con la profundidad del circuito.

En este trabajo empleamos un modelo de ruido depolarizante independiente por puerta. Este modelo, ampliamente utilizado en estudios teóricos y simulaciones de algoritmos variacionales [5,2], aproxima el efecto agregado de múltiples fuentes de error, como descoherencia e imprecisiones de control, mediante una contracción isotrópica del estado cuántico tras cada operación.

Bajo este modelo, tras cada puerta cuántica el estado evoluciona correctamente con probabilidad $(1-p)$ y, con probabilidad p, se reemplaza por el estado maximamente mezclado. Formalmente, el canal se describe como

$$\mathcal{E}(\rho) = (1-p)\rho + p\frac{I}{d},$$

donde ρ representa el estado ideal sin errores, p es la probabilidad de error por puerta, I es la matriz identidad y d es la dimensión del sistema.

La aplicación repetida de este canal a lo largo de L capas introduce una contracción multiplicativa de las componentes no proporcionales a la identidad del estado. Bajo la hipótesis de independencia por puerta, la contribución coherente del estado tras L capas queda escalada aproximadamente por un factor $(1-p)^L$, lo que induce una dependencia exponencial efectiva con la profundidad del circuito. Este comportamiento es consistente con resultados que muestran atenuaciones exponenciales del gradiente en presencia de ruido físico [15].

Hipótesis de modelado.

A partir del comportamiento multiplicativo inducido por la composición sucesiva de canales ruidosos, se adopta como hipótesis operativa que, en un régimen dominado por ruido físico, la magnitud coherente efectiva del gradiente puede modelarse como una función aproximadamente exponencial de la profundidad,

$$\|\mathbb{E}[\nabla]\|(L) \propto e^{-\epsilon_{\text{eff}} L}.$$

Esta relación se entiende válida en el régimen en el que la pérdida de coherencia domina la dinámica del entrenamiento. En la expresión anterior, $\|\mathbb{E}[\nabla]\|(L)$ denota la magnitud esperada del gradiente asociada a un circuito de profundidad L, y ϵ_{eff} representa una tasa efectiva de atenuación por capa. Este parámetro

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

no se interpreta como una constante universal, sino como una cantidad operativa que resume el efecto combinado del ruido por puerta y de la complejidad estructural del ansatz, incluyendo el número de qubits y la densidad de puertas por capa.

Este modelo se introduce como una descripción efectiva del régimen dominado por pérdida de coherencia. El objetivo central del presente trabajo consiste en estimar empíricamente la tasa ϵ_{eff} y analizar cómo su magnitud varía con la complejidad estructural del circuito.

Por otra parte, en algoritmos variacionales implementados en hardware NISQ, los valores esperados de los observables se estiman mediante muestreo finito. Bajo supuestos estándar de estimación estadística, la desviación típica del estimador de un valor esperado escala como $\mathcal{O}(1/\sqrt{N_{\text{shots}}})$ [7], donde N_{shots} denota el número de mediciones por circuito. En consecuencia, la incertidumbre asociada al gradiente estimado disminuye de forma sublineal con el presupuesto de medición.

Combinando la dependencia exponencial asumida para la magnitud coherente del gradiente con el escalado estadístico del estimador respecto al número de mediciones, puede definirse un cociente señal–ruido (signal-to-noise ratio, SNR) como

$$\text{SNR}(L, N_{\text{shots}}) = \frac{\|\mathbb{E}[\nabla]\|(L)}{\sigma_{\text{est}}(N_{\text{shots}})},$$

donde $\sigma_{\text{est}}(N_{\text{shots}})$ denota la desviación típica del estimador asociada al muestreo finito.

Bajo los supuestos anteriores, $\sigma_{\text{est}} \sim \mathcal{O}(1/\sqrt{N_{\text{shots}}})$ y $\|\mathbb{E}[\nabla]\|(L) \propto e^{-\epsilon_{\text{eff}} L}$, lo que conduce a la relación proporcional

$$\text{SNR}(L, N_{\text{shots}}) \propto \sqrt{N_{\text{shots}}}\, e^{-\epsilon_{\text{eff}} L}.$$

Hipótesis central del trabajo.

Bajo el modelo anterior, se plantea que la profundidad útil del ansatz puede interpretarse como una propiedad emergente del régimen señal–ruido. En particular, se introduce una profundidad efectiva L^* definida como el valor de L para el cual el cociente señal–ruido alcanza un umbral operacional prefijado SNR^*, es decir,

$$\text{SNR}(L^*, N_{\text{shots}}) = \text{SNR}^*.$$

El umbral SNR^* representa el valor mínimo de relación señal–ruido a partir del cual el gradiente se considera operativamente distinguible del ruido estadístico y, por tanto, potencialmente informativo para la optimización.

Sustituyendo la dependencia proporcional obtenida anteriormente y tomando logaritmos, se obtiene una relación aproximada de la forma

$$L^*(N_{\text{shots}}, q) \approx \frac{1}{\epsilon_{\text{eff}}(q)} \left(\frac{1}{2} \log N_{\text{shots}} - C \right),$$

donde el término $\frac{1}{2} \log N_{\text{shots}}$ proviene del escalado estadístico $\sigma_{\text{est}} \sim 1/\sqrt{N_{\text{shots}}}$, y la constante C incorpora tanto el umbral SNR^* como los factores de escala in-

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

dependientes de N_{shots}, tales como la amplitud inicial del gradiente y la varianza intrínseca del estimador.

Intuitivamente, este modelo sugiere que, para una tasa de atenuación dada ϵ_{eff}, el aumento del número de *shots* solo incrementa de forma logarítmica la profundidad efectiva alcanzable. Cuando la profundidad L supera el valor L^*, la contribución coherente del gradiente se encuentra fuertemente atenuada, de modo que incrementos adicionales en el presupuesto de medición reducen la varianza del estimador, pero no recuperan la estructura física degradada por el ruido.

Por el contrario, cuando $L < L^*$, la señal coherente aún es suficientemente significativa como para que la reducción de varianza mediante más *shots* se traduzca en mejoras efectivas del entrenamiento.

En este marco, L^* no constituye una propiedad fija del ansatz, sino una cantidad dependiente del número de qubits, de la complejidad por capa,que determina ϵ_{eff}, y del número de mediciones por circuito.

El objetivo empírico del presente trabajo consiste en estimar la tasa efectiva ϵ_{eff} y analizar cómo varía con la complejidad estructural del circuito, evaluando hasta qué punto este modelo explica el comportamiento observado bajo distintas configuraciones de profundidad y presupuesto de medición.

En la práctica experimental no se accede directamente a $\|\mathbb{E}[\nabla]\|(L)$, sino a estimadores del gradiente afectados por ruido de muestreo. Bajo una descomposición aditiva del tipo $\hat{g} = s_L + \eta$, donde s_L representa la contribución coherente estructural y η el ruido estadístico asociado al muestreo, la varianza observada del gradiente puede reflejar indirectamente la magnitud de s_L siempre que esta no quede completamente dominada por η. En este sentido, la varianza media del gradiente en las últimas iteraciones del entrenamiento se emplea como proxy operacional de la atenuación estructural modelada por ϵ_{eff}, bajo el supuesto de que la contribución coherente del gradiente no está completamente dominada por el ruido de estimación, de modo que la varianza observada refleja indirectamente la magnitud de dicha señal estructural.

3. Configuración experimental

Para contrastar empíricamente la hipótesis anterior, se realizaron simulaciones con Qiskit Aer bajo un modelo de ruido depolarizante independiente por puerta. Se consideraron probabilidades de error diferenciadas para puertas de un qubit y de dos qubits, aplicadas respectivamente a rotaciones y compuertas `cx`.

Se consideraron tres regímenes de ruido. El régimen `none` corresponde a simulación ideal sin errores. En `light` se fijaron probabilidades $p_1 = 0.005$ para puertas de un qubit y $p_2 = 0.01$ para puertas de dos qubits, mientras que en `heavy` estas se incrementaron a $p_1 = 0.02$ y $p_2 = 0.05$.

Dado que en hardware real las puertas de dos qubits presentan mayores tasas de error, se consideraron valores $p_2 > p_1$ en ambos regímenes. Estos parámetros

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

no pretenden reproducir un dispositivo específico, sino modelar de forma controlada distintos niveles de acumulación de error característicos del régimen NISQ.

El análisis principal se centra en el régimen `heavy`, donde la acumulación de error por puerta permite observar una degradación estructural del gradiente. Los regímenes `none` y `light` se emplearon como controles para verificar la ausencia del fenómeno en condiciones de ruido nulo o moderado.

Se empleó un ansatz *hardware-efficient* con rotaciones parametrizadas por qubit y compuertas `cx` entre qubits adyacentes. La tarea fue clasificación binaria sintética (`make_moons`) con entradas reescaladas a $[0, 2\pi]$.

El entrenamiento se realizó mediante SPSA, seleccionado por su robustez frente a gradientes ruidosos y su bajo coste en evaluaciones del circuito, características relevantes en régimen NISQ con alta varianza de estimación [13]. Sobre este optimizador se implementó un controlador externo con capacidad de intervención adaptativa durante el entrenamiento. En este contexto, el gradiente no se calcula mediante reglas analíticas como parameter-shift, sino que se estima mediante perturbaciones estocásticas simultáneas propias de SPSA, lo que permite reducir el número de evaluaciones del circuito necesarias en presencia de ruido.

En cada ejecución se registró la varianza media del gradiente estimado durante las últimas K iteraciones del entrenamiento, denotada por $\bar{V}_{\nabla}^{(K)}$ (implementada como `avg_grad_var_lastK`). En los experimentos se fijó $K = 10$. Esta magnitud se emplea como proxy operacional de la señal estructural del gradiente y sirve de base para la estimación empírica de la tasa efectiva de atenuación ϵ_{eff}.

Para cada configuración se evaluaron tres modos de entrenamiento:

- `baseline_noisy`: entrenamiento sin intervención externa.
- `OF_noisy_shots`: incremento adaptativo del número de *shots* ante detección de estancamiento.
- `OF_noisy_inject`: intervención adaptativa basada en la varianza estimada del gradiente y la evolución de la pérdida.

En el modo `OF_noisy_inject`, el controlador distingue entre un régimen dominado por ruido de estimación y uno de posible debilitamiento estructural del gradiente. En el primer caso se incrementan los *shots*; en el segundo se activan mecanismos de exploración paramétrica mediante perturbación gaussiana

$$\theta \leftarrow \theta + \mathcal{N}(0, \sigma^2_{\text{inject}}),$$

o, cuando la planitud es más pronunciada, reinicialización parcial de las capas más profundas del ansatz.

Aunque el controlador puede ejecutar tres acciones, duplicación de *shots*, perturbación aditiva y reinicialización parcial, su lógica se organiza en dos regímenes. Uno está orientado a la reducción de varianza cuando la limitación es esencialmente estadística y el otro a la exploración paramétrica cuando la limitación es estructural. La transición entre ambos se determina mediante umbrales sobre la varianza estimada del gradiente. Cuando la varianza se sitúa en un rango intermedio se aplica perturbación gaussiana con desviación típica σ_{inject}. Cuando

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

cae por debajo del umbral de planitud se activa la reinicialización parcial de los últimos n parámetros con desviación σ_{reinit}.

Los parámetros del controlador se fijaron *a priori*. Se utilizó un parámetro de paciencia de 5 épocas para detectar estancamiento, junto con un umbral mínimo de mejora de 10^{-4} y un valor objetivo de pérdida de 0.28.

Los umbrales de varianza se fijaron en $g_{\text{noise}} = 5 \times 10^{-5}$ y $g_{\text{flat}} = 10^{-6}$, que delimitan respectivamente los regímenes dominados por ruido estadístico y por posible planitud estructural.

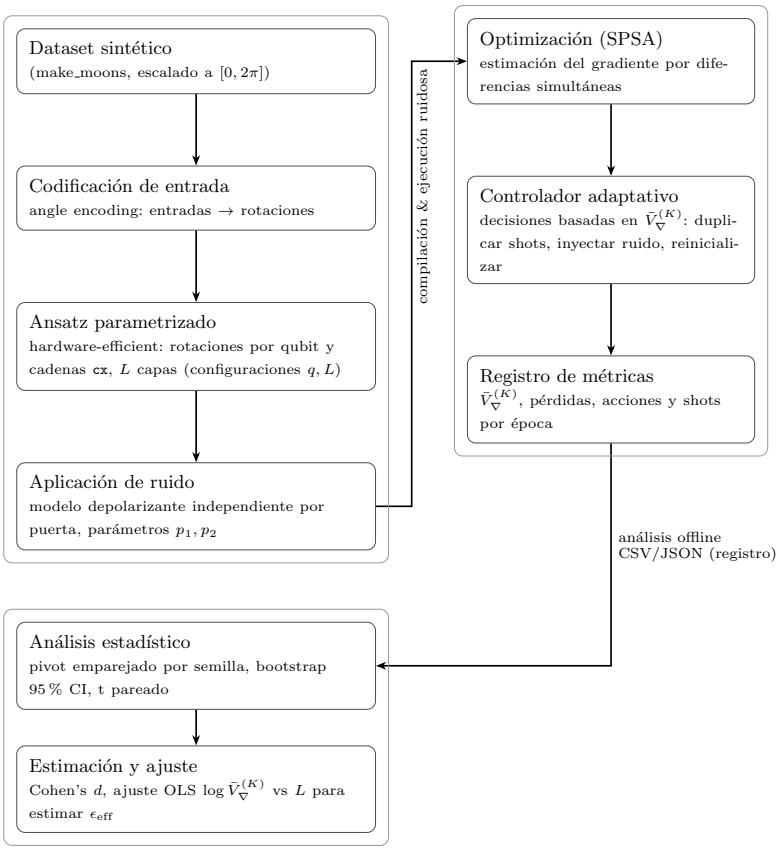


Figura 1: Flujo experimental. Arriba a la izquierda: preparación del circuito y modelo de ruido. A la derecha: optimización (SPSA), controlador adaptativo y registro de métricas. Abajo: análisis estadístico y estimación offline de la tasa efectiva ϵ_{eff}. Las decisiones del controlador se mantuvieron fijas durante las ejecuciones; el análisis posterior no alimentó el proceso online.

Se estableció un máximo de 1024 mediciones por circuito. En el régimen exploratorio se reinicializaron los últimos $n = 4$ parámetros y se emplearon

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

Esta obra está bajo una licencia Creative Commons "Atribución-NoComercial-SinDerivadas 4.0 Internacional".

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

desviaciones típicas $\sigma_{\text{inject}} = 0.01$ y $\sigma_{\text{reinit}} = 0.02$. En los experimentos se fijó además $K = 10$ para el cálculo del proxy $\bar{V}_{\nabla}^{(K)}$.

El diseño experimental, cuyo flujo completo se resume en la Figura 1, se organizó en torno a la hipótesis estructural del trabajo y no como un barrido simétrico exhaustivo de todos los factores. Se analizaron configuraciones con $q \in \{4, 5\}$ y profundidades crecientes del ansatz. Para $q = 5$ se consideró $L \in \{2, 3, 4, 5, 6\}$, lo que introduce un nivel elevado de complejidad por capa debido al mayor número de puertas. Para $q = 4$ se extendió el rango hasta $L \in \{7, 8, 9\}$ con el fin de alcanzar niveles de complejidad estructural comparables y evaluar la posible emergencia de un régimen de atenuación efectiva del gradiente.

De forma exploratoria se evaluó también la configuración con $q = 6$ bajo el régimen `heavy`, manteniendo un rango reducido de profundidades debido al mayor coste computacional.

El presupuesto inicial de medición se fijó en $N_{\text{shots}} \in \{16, 32, 60, 128\}$. Cada combinación de parámetros se ejecutó con múltiples semillas independientes. En cada semilla se generó el conjunto de datos sintético y se fijaron los generadores pseudoaleatorios del entorno de simulación y del optimizador, garantizando reproducibilidad y permitiendo análisis pareados entre estrategias.

Los resultados por ejecución se almacenaron en ficheros estructurados que incluyen métricas finales de rendimiento, diferencias emparejadas respecto al `baseline`, número de intervenciones del controlador y estadísticas del gradiente.

El análisis estadístico incluyó comparaciones emparejadas de precisión entre estrategias por semilla, estimación de medias e intervalos de confianza bootstrap al 95 %, pruebas t pareadas bilaterales y cálculo de tamaños de efecto mediante el estadístico de Cohen d. Asimismo, se realizó un ajuste lineal de $\log \bar{V}_{\nabla}^{(K)}$ en función de L para estimar ϵ_{eff}.

El código utilizado para la generación de datos, ejecución de experimentos y análisis estadístico se encuentra disponible públicamente en el repositorio `https://github.com/GSYAtools/QML`, con el objetivo de facilitar la reproducibilidad de los resultados que se presentan en la siguiente sección.

4. Resultados

El análisis se centra en la estimación de la tasa efectiva de atenuación del gradiente ϵ_{eff} y en la evaluación del comportamiento relativo de las estrategias de control bajo distintos regímenes de profundidad y ruido.

Para contrastar la hipótesis

$$\|\mathbb{E}[\nabla]\|(L) \propto e^{-\epsilon_{\text{eff}} L},$$

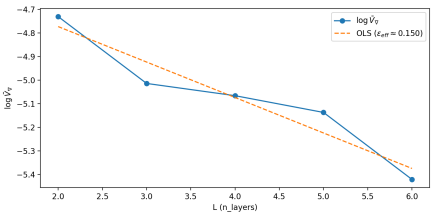
se analizó la dependencia de $\log \bar{V}_{\nabla}^{(K)}$ con la profundidad L, centrándose en el régimen `heavy`, donde la acumulación estructural de error es más pronunciada. La estimación de ϵ_{eff} se obtuvo mediante ajuste lineal OLS del modelo

$$\log \bar{V}_{\nabla}^{(K)} = \beta_0 + \beta_1 L,$$

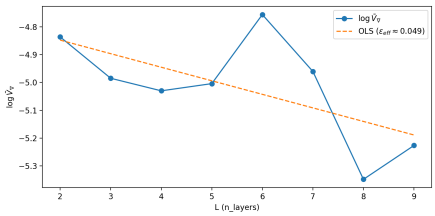
Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

con $\beta_1 = -\epsilon_{\text{eff}}$, contrastando $H_0 : \beta_1 = 0$ mediante estadístico t bilateral.

Las Figuras 2a y 2b muestran ejemplos representativos para 60 *shots* en $q = 5$ y $q = 4$.



(a) $q = 5$, 60 shots, `heavy`.



(b) $q = 4$, 60 shots, `heavy`.

Figura 2: Logaritmo de la varianza media del gradiente, $\log \bar{V}_\nabla$, en función de L. Las rectas corresponden al ajuste lineal OLS.

Para $q = 5$ se observa una relación aproximadamente lineal en escala logarítmica, consistente con una atenuación exponencial efectiva. Las estimaciones obtenidas para distintos presupuestos de medición se sitúan en $\epsilon_{\text{eff}} \approx 0.14$–$0.17$, con R^2 entre 0.30 y 0.40 y pendientes altamente significativas ($p < 10^{-3}$). La estabilidad frente a variaciones en el número de *shots* sugiere que ϵ_{eff} captura una propiedad estructural del circuito.

En $q = 4$ la pendiente estimada es positiva pero considerablemente menor, con $\epsilon_{\text{eff}} \approx 0.03$–$0.04$, R^2 del orden de 0.04–0.05 y significación marginal. Esto indica que la atenuación estructural es más débil y difícilmente detectable en configuraciones de menor complejidad, en un régimen donde la señal coherente del gradiente es comparable al ruido de estimación, lo que dificulta la identificación empírica de una tendencia exponencial clara.

De forma exploratoria se evaluó también la configuración $q = 6$ bajo `heavy`. Las estimaciones de ϵ_{eff} resultaron positivas y de magnitud intermedia, pero con coeficientes de determinación reducidos y ausencia de significación estadística robusta con el número de ejecuciones considerado. Estos resultados son compatibles con el marco estructural propuesto, aunque deben interpretarse como evidencia preliminar.

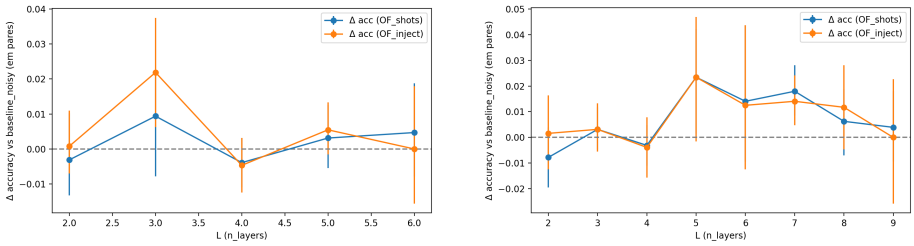
En conjunto, los resultados para $q = 4$ y $q = 5$ apoyan la hipótesis de que la tasa efectiva de atenuación aumenta con la complejidad estructural del circuito.

La Figura 3 muestra la diferencia media de precisión entre estrategias adaptativas y el `baseline` ruidoso para $q = 5$ y $q = 4$ bajo `heavy` y 60 *shots*.

En $q = 5$, donde ϵ_{eff} es mayor, las variaciones en rendimiento relativo son más pronunciadas. El análisis emparejado por semilla indica que la estrategia basada en incremento de *shots* no produce mejoras estadísticamente significativas en ninguna profundidad considerada. En cambio, la estrategia de inyección paramétrica muestra un efecto positivo localizado en $L = 3$, con diferencia me-

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.



(a) $q = 5$, 60 shots. (b) $q = 4$, 60 shots.

Figura 3: Diferencia media de precisión entre estrategias adaptativas y `baseline` ruidoso en función de L. Barras con intervalo bootstrap al 95 %.

dia significativa sin corrección múltiple y un tamaño de efecto elevado (Cohen's $d \approx 0.79$), aunque dicho efecto no permanece significativo tras corrección por comparaciones múltiples ni se observa de forma sistemática en el resto de profundidades.

En $q = 4$ las diferencias son más moderadas y los intervalos de confianza presentan mayor solapamiento, en coherencia con la menor magnitud estimada de ϵ_{eff}. En conjunto, no se identifica una transición abrupta entre estrategias, sino variaciones dependientes del régimen señal–ruido asociado a la profundidad.

Finalmente, el análisis bajo distintos niveles de ruido confirma que el patrón exponencial emerge únicamente en el régimen `heavy`. En `none` y `light` las pendientes estimadas son pequeñas y no muestran evidencia consistente de atenuación estructural. Esto sugiere que el fenómeno no es artefacto del optimizador ni del dataset, sino consecuencia directa de la acumulación de error por puerta.

En conjunto, los resultados empíricos son coherentes con el modelo propuesto y respaldan la interpretación de la profundidad útil del ansatz como una propiedad emergente del equilibrio entre atenuación estructural del gradiente y varianza de estimación.

5. Discusión

Los resultados obtenidos proporcionan evidencia empírica consistente con la hipótesis de que la degradación del entrenamiento en régimen NISQ puede describirse mediante una tasa efectiva de atenuación del gradiente por capa, ϵ_{eff}. En el régimen de ruido `heavy`, la relación aproximadamente lineal entre $\log \bar{V}_{\nabla}^{(K)}$ y la profundidad L observada en $q = 5$ es consistente con un comportamiento exponencial efectivo de la señal coherente del gradiente. La estabilidad de la pendiente frente a variaciones en el número de *shots* indica que ϵ_{eff} refleja principalmente una propiedad estructural del circuito bajo ruido físico, asociada a la acumulación de error por puerta y a la complejidad por capa.

En configuraciones con $q = 4$ se detecta el mismo patrón, aunque con magnitud sustancialmente menor y menor capacidad explicativa del ajuste lineal en

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

Esta obra está bajo una licencia Creative Commons "Atribución-NoComercial-SinDerivadas 4.0 Internacional".

escala logarítmica. Esto sugiere que la atenuación estructural no desaparece en circuitos de menor tamaño, sino que se manifiesta con intensidad dependiente de la complejidad del circuito. Las estimaciones exploratorias para $q = 6$ bajo `heavy` resultan positivas pero con bajo poder explicativo y sin significación estadística robusta, lo que impide extraer conclusiones firmes sobre el escalado con el número de qubits, aunque no contradice el marco estructural propuesto.

El modelo efectivo establece que

$$\mathrm{SNR}(L, N) \propto \sqrt{N}\, e^{-\epsilon_{\mathrm{eff}} L},$$

lo que conduce a una profundidad útil

$$L^*(N, q) \sim \frac{1}{\epsilon_{\mathrm{eff}}(q)} \left(\frac{1}{2} \log N - C \right).$$

En este contexto, ϵ_{eff} se comporta como un parámetro estructural aproximadamente independiente del presupuesto de medición, mientras que la profundidad efectiva L^* depende tanto de dicha tasa como del número de *shots*, cuya influencia aparece únicamente de forma logarítmica. El aumento del presupuesto de medición puede retrasar la pérdida de señal útil, pero no elimina la degradación estructural inducida por la profundidad.

El comportamiento relativo de las estrategias adaptativas es coherente con esta interpretación. No se identifica una transición abrupta entre regímenes discretos, sino variaciones progresivas cuya intensidad depende del equilibrio entre atenuación estructural y varianza de estimación. En particular, la mejora puntual observada con inyección paramétrica en $q = 5$, $L = 3$ muestra que intervenciones adaptativas pueden resultar efectivas en determinados regímenes, aunque no de forma universal ni sistemática en todas las profundidades consideradas.

El estudio presenta limitaciones que deben considerarse. El modelo de ruido empleado es depolarizante e independiente por puerta y no captura todos los mecanismos físicos presentes en hardware real. La extensión del análisis a modelos de ruido calibrados a hardware real y su validación experimental en dispositivos cuánticos constituyen una línea relevante de trabajo futuro. El número de semillas por configuración es limitado, lo que restringe la potencia estadística en algunos contrastes. La métrica $\bar{V}_{\nabla}^{(K)}$ actúa como proxy operacional de la magnitud coherente del gradiente y no constituye una medición directa de la norma esperada. Asimismo, los experimentos se han realizado en simulación. El número de cúbits considerado se mantiene en un rango reducido debido al coste computacional asociado a las simulaciones ruidosas, lo que limita la evaluación del comportamiento del modelo en escalas mayores..

6. Conclusiones

En este trabajo se ha propuesto y contrastado un modelo efectivo para describir la degradación del entrenamiento en VQAs bajo ruido NISQ. En lugar de asumir la existencia de una profundidad crítica universal, se introdujo la tasa

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

efectiva de atenuación del gradiente por capa, ϵ_{eff}, como parámetro estructural que gobierna la profundidad útil del ansatz.

Los resultados en simulación bajo ruido depolarizante `heavy` muestran evidencia consistente de una atenuación aproximadamente exponencial del gradiente para $q = 5$ qubits, con estimaciones de ϵ_{eff} en el rango 0.15–0.20 y pendientes estadísticamente significativas. Estas estimaciones se mantienen estables frente a variaciones en el presupuesto de medición, lo que indica que ϵ_{eff} captura principalmente una propiedad estructural del circuito asociada a la acumulación de error por puerta y a la complejidad por capa.

En configuraciones con $q = 4$ se observa el mismo patrón, aunque con magnitud considerablemente menor y menor capacidad explicativa del modelo lineal en escala logarítmica. Las estimaciones exploratorias para $q = 6$ resultan positivas pero no concluyentes con la potencia estadística actual, lo que sugiere que la validación del escalado con el número de qubits requiere estudios adicionales.

La reinterpretación de la profundidad crítica como una cantidad emergente del equilibrio entre atenuación estructural y varianza de estimación permite formalizar una relación operativa del tipo

$$L^*(N, q) \sim \frac{1}{\epsilon_{\text{eff}}(q)} \left(\frac{1}{2} \log N - C \right),$$

en la que la tasa estructural ϵ_{eff} se comporta de manera aproximadamente independiente del presupuesto de medición, mientras que la profundidad útil depende de este de forma logarítmica. Este marco desplaza el foco desde la búsqueda de umbrales universales hacia la caracterización cuantitativa del régimen efectivo bajo ruido.

Desde una perspectiva práctica, la estimación empírica de ϵ_{eff} proporciona una herramienta para comparar configuraciones en términos de robustez estructural y para orientar decisiones de diseño tales como limitar la profundidad del circuito o ajustar el presupuesto de medición.

Como pasos futuros se identifican dos líneas complementarias. Por un lado, resulta relevante reforzar la validación estructural del modelo mediante campañas con mayor potencia experimental en configuraciones de mayor número de qubits. Por otro, la mejora puntual observada con intervenciones paramétricas adaptativas sugiere explorar de forma sistemática estrategias de control que puedan estabilizar el entrenamiento en regímenes específicos de profundidad.

En conjunto, el estudio ofrece un marco conceptual y experimental para analizar la profundidad útil de arquitecturas variacionales bajo ruido, integrando de manera explícita la degradación estructural del gradiente en el análisis del rendimiento de VQAs en régimen NISQ.

Agradecimientos Esta publicación no habría sido posible sin el apoyo de los siguientes proyectos: ATHENA-HYGIEIA (PID2024-155693NB-C42) financiado por MCIU/ AEI/10.13039/501100011033 / FEDER, UE, UEDi4SPDS (PCI2023145980-2), financiado por MCIN/AEI/ 10.13039/ 501100011033 Unión Europea (Chist-Era Program); AURORA (SBPLY/24/180225/ 000074), financiado por la Consejería de Educación,

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.

Cultura y Deportes de la Junta de Comunidades de Castilla La Mancha, y el Fondo Europeo de Desarrollo Regional FEDER (Programa Operativo de Castilla-La Mancha 2021–2027); KOSMOS-UCLM (PID2024-155363OB-C44) financiado por MCIN/AEI/ 10.13039/501100011033/FEDER, UE; RADAR (2025-GRIN-38447) financiado por FEDER; y RED2024-154240-T financiado por MICIU/AEI/ 10.13039/501100011033.

Referencias

1. Bharti, K., Cervera-Lierta, A., Kyaw, T.H., Haug, T., Alperin-Lea, S., Anand, A., Degroote, M., Heimonen, H., Kottmann, J.S., Menke, T., et al.: Noisy intermediate-scale quantum algorithms. Reviews of Modern Physics **94**(1), 015004 (2022)
2. Cerezo, M., Sone, A., Volkoff, T., Cincio, L., Coles, P.J.: Cost function dependent barren plateaus in shallow parametrized quantum circuits. Nature communications **12**(1), 1791 (2021)
3. Farhi, E., Goldstone, J., Gutmann, S.: A quantum approximate optimization algorithm. arXiv preprint arXiv:1411.4028 (2014)
4. Havlíček, V., Córcoles, A.D., Temme, K., Harrow, A.W., Kandala, A., Chow, J.M., Gambetta, J.M.: Supervised learning with quantum-enhanced feature spaces. Nature **567**(7747), 209–212 (2019)
5. Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J.M., Gambetta, J.M.: Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. nature **549**(7671), 242–246 (2017)
6. McClean, J.R., Boixo, S., Smelyanskiy, V.N., Babbush, R., Neven, H.: Barren plateaus in quantum neural network training landscapes. Nature communications **9**(1), 4812 (2018)
7. Nielsen, M.A., Chuang, I.L.: Quantum computation and quantum information. Cambridge university press (2010)
8. Ortiz Marrero, C., Kieferová, M., Wiebe, N.: Entanglement-induced barren plateaus. PRX quantum **2**(4), 040316 (2021)
9. Pesah, A., Cerezo, M., Wang, S., Volkoff, T., Sornborger, A.T., Coles, P.J.: Absence of barren plateaus in quantum convolutional neural networks. Physical Review X **11**(4), 041011 (2021)
10. Preskill, J.: Quantum computing in the nisq era and beyond. Quantum **2**, 79 (2018)
11. Proctor, T., Rudinger, K., Young, K., Nielsen, E., Blume-Kohout, R.: Measuring the capabilities of quantum computers. Nature Physics **18**(1), 75–79 (2022)
12. Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., Killoran, N.: Evaluating analytic gradients on quantum hardware. Physical Review A **99**(3), 032331 (2019)
13. Spall, J.C.: Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE transactions on automatic control **37**(3), 332–341 (2002)
14. Tilly, J., Chen, H., Cao, S., Picozzi, D., Setia, K., Li, Y., Grant, E., Wossnig, L., Rungger, I., Booth, G.H., et al.: The variational quantum eigensolver: a review of methods and best practices. Physics Reports **986**, 1–128 (2022)
15. Wang, S., Fontana, E., Cerezo, M., Sharma, K., Sone, A., Cincio, L., Coles, P.J.: Noise-induced barren plateaus in variational quantum algorithms. Nature communications **12**(1), 6961 (2021)

Braga Ortuño, C. M., Horneros, A. J., Serrano, M. A., Fernandez-Medina, E.: Profundidad efectiva en algoritmos de Quantum Machine Learning: atenuación del gradiente en VQAs bajo ruido NISQ. In: Cetina, C. (ed.) Actas de las XXX Jornadas de Ingeniería del Software y Bases de Datos (JISBD 2026). Sistedes (2026). https://hdl.handle.net/11705/JISBD/2026/140

Esta obra está bajo una licencia Creative Commons “Atribución-NoComercial-SinDerivadas 4.0 Internacional”.