مجلة جامعة حمص

سلسلة العلوم الهندسية الميكانيكية والكهربائية والمعلوماتية

مجلة علمية محكمة دورية

المجلد 48 . العدد 5

1448 هـ - 2026 م

الأستاذ الدكتور طارق حسام الدين رئيس جامعة حمص **المدير المسؤول عن المجلة**

أ. د. وليد حمادة	**رئيس تحرير مجلة جامعة حمص للعلوم الإنسانية**
د.نعيمة عجيب	**رئيس تحرير مجلة جامعة حمص للعلوم الطبية والهندسية والأساسية والتطبيقية**

د.محمد فراس رمضان	عضو هيئة التحرير
د. مضر سعود	عضو هيئة التحرير
د. ممدوح عبارة	عضو هيئة التحرير
د. موفق تلاوي	عضو هيئة التحرير
د. طلال رزوق	عضو هيئة التحرير
د. أحمد الجاعور	عضو هيئة التحرير
د. الياس خلف	عضو هيئة التحرير
د. روعة الفقس	عضو هيئة التحرير
د. محمد الجاسم	عضو هيئة التحرير
د. خليل الحسن	عضو هيئة التحرير
د. هيثم حسن	عضو هيئة التحرير
د. أحمد حاج موسى	عضو هيئة التحرير

تهدف المجلة إلى نشر البحوث العلمية الأصيلة، ويمكن للراغبين في طلبها الاتصال بالعنوان التالي:

رئيس تحرير مجلة جامعة حمص

سورية . حمص . جامعة حمص . الإدارة المركزية . ص . ب (77)

. هاتف / فاكس : 2138071 31 963 ++

. موقع الإنترنت : www.homs-univ.edu.sy

. البريد الالكتروني : journal.homs-univ.edu.sy

ISSN: 1022-467X

شروط النشر في مجلة جامعة حمص

الأوراق المطلوبة:

- 2 نسخة ورقية من البحث بدون اسم الباحث / الكلية / الجامعة) + CD / word من البحث منسق حسب شروط المجلة.
- طابع بحث علمي + طابع نقابة معلمين.
- **اذا كان الباحث طالب دراسات عليا:**

يجب إرفاق قرار تسجيل الدكتوراه / ماجستير + كتاب من الدكتور المشرف بموافقته على النشر في المجلة.

- **اذا كان الباحث عضو هيئة تدريسية:**

يجب إرفاق قرار المجلس المختص بإنجاز البحث أو قرار قسم بالموافقة على اعتماده حسب الحال.

- **اذا كان الباحث عضو هيئة تدريسية من خارج جامعة البعث :**

يجب إحضار كتاب من عمادة كليته تثبت أنه عضو بالهيئة التدريسية و على رأس عمله حتى تاريخه.

- **اذا كان الباحث عضواً في الهيئة الفنية :**

يجب إرفاق كتاب يحدد فيه مكان و زمان إجراء البحث ، وما يثبت صفته وأنه على رأس عمله.

- **يتم ترتيب البحث على النحو الآتي بالنسبة لكليات (العلوم الطبية والهندسية والأساسية والتطبيقية):**

عنوان البحث . . ملخص عربي و إنكليزي (كلمات مفتاحية في نهاية الملخصين).

1- مقدمة

2- هدف البحث

3- مواد وطرق البحث

4- النتائج ومناقشتها .

5- الاستنتاجات والتوصيات .

6- المراجع.

- **يتم ترتيب البحث على النحو الآتي بالنسبة لكليات (الآداب – الاقتصاد– التربية – الحقوق – السياحة – التربية الموسيقية وجميع العلوم الإنسانية):**
- عنوان البحث . . ملخص عربي و إنكليزي (كلمات مفتاحية في نهاية الملخصين).

1. مقدمة.
2. مشكلة البحث وأهميته والجديد فيه.
3. أهداف البحث و أسئلته.
4. فرضيات البحث و حدوده.
5. مصطلحات البحث و تعريفاته الإجرائية.
6. الإطار النظري و الدراسات السابقة.
7. منهج البحث و إجراءاته.
8. عرض البحث و المناقشة والتحليل
9. نتائج البحث.
10. مقترحات البحث إن وجدت.
11. قائمة المصادر والمراجع.

7- يجب اعتماد الإعدادات الآتية أثناء طباعة البحث على الكمبيوتر:

أ- قياس الورق 17.5×25 B5.

ب- هوامش الصفحة: أعلى 2.54- أسفل 2.54 – يمين 2.5- يسار 2.5 سم

ت- رأس الصفحة 1.6 / تذييل الصفحة 1.8

ث- نوع الخط وقياسه: العنوان . Monotype Koufi قياس 20

ـ كتابة النص Simplified Arabic قياس 13 عادي ـ العناوين الفرعية Simplified Arabic قياس 13 عريض.

ج . يجب مراعاة أن يكون قياس الصور والجداول المدرجة في البحث لا يتعدى 12سم.

8- في حال عدم إجراء البحث وفقاً لما ورد أعلاه من إشارات فإن البحث سيهمل ولا يرد البحث إلى صاحبه.

9- تقديم أي بحث للنشر في المجلة يدل ضمناً على عدم نشره في أي مكان آخر، وفي حال قبول البحث للنشر في مجلة جامعة البعث يجب عدم نشره في أي مجلة أخرى.

10- الناشر غير مسؤول عن محتوى ما ينشر من مادة الموضوعات التي تنشر في المجلة

11- تكتب المراجع ضمن النص على الشكل التالي: [1] ثم رقم الصفحة ويفضل استخدام التهميش الإلكتروني المعمول به في نظام وورد WORD حيث يشير الرقم إلى رقم المرجع الوارد في قائمة المراجع.

تكتب جميع المراجع باللغة الانكليزية (الأحرف الرومانية) وفق التالي:

آ . إذا كان المرجع أجنبياً:

الكنية بالأحرف الكبيرة ـ الحرف الأول من الاسم تتبعه فاصلة ـ سنة النشر ـ وتتبعها معترضة (–) عنوان الكتاب ويوضع تحته خط وتتبعه نقطة ـ دار النشر وتتبعها فاصلة ـ الطبعة (ثانية . ثالثة) . بلد النشر وتتبعها فاصلة . عدد صفحات الكتاب وتتبعها نقطة.

وفيما يلي مثال على ذلك:

–MAVRODEANUS, R1986– **Flame Spectroscopy**. Willy, New York, 373p.

ب . إذا كان المرجع بحثاً منشوراً في مجلة باللغة الأجنبية:

ـــ بعد الكنية والاسم وسنة النشر يضاف عنوان البحث وتتبعه فاصلة، اسم المجلد ويوضع تحته خط وتتبعه فاصلة ـــ المجلد والعدد (كتابة مختزلة) وبعدها فاصلة ـــ أرقام الصفحات الخاصة بالبحث ضمن المجلة.

مثال على ذلك:

BUSSE,E 1980 Organic Brain Diseases **Clinical Psychiatry News** , Vol. 4. 20 – 60

ج. إذا كان المرجع أو البحث منشوراً باللغة العربية فيجب تحويله إلى اللغة الإنكليزية و التقيد بالبنود (أ و ب) ويكتب في نهاية المراجع العربية: (المراجع In Arabic)

رسوم النشر في مجلة جامعة حمص

1. **دفع رسم نشر (50000) ل.س أربعون ألف ليرة سورية عن كل بحث لكل باحث يريد نشره في مجلة جامعة البعث.**
2. **دفع رسم نشر (200000) ل.س مئة الف ليرة سورية عن كل بحث للباحثين من الجامعة الخاصة والافتراضية .**
3. **دفع رسم نشر (200) مئتا دولار أمريكي فقط للباحثين من خارج القطر العربي السوري .**
4. **دفع مبلغ (15000) ل.س ستة آلاف ليرة سورية رسم موافقة على النشر من كافة الباحثين.**

المحتوى

اسم البحث	اسم الباحث	الصفحة
هندسة الحصر الكمي لتحسين كفاءة الطاقة في ترانزستورات أنابيب الكربون النانوية المتعرجة فائقة الانخفاض في استهلاك الطاقة	د.خضر بشور	58-11
تقييم أداء شيفرة HAPPY/Steane الكمومية الهولوغرافية تحت أخطاء باولي النقية (X,Z,Y) وخطأ إزالة الاستقطاب (Depolarizing Error)	مجد سالم د. اسبر نسيم إبراهيم	92-59
تحسين ترتيب الربط باستخدام خوارزمية MCTS الموجَهة بمعلومات الاستعلام في قواعد البيانات	محمد حاج حسن أ.د. يُسر السيد سليمان الأتاسي	136-93
تحسين الصيانة التنبؤية لآلات التفريز باستخدام هندسة الميزات ونماذج تعلّم الآلة الهجينة	صفاء مصطفى محسن د. حسام الوفائي	164-137

هندسة الحصر الكمي لتحسين كفاءة الطاقة في ترانزستورات أنابيب الكربون النانوية المتعرجة فائقة الانخفاض في استهلاك الطاقة

اعداد : د.خضر بشور

الملخص:

تقدم هذه الدراسة منهجاً في هندسة الحصر الكمي لتحسين كفاءة الطاقة في ترانزستورات تأثير المجال المصنعة من أنابيب الكربون النانوية(CNTFETs)، من خلال الاستفادة من الحصر الكمي المعتمد على القطر في أنابيب الكربون النانوية المتعرجة. وقد تم استخدام نموذج تحليلي يعتمد على صيغة لانداور لمحاكاة خصائص التيار–الجهد، وعلى العلاقات النظرية لفجوة النطاق المستمدة من نموذج الرباط المحكم.

تم توسيع النموذج ليشمل مقاومة التلامس(100 kΩ)، والسعات الطفيلية، وحساب حاصل (نتاج) القدرة–التأخير(PDP) ، ودراسة سرعة التبديل من خلال زمن التأخير (τ) وتردد القطع(f_T) ، ونمذجة التسخين الذاتي(Self-heating) ، وتحليل حساسية الأداء لانحرافات التصنيع. كما تمت مقارنة عازلين □SiO و □HfO ودراسة تأثير ذلك على الأداء.

تظهر النتائج أن تقليل قطر الأنبوب النانوي من 1.588 نانومتر إلى 0.556 نانومتر يزيد فجوة النطاق من 0.537eV إلى 1.533eV، مما يكبح تيار التسريب (I_{OFF}) بأكثر من ثمانية مراتب قياسية، محققاً نسبة I_{ON}/I_{OFF} تتجاوز 10^{12}. كما يقلل عازل □HfO من I_{OFF} بمقدار 4.4 مرات ولكنه يزيد PDP بمقدار 6.4 مرات. تتراوح سرعة التبديل بين 0.27-2.40ps وتردد القطع بين 66-595GHz، ويرتفع التسخين الذاتي بمقدار 19-30K . كما تم تحديد النطاق القطري 1.0-1.2nm كمنطقة تصميم مثلى، مع توصيات بالتكامل الهجين مع تقنيات CMOS الحالية في تطبيقات SRAM وأجهزة IoT . تؤكد هذه النتائج أن هندسة الحصر الكمي تتيح مفاضلة تصميمية مرنة بين التسريب، السرعة، واستهلاك الطاقة الديناميكية، مما يسمح بتخصيص أداء الترانزستور وفقاً لمتطلبات التطبيق المحددة، بدءاً من أجهزة الاستشعار منخفضة الطاقة وصولاً إلى المعالجات عالية التردد.

الكلمات المفتاحية:

الحصر الكمي، ترانزستورات أنابيب الكربون النانوية المتعرجة، فجوة النطاق وكبح التسريب، حاصل القدرة–التأخير، سرعة التبديل والتسخين الذاتي.

Quantum Confinement Engineering for Improving Energy Efficiency in Ultra-Low-Power Zigzag Carbon Nanotube Transistors

Abstract

This study introduces a quantum confinement engineering approach for enhancing the energy efficiency of Carbon Nanotube Field Effect Transistors (CNTFETs), through the utilization of diameter-dependent quantum confinement in zigzag carbon nanotubes. An analytical model based on the Landauer formula was employed to simulate current-voltage characteristics, relying on theoretical bandgap relations derived from the tight-binding model.

The model was extended to include contact resistance (100 kΩ), parasitic capacitances, power-delay product (PDP) calculation, switching speed analysis through propagation delay (τ) and cut-off frequency (f_T), self-heating modeling, and sensitivity analysis for fabrication tolerances. Two gate dielectrics (SiO_2 and HfO_2) were compared to investigate their impact on device performance.

Our results show that reducing the nanotube diameter from 1.588 nm to 0.556 nm increases the bandgap from 0.537 eV to 1.533 eV, leading to exponential suppression of leakage current (I_{OFF}) by over eight orders of magnitude, achieving an I_{ON}/I_{OFF} ratio exceeding 10^{12}. HfO_2 reduces I_{OFF} by 4.4× but increases PDP by 6.4× compared to SiO_2. The switching speed ranges from 0.27-2.40 ps with cut-off frequencies between 66-595

GHz, while self-heating raises the operating temperature by 19-30 K. The optimal diameter range for digital applications is identified as 1.0-1.2 nm, with recommendations for hybrid integration with existing CMOS technologies in SRAM and IoT applications. These findings confirm that quantum confinement engineering enables a flexible design trade-off between leakage current, switching speed, and dynamic power consumption, allowing transistor performance to be tailored according to specific application requirements, from low-power sensors to high-frequency processors.

Keywords:
Quantum confinement, zigzag carbon nanotube transistors, bandgap and leakage suppression, power-delay product, switching speed and self-heating..

1- المقدمة:

ما زالت عملية التصغير المستمرة للإلكترونيات القائمة على السيليكون تواجه تحديات أساسية في إدارة الطاقة، خاصة مع اقتراب أبعاد الأجهزة من المقاييس الذرية. عند هذه الحدود، تصبح الظواهر الميكانيكا الكمية، ولا سيما تيارات التسريب النفقية، أكثر هيمنة، مما يؤدي إلى تدهور كبير في الأداء وكفاءة الطاقة على حد سواء. وقد دفعت هذه التحديات إلى استكشاف مواد بقنوات بديلة، حيث برزت أنابيب الكربون النانوية (CNTs) كمرشح واعد بشكل خاص [1]. فبنيتها شبه أحادية البعد الفريدة، المقترنة بالنقل شبه الباليستي لحاملات الشحنة، تقدم إمكانات كبيرة للتطبيقات الإلكترونية عالية السرعة ومنخفضة الطاقة [2].

تستفيد ترانزستورات تأثير المجال المصنوعة من أنابيب الكربون النانوية (CNTFETs) من هذه الخصائص المميزة لتحقيق تحكم إلكتروستاتيكي محسن. يتيح هذا التحكم المُحسّن تنظيم الشحنة

بكفاءة مع التخفيف من التيارات النفقية المباشرة التي تؤثر بشكل متزايد على الترانزستورات السيليكونية التقليدية مع استمرار تصغيرها [3].

السمة الرئيسية التي تميز أنابيب الكربون النانوية هي التأثير البارز للحصر الكمي في خواصها الإلكترونية. على عكس أشباه الموصلات الحجمية التي تتميز بفجوة نطاق ثابتة، فإن فجوة النطاق لأنابيب الكربون النانوية شبه الموصِلة ترتبط عكسياً بقطره [4]. بالنسبة للأنابيب النانوية متعرجة النمط (موضوع هذا البحث)، تأخذ هذه العلاقة شكلاً رياضياً واضحاً يربط مباشرة بين البنية الهندسية والسلوك الإلكتروني. ينشأ هذا السلوك من حصر الدوال الموجية الإلكترونية داخل الهيكل الأسطواني، مما يُكمم الحالات الإلكترونية المتاحة. تُحول العلاقة العكسية بين فجوة النطاق والقطر الأبعاد الهيكلية للأنابيب النانوية المتعرجة إلى معامل تصميم كمي دقيق يمكن هندسته بشكل منهجي.

تشكل أنابيب الكربون النانوية المتعرجة، التي يتراوح قطرها النموذجي من 0.5 إلى 2 نانومتر، فئة مهمة من الأنظمة الميزوسكوبية حيث تكون ظواهر الحصر الكمي واضحة بشكل خاص. تجعل العلاقة الواضحة بين قطرها وخصائصها الإلكترونية منها منصات مثالية لاستكشاف كيفية توظيف التأثيرات الكمية لتحسين الأجهزة عملياً [5].

على الرغم من أن أداء ترانزستورات تأثير المجال لأنابيب الكربون النانوية مُوثق جيداً في الأدبيات العلمية، فإن التوظيف المنهجي للحصر الكمي في الأنابيب النانوية المتعرجة لتحسين كفاءة الطاقة لا يزال غير مُستكشف نسبياً. تعامل العديد من الدراسات مع فجوة النطاق على أنها خاصية مادية أصيلة متجاهلة إمكانية ضبطها عبر اختيار القطر المناسب.

بالإضافة إلى ذلك، ركزت معظم الدراسات على الخصائص الساكنة، متجاهلة معايير الأداء الديناميكي مثل سرعة التبديل، حاصل القدرة–التأخير(PDP) ، والتسخين الذاتي، والتي تعتبر حاسمة للتطبيقات الرقمية الحديثة. كما أن تأثير اختيار العازل مثل HfO□ (عالي السماحية) وتحسين التلامس المعدني–النانوي على الأداء لم يحظَ بالاهتمام الكافي.

تركز هذه الدراسة على أنابيب الكربون النانوية المتعرجة لإظهار كيف تتيح هندسة الحصر الكمي تحقيق تحسينات ملحوظة في كفاءة الطاقة لترانزستورات .CNTFET من خلال التحقيق المنهجي في الأنماط المتعرجة عبر نطاق قطر يتراوح بين 0.476-1.588nm، تم تحقيق تعزيز كبير في

فجوة النطاق وكبح تيارات التسريب بأكثر من ثمانية مراتب قياسية. كما تم إدراج مقاومة التلامس (100 kΩ)والسعات الطفيلية لحسابPDP ، ودراسة سرعة التبديل من خلال زمن التأخير (τ) وتردد القطع(f_T) ، ونمذجة التسخين الذاتي(Self-heating) ، وتحليل حساسية الأداء لانحرافات التصنيع، ومقارنة عازلي SiO و HfO كما تم دراسة تأثير دالة الشغل للأقطاب وكثافة الحالات (DOS) وتوزيع الشحنات. أخيراً، تم اقتراح منهجية للتكامل الهجين مع تقنيات CMOS الحالية في تطبيقات SRAM وأجهزة IoT . تكشف النتائج أن الحصر الكمي في الأنابيب النانوية المتعرجة ليس مجرد ظاهرة فيزيائية مثيرة للاهتمام، بل هو مبدأ تصميم دقيق للإلكترونيات النانوية المستقبلية منخفضة الطاقة.

2- هدف الدراسة:

هدف هذه الدراسة إلى استكشاف إمكانية توظيف هندسة الحصر الكمي في ترانزستورات أنابيب الكربون النانوية المتعرجة كأداة تصميم فعالة لتحسين كفاءة الطاقة، وذلك من خلال:

- دراسة تأثير تغيير قطر الأنبوب النانوي على فجوة النطاق، وتحليل كيفية انعكاس ذلك على تيارات التسريب (I_{OFF}) ونسب تيار التشغيل/الإيقاف (I_{ON}/I_{OFF}).
- إدراج مقاومة التلامس ($R_{contact} = 100k\Omega$) والسعات الطفيلية لحساب حاصل القدرة–التأخير (Power-Delay Product - PDP) ، وهو المعيار الحقيقي لتقييم كفاءة الطاقة في الدوائر الرقمية.
- دراسة سرعة التبديل والاستجابة الترددية للجهاز من خلال حساب زمن التأخير (τ) وتردد القطع(f_T) ، لضمان أن الترانزستور سريع بما يكفي للتطبيقات الرقمية الحديثة.
- نمذجة التسخين الذاتي (Self-heating) لدراسة كيفية تصريف الحرارة الناجمة عن كثافة التيار العالية في الأقطار الصغيرة جداً، وضمان استقرار الأداء تحت ضغط العمل المستمر.
- مقارنة أداء عازلين SiO التقليدي و HfO عالي السماحية(High-k) ، لمواجهة الزيادة الكبيرة في تيار التسريب عند تقليل طول القناة.

- تحليل حساسية الأداء لانحرافات التصنيع (Fabrication Tolerances) باستخدام طريقة الاستيفاء الخطي، لتحديد هامش الخطأ المسموح به في قطر الأنبوب النانوي.
- دراسة تأثير دالة الشغل (Work Function) للأقطاب على أداء الترانزستور، وتحديد المعدن الأمثل للتلامس.
- اقتراح منهجية للتكامل الهجين (Hybrid Integration) مع تقنيات CMOS الحالية مثلاً في وحدات الذاكرة SRAM وأجهزة IoT بدلاً من طرحها كبديل كامل، مما يعزز واقعية البحث وقابليته للتطبيق التجاري.
- تحديد النطاق القطري الأمثل الذي يحقق توازناً بين الأداء العالي، سرعة التبديل، واستهلاك الطاقة المنخفض، مع تقييم استقرار الأداء عبر درجات حرارة تشغيل مختلفة (200–400 كلفن).

3- مواد البحث وطرائقه:

تم في هذا البحث حساب فجوات النطاق باستخدام العلاقة النظرية ($E_g = 0.852/d$)، وهي مستمدة من نموذج الترابط القوي (Tight-Binding Model) للأنابيب النانوية المتعرجة، في حين تمت محاكاة خصائص التيار–الجهد باستخدام محاكي FETToy القائم على صيغة لانداور. طُبقت المعاملات الموحدة الآتية: طول قناة 20nm، سمك عازل (SiO_2) 1.5nm، وجهد V_{DS} = 1V، مع دراسة تأثير تغير طول القناة ودرجة الحرارة (200–400 كلفن) على الأداء.

كما تم أيضاً دراسة مقاومة التلامس والسعات الطفيلية وتم حساب سعة البوابة وحساب حاصل القدرة–التأخير (Power-Delay Product (PDP. و تمت المقارنة بين الأداء عند استخدام عازلين SiO_2 و HfO_2 بنفس السماكة t_{ox} = 1.5nm .كما تمت دراسة سرعة التبديل من خلال حساب زمن التأخير وتردد القطع. ونمذجة التسخين الذاتي (Self-heating) عبر حساب القدرة الحرارية والمقاومة الحرارية. بالإضافة إلى ذلك تمت دراسة كثافة الحالات (DOS) للأنبوب النانوي وحساب توزيع الشحنات. وكذلك تأثير الأقطاب (Source/Drain Contacts) على أداء الترانزستور.

4- نمذجة ترانزستورات CNTFET :

يحدد هذا القسم المبادئ الفيزيائية الأساسية التي تحكم الحصر الكمي في أنابيب الكربون النانوية المتعرجة، ويقدم النماذج الكمية المستخدمة لتقييم تأثيرها على أداء الترانزستور.

4-1 الحصر الكمي وتشكيل فجوة النطاق في الأنابيب النانوية متعرجة النمط:

يؤدي تقليل الأبعاد في المواد النانوية – من الطبقات ثنائية الأبعاد إلى الأسلاك النانوية أحادية البعد، وصولاً إلى النقاط الكمية صفرية الأبعاد – إلى تعديلات عميقة في خصائصها الإلكترونية بسبب الحصر الكمي [6].

على عكس أشباه الموصلات التقليدية ذات فجوة النطاق الجوهرية الثابتة، فإن فجوة النطاق لأنابيب الكربون النانوية هي خاصية قابلة للتصميم تحددها هندستها. تنشأ هذه القابلية للضبط من الحصر الكمي، حيث يؤدي تقييد حركة حاملات الشحنة إلى أبعاد مماثلة لطولها الموجي الكمي إلى ظهور حالات طاقة منفصلة [7]. في أنابيب الكربون النانوية المتعرجة، يكون هذا الحصر واضحاً بسبب بنيتها الأسطوانية. يجب أن تفي الدوال الموجية للإلكترونات بشروط دورية حول محيط الأنبوب، مما يُكمم الزخم المحيطي، ويربط بشكل مباشر بين البنية الفيزيائية والخصائص الإلكترونية4] ،[5. بالنسبة للأنابيب النانوية المتعرجة شبه الموصلة، والمعرفة بمتجه النمط الحلزوني(n,0) ، هناك علاقة عكسية أساسية بين فجوة النطاق (E_g) والقطر (d) [8]:

$$E_g = \frac{2\gamma_0 a_{c-c}}{d} = \frac{0.852}{d} \quad (1)$$

حيث أن γ هي طاقة الترابط ($\gamma \approx 3.0$ eV)، و a_{c-c} هي طول الرابطة الكربون–كربون(a_{c-c} = 1.42 Å)، و d هو قطر الأنبوب النانوي.

تتفق هذه العلاقة العكسية مع ما ورد في دراسات سابقة[9] ، حيث تم التأكيد على أن فجوة النطاق في أنابيب الكربون النانوية ترتبط عكسياً مع القطر، وأن زيادة القطر تؤدي إلى تقليل فجوة النطاق. هذه التناسبية تعني أن فجوة النطاق قابلة للضبط، فاختيار قطر أصغر يؤدي مباشرة إلى فجوة نطاق أكبر. بالنسبة للأنابيب النانوية متعرجة النمط، فإن شرط السلوك شبه الموصِل هو (n $\bmod 3) \neq 0$ [1].

4–2 إطار عمل النقل الكمي:

لحساب خصائص التيار–الجهد في ترانزستورات تأثير المجال لأنابيب الكربون النانوية (CNTFETs)، تم استخدام محاكيFETToy ، وهو نموذج تحليلي يعتمد على صيغة لانداور في إطار النقل الباليستي، يقوم محاكي FETToy بحساب التيار الكهربائي باستخدام صيغة لانداور (Landauer formula) والتي تعطى بالعلاقة التالية[10]:

$$I = \frac{4e}{h}\int_{-\infty}^{+\infty} T\left(E, V_g\right)[f_s(E) - f_D(E)]dE \qquad (2)$$

حيث أن $T(E,V_g)$ هي احتمالية النفاذ عبر الجهاز تحت جهد البوابةVg ، و $f_S(E), f_D(E)$ هي توزيعات فيرمي–ديراك لوصلة المصرف – المنبع.

3–4 نمذجة تيار التسريب:

المصدر المهيمن للطاقة الساكنة في الترانزستورات النانوية هو تيار التسريب في حالة الإيقاف (I_{OFF})، والذي يحكمه النفق الكمي من نطاق إلى نطاق (BTBT) (Band-to-Band Tunneling)، تعتمد احتمالية النفق أُسياً على فجوة النطاق. يُعطي النموذج القياسي المُشتق من تقريب WKB (Wentzel–Kramers–Brillouin approximation) لترانزستورات تأثير المجال لأنابيب الكربون النانوية (CNTFETs) بالعلاقة [11] :

$$I_{off} = I_o.\exp(-\frac{\pi\sqrt{m^*}.{E_g}^{3/2}}{2\sqrt{2}.e.\hbar.F}) \qquad (3)$$

حيث أن: I_{OFF} هو تيار التسريب في حالة الإيقاف (V_{GS}=0V)، و I_O هو معامل التيار الأساسي (يتناسب مع مساحة المقطع العرضي وكثافة الحالات)، و*m هي الكتلة الفعالة (كتلة الإلكترون في الوسط البلوري لأنابيب الكربون النانوية)، وF هو شدة المجال الكهربائي، و e هي شحنة الإلكترون وتساوي C □10□¹□ × 1.6، و □ هو ثابت بلانك المخفض (× h/2π = 1.055 10□³□ J·s).

يكشف الجمع بين المعادلتين (2.1) و (2.3) عن علاقة متتالية قوية:

$$d \downarrow \Rightarrow E_g \uparrow \Rightarrow I_{off} \downarrow \qquad (4)$$

وبالتالي، فإن تعزيز الحصر الكمي عن طريق تقليل القطر يزيد من فجوة النطاق، مما ينعكس بدوره على كبح تيار التسريب بشكل أسي. شكل هذا التأثير المضاعف القوي محور استراتيجية تحسين الطاقة في هذه الدراسة.

4-4 تأثير درجة الحرارة على الحصر الكمي:

يجب أن يظل أداء ترانزستورات CNTFET مستقرًا عبر درجات حرارة التشغيل، ويعتبر تضييق فجوة النطاق مع زيادة درجة الحرارة، بسبب تفاعلات الإلكترون–فونون، هو أمر راسخ في فيزياء أشباه الموصلات. استنادًا إلى بيانات فجوة النطاق المعتمدة على درجة الحرارة المبلغ عنها في الأدبيات [12]، تمت معايرة هذا الاعتماد باستخدام تقريب خطي:

$$E_g = E_g(300) + \alpha \times (\mathrm{T} - 300) \qquad (5)$$

حيث إن $\alpha = -4.7 \times 10^{-4}\ eV/K$ وهو معامل درجة الحرارة.

يقدم هذا النموذج الخطي تقريبًا معقولاً على نطاق درجة الحرارة (200-400K) المدروس، حيث يؤدي هذا التضييق الحراري إلى زيادة التسريب عند درجات الحرارة الأعلى.

4-5: نمذجة المقاومة الطفيلية والسعات والتأثيرات العابرة:

أولاً: المقاومة الكلية ومقاومة التلامس:

وفقاً للنموذج المرجعي لترانزستورات CNTFET الصادر عن جامعة ستانفورد [13]، تتكون المقاومة التسلسلية للترانزستور من مقاومة التلامس المعدني–النانوي (R_C) ومقاومة امتدادات المصرف والمنبع (R_{ext}).

$$R_S = R_C + R_{ext} \qquad (6)$$

في هذا البحث، ولأن طول القناة قصير 10nm أو 20nm، تم دمج مقاومة الامتدادات ضمن مقاومة القناة، لتصبح المقاومة الكلية للترانزستور في حالة التشغيل:

$$R_{total} = R_{contact} + R_{channel}$$

تم اعتماد قيمة $R_{contact} = 100K\Omega$ بناءً على نموذج جامعة ستانفورد VS-CNFET الذي يعاير هذه القيمة تجريبياً [13]، حيث يتراوح النطاق النموذجي لمقاومة التلامس في أنابيب الكربون النانوية بين 10KΩ و 1MΩ.

وبالنسبة لـ $R_{channel}$ فهي تساوي

$$R_{channel} = \frac{V_{DS}}{I_{ON}}$$

ثانياً: سعة البوابة الأسطوانية:

في حالة الهيكل الأسطواني المحيطي(Gate-All-Around) ، تحسب سعة البوابة (لكل وحدة طول) بالعلاقة [14]:

$$C_{gs} = \frac{2\pi\varepsilon_0\varepsilon_{ox}}{\ln[\frac{2h+d}{d}]} \quad (7)$$

حيث: h: سماكة العازل، d قطر الأنبوب، ε_{ox} هو ثابت العازلية له، ε_0 هو ثابت العازلية في الفراغ.

ثالثاً: حاصل (نتاج) القدرة–التأخير (PDP):

يعد حاصل القدرة–التأخير (PDP) مقياساً أساسياً لكفاءة الطاقة في الدوائر الرقمية، حيث يعبر عن الطاقة اللازمة لتبديل الترانزستور من حالة إلى أخرى، يعطى PDP في الدوائر الرقمية بالعلاقة [15]:

$$PDP = C_{Load} \times {V^2}_{DD}$$

حيث C_{load} هي سعة الحمل الكلية، و V_{DD} هو جهد التغذية (1V في هذه الدراسة) في النماذج المبسطة ولأن السعة الطفيلية صغيرة نسبياً مقارنة بسعة البوابة في الترانزستورات فائقة الصغر، تم استخدام سعة البوابة C_{gate} كتقريب لـ C_{load} ، أي $C_{Load} \approx C_{gate}$ لتصبح المعادلة المستخدمة في هذا العمل على الشكل التالي:

$$PDP = C_{gate} \times {V^2}_{DD} \quad (8)$$

رابعاً: التأثيرات العابرة (سرعة التبديل):

لتقييم الأداء الديناميكي للترانزستور وقدرته على العمل في التطبيقات عالية التردد، تم حساب معلمتين رئيسيتين: زمن التأخير (τ) وتردد القطع (f_T): باستخدام العلاقتين التاليتين [16]:

$$\tau = \frac{C_{gate} \times V_{DD}}{I_{ON}} \quad (9)$$

$$f_T = \frac{I_{ON}}{2 \times \pi \times C_{total} \times V_{DD}} \quad (10)$$

خامساً: التسخين الذاتي (Self-heating) :

تنتج الحرارة في الترانزستور أثناء التشغيل بسبب مرور التيار عبر مقاومة القناة، مما يؤدي إلى ارتفاع درجة حرارة الجهاز. يمكن حساب الطاقة الحرارية المتولدة في الترانزستور أثناء التشغيل بالعلاقة:

$$P_{heat} = I_{ON} \times V_{DS} \quad (11)$$

حيث P_{heat} هي القدرة الحرارية المتولدة بوحدة الواتW ، I_{ON} هو تيار التشغيل بوحدة الأمبير A، و V_{DS} هو جهد المصرف –المنبع (1.0 V).

ويعطى ارتفاع درجة حرارة الأنبوب النانوي عن درجة حرارة المحيط يعطى بالعلاقة [17]:

$$\Delta T = P_{heat} \times R_{th} \quad (12)$$

حيث ΔT هو ارتفاع درجة الحرارة بوحدة الكلفنK ، و R_{th} هي المقاومة الحرارية الكلية للجهاز (Thermal Resistance)والتي تعتمد على قطر الأنبوب النانوي وخصائص المواد المحيطة، وهي مدى صعوبة خروج الحرارة من الأنبوب النانوي إلى المحيط. تحسب بالعلاقة [17]:

$$R_{th} = R_{th}(CNT) + R_{th}(Oxide) + R_{th}(Substrate) \quad (13)$$

حيث إن:

$R_{th}(CNT)$: هي المقاومة الحرارية للأنبوب النانوي نفسه.

$R_{th}(Oxide)$ هي المقاومة الحرارية لطبقة العازل (SiO_2) أو (HfO_2).

$R_{th}(Substrate)$ هي المقاومة الحرارية للركيزة (القاعدة السفلية).

للتبسيط تم اعتماد قيمة نموذجية لـ R_{th} كما يلي: للأنبوب (7,0): 10^7 K/W و للأنبوب (13,0): 5×10^6 K/W وللأنبوب (20,0): 3×10^6 K/W.

وتكون درجة حرارة التشغيل الفعلية للترانزستور هي [17]:

$$T_{\text{operation}} = T_{\text{ambient}} + \Delta T \quad (14)$$

حيث $T_{ambient}$ هي درجة حرارة المحيط 300 K (في الظروف العادية).

4-6 كثافة الحالات (DOS) وتوزيع الشحنات:

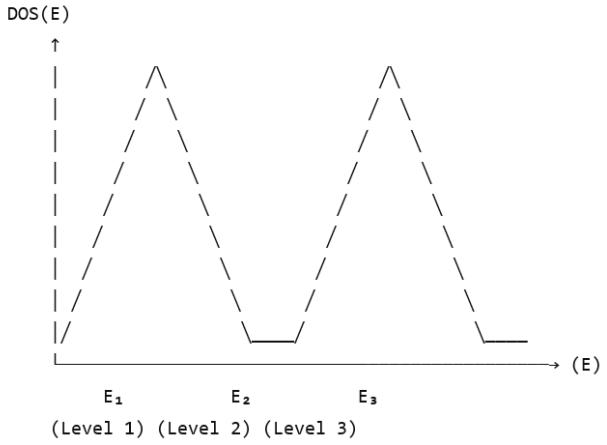
تصف كثافة الحالة (DOS) لنظام ما عدد الحالات الممكن شغلها لكل فترة من الطاقة في كل مستوى طاقة متاح، حيث يعني ارتفاع مستوىDOS عند مستوى طاقة معين أن هناك العديد من الحالات المتاحة ليتم شغلها، فعندما DOS يساوي الصفر ذلك يعني أنه لا يمكن شغل أي حالة على مستوى الطاقة هذا.

تعمل الأنابيب النانوية كمواد نصف ناقلة إذا كانت تحتوي فجوة طاقة (Energy Band) بين حزمة الناقلية وحزمة التكافؤ. وتتصرف كمواد ناقلة إذا تداخلت حزمة الناقلية مع حزمة التكافؤ وعندها يمكن أن يثار الإلكترون بسهولة أكبر من مستوى حزمة التكافؤ إلى حزمة الناقلية. وبالتالي فإن DOS وأنواع CNT هي التي تحدد الناقلية للمادة. بالنسبة للإلكترون في حزمة الناقلية، فإن زيادة طاقة الإلكترون تتسبب في توفير المزيد من الحالات ليتم شغلها، عندما تكون كثافة الحالة متقطعة لفاصل زمني معين للطاقة فهذا يعني أنه لا توجد حالات متاحة للإلكترونات لكي تشغلها في فجوة الطاقة. وهذا يعني أن الإلكترون الذي يكون على حافة حزمة الناقلية يجب أن يفقد طاقة تساوي فجوة الطاقة للانتقال إلى حزمة التكافؤ. إن كثافة الحالات موجودة فقط للإلكترونات الحرة، وتكون كثافة الحالات هي صفر داخل فجوة الطاقة [12].

عندما نطبق جهداً على البوابة(V_{GS}) ، نرفع مستوى طاقة فيرمي (E_F) داخل الأنبوب. وكلما ارتفعتE_F ، زاد عدد الإلكترونات التي تدخل القناة. وتؤثر كثافة الحالات على تيار التشغيل (I_{ON}) حيث أن: التيار = (شحنة الإلكترون) × (السرعة) × (عدد الإلكترونات) وباعتبار أن السرعة ثابتة تقريباً في الأنابيب النانوية (سرعة فيرمي $8 \times 10^5 m/s \approx$)، فإنه عندما تكون كثافة الحالات أكبر يكون عدد الالكترونات أكبر وبالتالي نحصل على تيار أكبر.

في الأنابيب النانوية أحادية البعد، تظهر ما يعرف بتفردات فان هوف (Van Hove singularities)في كثافة الحالات، حيث تبلغ كثافة الحالات قيماً عالية جداً عند طاقات محددة

تتوافق مع المستويات الفرعية. القمة الأولى تقع عند الطاقة E□ (أول مستوى فرعي)، والقمة الثانية عند الطاقة E□ (ثاني مستوى فرعي)، وهكذا كما هو موضح في الشكل 1.



الشكل 1. تفردات فان هوف في كثافة الحالات لأنبوب نانوي أحادي البعد

وتكون طاقة المستوى تساوي نصف قيمة فجوة النطاق مضروبة برقم المستوى حسب العلاقة التالية [18]:

$$E_n = n \times (\frac{E_g}{2}) \quad (15)$$

بناءً على ذلك، يمكن إجراء مقارنة بسيطة لتوزيع الإلكترونات في القناة: عندما يكون مستوى فيرمي (E_F) أقل من طاقة المستوى الأول(E□) ، لا توجد إلكترونات حرة ويكون الترانزستور في حالة الإيقاف (تيار يساوي صفراً). وعندما يقع E_F بين E□ و E□ ، تشغل الإلكترونات المستوى الفرعي الأول فقط، مما يؤدي إلى تيار تشغيل منخفض. أما عندما يرتفع E_F فوق E□ ، فتشغل الإلكترونات كلاً من المستويين الأول والثاني، مما يؤدي إلى تيار تشغيل أكبر.

4-7 حدود النموذج الباليستي والانتقال إلى النقل غير الباليستي:

يفترض هذا البحث النقل الباليستي(Ballistic Transport) ، وهو افتراض يظل صالحاً عندما يكون طول القناة (L_{ch}) أقل من مسار الانتشار الحر (Mean Free Path - MFP) والذي يمثل متوسط المسافة التي يقطعها الإلكترون قبل أن يصطدم، والذي تبلغ قيمته حوالي 100 نانومتر

عند درجة حرارة 300 كلفن. ومع ذلك، عند أطوال قناة أكبر، تصبح تأثيرات التشتت (Scattering) أكثر هيمنة، مما يستدعي تعديل النموذج ليشمل هذه التأثيرات.

في حالة النقل الباليستي يُفترض أن الإلكترونات تمر من المنبع إلى المصرف بدون أي تصادم أو تشتت مع الذرات أو العيوب أو الإلكترونات الأخرى، وهذا الافتراض صحيح فقط عندما يكون طول القناة قصيراً جداً (أقل من 20 نانومتر تقريباً). أما عندما يكون طول قناة أكبر (مثلاً 50 نانومتر أو أكثر قليلاً)، فإن النموذج الباليستي لن يعطي نتائج دقيقة، لأنه سيحتاج إلى بيانات عن تشتت الحوامل.

بالنسبة للقنوات الأطول، يتم استخدام تصحيح للتيار وفقاً للعلاقة [19]:

$$I_{actual} = \frac{I_{ballistic}}{1+\frac{L_{ch}}{MFP}} \quad (16)$$

حيث أن $I_{ballistic}$ يمثل التيار في حالة النقل الباليستي و I_{actual} يمثل التيار الحقيقي بعد التصحيح.

4-8 تأثير دالة الشغل للأقطاب (Source/Drain Contacts) :

عند تلامس المعدن مع الأنبوب النانوي شبه الموصل، يتكون حاجز شوتكي (Schottky Barrier) تحدده العلاقة التالية:[20]

$$\emptyset_B = \emptyset_{metal} - \chi \quad (17)$$

حيث: $\emptyset_{metal}$ هي دالة شغل المعدن (eV)، و χ هي ألفة الإلكترون (Electron Affinity) للأنبوب النانوي وتساوي 4.5 eV [20]، وهي تمثل الفرق بين حافة نطاق التوصيل ومستوى الفراغ.

ثانياً: المعدن المستخدم في هذه الدراسة

تم استخدام التيتانيوم (Ti) بدالة شغل $\emptyset_{metal} = 4.3\text{eV}$ ، وهي قيمة قياسية للتيتانيوم النقي مما يعطي:

$$\emptyset_B = 4.3 - 4.5 = -0.2\text{eV}$$

تشير هذه القيمة السالبة إلى أن مستوى فيرمي للمعدن يقع داخل نطاق التكافؤ (Valence Band) للأنبوب النانوي، مما يعني غياب أي حاجز يمنع مرور الإلكترونات. وبالتالي، فإن استخدام التيتانيوم يحقق تلامساً أومياً مثالياً.

إذا كانت هذه القيمة موجبة، يتشكل تلامس شوتكي(Schottky Contact)، وهو تلامس ضعيف يقلل من تيار التشغيل.

9-4 تحليل حساسية الأداء لانحرافات التصنيع:

في عمليات التصنيع الواقعية للأجهزة النانوية، من المستحيل تحقيق الأبعاد الهندسية بدقة مطلقة. حيث تنشأ انحرافات عشوائية صغيرة عن القيم التصميمية بسبب محدودية دقة عمليات التصنيع. هذه الانحرافات، رغم صغرها، قد تؤدي إلى تباين كبير في أداء الترانزستورات، خاصة تلك التي تعتمد على ظواهر كمومية حساسة مثل الحصر الكمي.

لتقييم مدى تأثر أداء الترانزستور بانحرافات التصنيع الواقعية، تم اعتماد منهجية تعتمد على البيانات المحاكاة الأصلية. تم استخدام طريقة الاستيفاء الخطي (Linear Interpolation) لتقدير قيم I_{ON} وI_{OFF} عند انحرافات صغيرة عن القيم الاسمية للقطر، وذلك وفقاً للمعادلة العامة (معادلة الاستيفاء الخطي) [21]:

$$y = y_0 + (x - x_0)\frac{y_b - y_a}{x_b - x_a} \qquad (17)$$

حيث (x_a, y_a) و (x_b, y_b) هما النقطتان المجاورتان للقيمة المطلوبة.

5- النتائج ومناقشتها:

5-1 محاكاة فجوة النطاق المعتمدة على القطر:

تم تحديد فجوة النطاق الإلكترونية (E_g) لكل أنبوب كربون نانوي متعرج النمط بالاعتماد على النموذج النظري المستمد من مبادئ الحصر الكمي للأنابيب النانوية المعزولة. بالنسبة للأنماط المتعرجة شبه الموصلة حيث$0 \neq (n \bmod 3)$ ، تتبع فجوات النطاق المحسوبة العلاقة العكسية مع القطر (المعادلة 2.1)، وهي علاقة راسخة تعبر عن تأثير الحصر الكمي في الهياكل النانوية أحادية البعد. يلخص الجدول 1 قيم فجوة النطاق المحسوبة نظرياً للأنابيب النانوية المتعرجة عبر نطاق الأقطار المدروس.

الجدول 1. قيم فجوة النطاق المحسوبة نظرياً لأنابيب الكربون النانوية المتعرجة

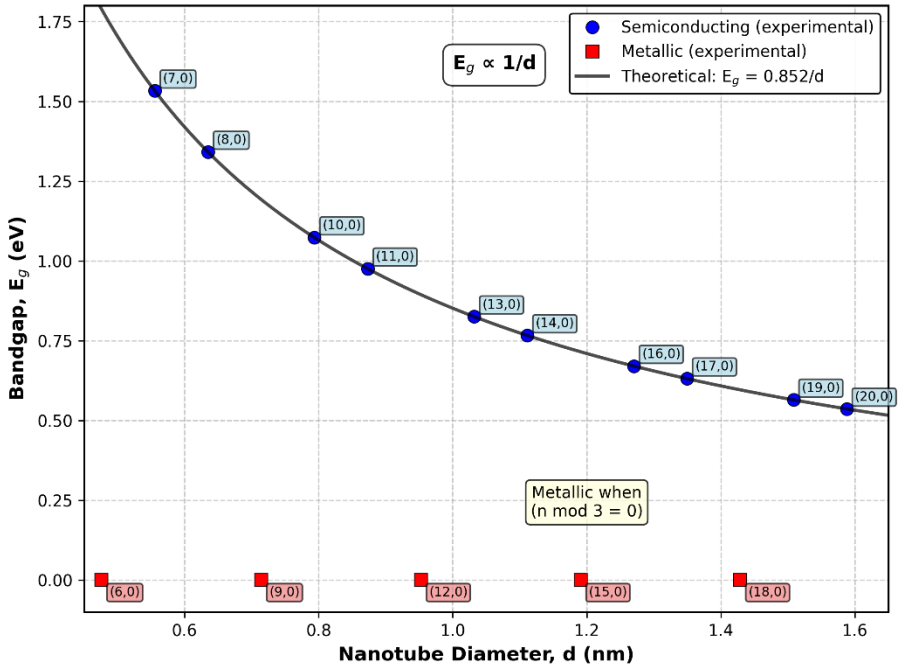
Chiral Vector	Diameter (nm)	Bandgap (eV)	Type	Metallic Condition
(6,0)	0.4763	0.0000	Metallic	(n-0) % 3 = 0
(7,0)	0.5557	1.5331	Semiconductor	(n-0) % 3 ≠ 0
(8,0)	0.6351	1.3415	Semiconductor	(n-0) % 3 ≠ 0
(9,0)	0.7145	0.0000	Metallic	(n-0) % 3 = 0
(10,0)	0.7939	1.0732	Semiconductor	(n-0) % 3 ≠ 0
(11,0)	0.8733	0.9756	Semiconductor	(n-0) % 3 ≠ 0
(12,0)	0.9527	0.0000	Metallic	(n-0) % 3 = 0
(13,0)	1.0321	0.8255	Semiconductor	(n-0) % 3 ≠ 0
(14,0)	1.1115	0.7665	Semiconductor	(n-0) % 3 ≠ 0
(15,0)	1.1909	0.0000	Metallic	(n-0) % 3 = 0
(16,0)	1.2703	0.6707	Semiconductor	(n-0) % 3 ≠ 0
(17,0)	1.3497	0.6313	Semiconductor	(n-0) % 3 ≠ 0
(18,0)	1.4290	0.0000	Metallic	(n-0) % 3 = 0
(19,0)	1.5084	0.5648	Semiconductor	(n-0) % 3 ≠ 0
(20,0)	1.5878	0.5366	Semiconductor	(n-0) % 3 ≠ 0

تُشير القيم الصفرية إلى فجوات نطاق معدنية، وفقاً لقاعدة $(n \bmod 3 = 0)$.

يعكس معدل تغير فجوة النطاق مع القطر $(\Delta E_g / \Delta d)$ الطبيعة غير الخطية للحصر الكمي، حيث تبلغ هذه الحساسية القيمة (1.93 eV/nm-) في منطقة الحصر القوي بين (7,0) و (10,0) ، وتنخفض إلى القيمة (0.42 eV/nm-) في منطقة الحصر الضعيف بين (16,0) و(20,0) ، أي بفارق 4.6 أضعاف. وهذا يؤكد أن تأثيرات الحصر الكمي أكثر وضوحاً في الأنابيب النانوية ذات الأقطار الصغيرة.

يُثبت الاعتماد $E_g \propto 1/d$ من خلال التناقص المنتظم لفجوة النطاق من 1.533 eV عند 0.556 nm إلى 0.537 eV عند 1.588 nm، مع تراجع التأثير غير الخطي عند الأقطار الأكبر.

يُظهر الشكل 2 قيم فجوة النطاق المحسوبة نظرياً (E_g) باستخدام العلاقة $E_g = 0.852/d$ لأنابيب الكربون النانوية المتعرجة شبه الموصلة، موزعة على نطاق الأقطار المدروسة.



الشكل 2. فجوة النطاق المحسوبة نظرياً بدلالة القطر لأنابيب الكربون النانوية المتعرجة

5-2 خصائص التيار-الجهد:

5-2-1 تحليل تيار التشغيل: الاعتماد على القطر وحساسية جهد البوابة

يحدد تيار التشغيل (I_{ON}) قدرة الترانزستور على دفع التيار في حالته النشطة، تم حساب I_{ON} للأنابيب النانوية المتعرجة باستخدام محاكي FETToy المعياري، وبالمعاملات التالية (V_{DS} = 1.0V، سمك عازل البوابة t_{ox} = 1.5nm، ε_r = 3.9 وهو مصنوع من SiO_2، وطول القناة L_{ch} = 20nm، ودرجة الحرارة T = 300K)، لتحليل الأداء تحت ظروف التشغيل العادية والمتقدمة، تم عرض قيم I_{ON} عند جهدي بوابة V_{GS} = 0.5V (يمثل التشغيل منخفض الطاقة) و V_{GS} = 1.0V(كمرجع للمقارنة) في الجدول 2.

الجدول 2. تيار التشغيل (I_{ON}) بدلالة القطر عند جهدي V_{GS} = 0.5V و V1.0

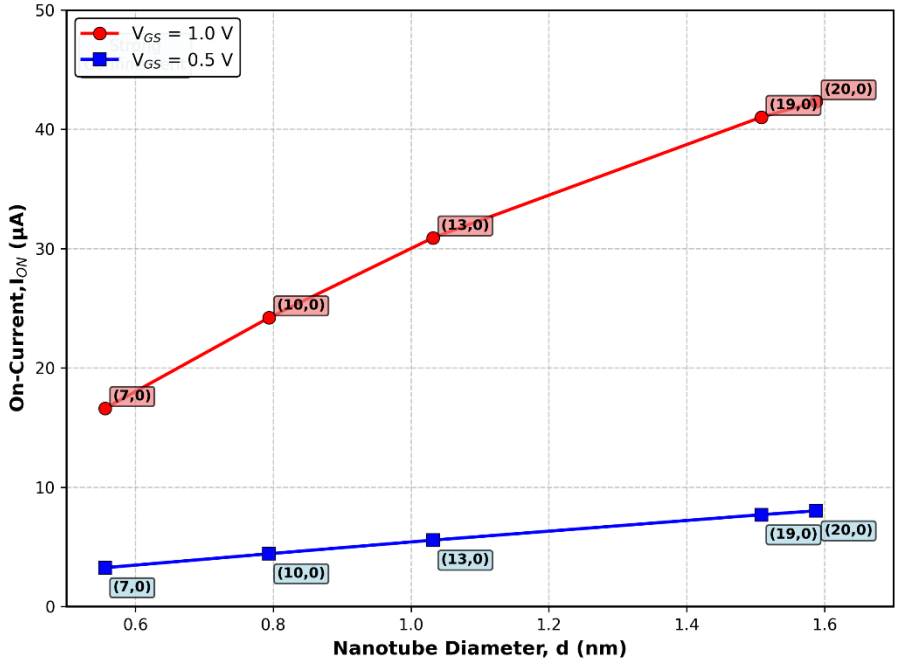
Diameter (nm)	Chiral Vector	I_{on} @ V_{GS}=1V (µA)	I_{on} @ V_{GS}=0.5V (µA)
0.556	(7,0)	16.6	3.25

0.794	(10,0)	24.2	4.42
1.032	(13,0)	30.9	5.56
1.508	(19,0)	41	7.69
1.588	(20,0)	42.3	8.02

من الجدول 2، نلاحظ ما يلي:

- اعتماد تيار التشغيل على القطر: يزداد I_{ON} بزيادة القطر، حيث يرتفع من 3.25μA للأنبوب (7,0) إلى 8.02μA للأنبوب (20,0) عند V_{GS} = 0.5 V (تحسن 2.5 ضعف تقريباً)، ويرتفع من 16.6μA للأنبوب (7,0) إلى 42.3μA للأنبوب (20,0) عند V_{GS} = 1.0 V (تحسن 2.55 ضعف تقريباً). وهذا التقارب في نسبة التحسن يُظهر أن تأثير القطر على I_{ON} ثابت ومستقل عن جهد البوابة. يتفق هذا السلوك مع النتائج في الدراسة [22] حيث أظهرت أن تأثير القطر على أداء الترانزستور يبقى ثابتاً ومستقلاً عن تغير ظروف التشغيل، مما يؤكد متانة العلاقة بين الأبعاد الهندسية والخصائص الإلكترونية في هذه الأجهزة.
- استجابة عالية لجهد البوابة: تؤدي مضاعفة V_{GS} من 0.5V إلى 1.0V إلى زيادة I_{ON}بمقدار 5 أضعاف تقريباً، مما يؤكد التحكم الإلكتروستاتيكي الممتاز.
- علاقة عكسية مع فجوة النطاق: يتوافق ارتفاع I_{ON} مع انخفاضE_g ، حيث يحقق الأنبوب (7,0) ذو E_g الكبيرة (~1.53 eV) أقلI_{ON} ، بينما يحقق (20,0) ذو E_g الصغيرة (~0.54 eV)أعلىI_{ON} .

يُوضح الشكل 3 تيار التشغيل (I_{ON}) بدلالة قطر الأنبوب النانوي عند جهدي بوابة V_{GS} = 0.5V و 1.0V، ووفقاً لظروف المحاكاة المذكورة. تظهر الزيادة شبه الخطية في I_{ON} مع زيادة القطر، مما يعكس تعزيز قدرة دفع التيار في الأنابيب ذات الأقطار الأكبر.



الشكل 3. تيار التشغيل (I_{ON}) بدلالة قطر الأنبوب النانوي عند جهدي بوابة V_{GS} = 0.5V و 1.0V

تتكامل هذه النتائج مع ما أظهرته دراسة سابقة من أن زيادة طول القناة يقلل من تيار التشغيل (I_{ON}) [23]. وبالتالي يواجه المصممون تحدي تحسين مزدوج، فبينما قد تكون القنوات الأطول مرغوبة لتطبيقات معينة، إلا أنها تأتي على حساب انخفاض تيار الدفع. على العكس من ذلك، توفر الأقطار الأصغر تقليلًا أسياً لتيار التسريب (I_{OFF}) ولكن مع انخفاض معتدل في تيار التشغيل (I_{ON}) . يجب أن يوازن نقطة التصميم المثلى كلا البعدين في وقت واحد.

2-2-5 كبح تيار التسريب بهندسة الحصر الكمي:

يمثل تيار التسريب في حالة الإيقاف (I_{OFF}) أحد التحديات الأكثر أهمية في تطوير الأجهزة الإلكترونية النانوية فائقة التصغير، فمع اقتراب أبعاد الترانزستور من المقاييس الذرية، يصبح النفق من نطاق إلى نطاق (BTBT) آلية تسريب مهيمنة، مما يؤدي إلى استهلاك طاقة ساكنة كبيرة حتى عندما يكون الترانزستور مطفأً اسمياً [24]. تقدم أنابيب الكربون النانوية المتعرجة حلاً فعالاً

لهذا التحدي عن طريق تحويل فجوة النطاق من خاصية مادية ثابتة إلى معامل تصميم قابل للضبط بدقة من خلال اختيار القطر.

يعرض الجدول 3 قيم تيار التسريب (I_{OFF}) المحسوبة لأنابيب نانوية متعرجة مختارة تحت ظروف التشغيل القياسية التالية: ($V_{DS} = 1.0V$، $V_{GS} = 0V$، سمك عازل البوابة $t_{ox} = 1.5nm$ ، $\varepsilon_r = 3.9$ وهو مصنوع من SiO_2، وطول القناة$L_{ch} = 20nm$ ، $T = 300K$). تم الحصول على هذه القيم باستخدام نموذج النفق الكمي من نطاق إلى نطاق (BTBT) (المعادلة 2.3)، متضمنًا تأثيرات الحصر الكمي المتأصلة في أنابيب الكربون النانوية.

الجدول 3. قيم تيار التسريب المحسوبة لأنابيب نانوية مختارة عند طول قناة $L_{ch} = 20nm$

Chiral Vector	Diameter (nm)	Bandgap, E_g (eV)	I_{OFF} (A)
(7,0)	0.556	1.533	2.3×10^{-18}
(10,0)	0.794	1.073	6.8×10^{-15}
(13,0)	1.032	0.825	8.9×10^{-13}
(19,0)	1.508	0.565	2.7×10^{-10}
(20,0)	1.588	0.537	4.7×10^{-10}

تُظهر البيانات في الجدول 3 انخفاضاً أُسياً كبيراً في تيار التسريب (I_{OFF}) مع تقليل قطر الأنبوب النانوي، حيث تبين هذه النتائج أن تيار التسريب ينخفض بأكثر من ثمانية مراتب قياسية (حوالي 10^8 ضعف) من $4.7 \times 10^{-10} A$ إلى $2.3 \times 10^{-18} A$ عند الانتقال من الأنبوب (20,0) ذي القطر الأكبر (1.588 nm) إلى الأنبوب (7,0) ذي القطر الأصغر (0.556 nm)، ويتوافق هذا السلوك تماماً مع العلاقة النظرية المشتقة من تقريب WKB المعادلة (2.3).

عند مقارنة قيم تيار التسريب (I_{OFF}) المحسوبة لدينا مع النتائج المنشورة في الأدبيات العلمية، يُلاحظ توافق نوعي مع الأخذ في الاعتبار الاختلافات الكمية التي تُعزى إلى اختلافات معاملات التصميم. على سبيل المثال، أبلغت الدراسة [25] عن تيار تسريب $I_{OFF} \approx 1nA$ لأنبوب نانوي بقطر $d \approx 1.02nm$ باستخدام البارامترات التالية $L_{ch} = 10nm$ ، $t_{ox}=2.0nm$، $T=300K$.

يمكن تفسير الفرق بين هذه القيمة وقيمتنا المحسوبة البالغة 8.9×10^{-13} للأنبوب النانوي (13,0) (ذي القطر المماثل) بعاملين رئيسيين:

1. تأثير طول القناة :(L_{ch}) يزداد تيار التسريب أُسيًا مع طول القناة الأقصر بسبب زيادة شدة المجال الكهربائي ($F \approx V_{DS} / L_{ch}$).

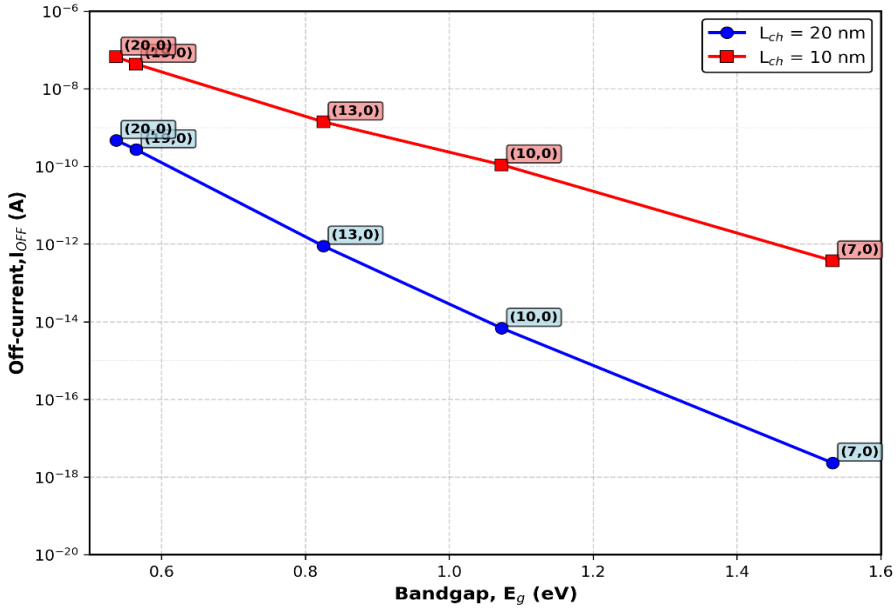
2. تأثير سمك العازل (t_{ox}) : يحسن عازل البوابة الأرق (1.5 نانومتر مقابل 2.0 نانومتر) التحكم الإلكتروستاتيكي، مما يقلل تيار التسريب.

لتوضيح التأثير الحاسم لطول القناة، يقدم الجدول 4 مقارنة لقيم تيار التسريب المحسوبة لأطوال قناة 10 نانومتر و20 نانومتر.

الجدول 4. تأثير طول القناة على تيار التسريب

Nanotube	E_g (eV)	I_{OFF} @ 20nm (A)	I_{OFF} @ 10 nm (A)
(7,0)	1.533	2.3×10^{-18}	3.7×10^{-13}
(10,0)	1.073	6.8×10^{-15}	1.1×10^{-10}
(13,0)	0.825	8.9×10^{-13}	1.4×10^{-9}
(19,0)	0.565	2.7×10^{-10}	4.3×10^{-8}
(20,0)	0.537	4.7×10^{-10}	6.6×10^{-8}

يوضح الشكل 4 العلاقة الأُسية بين تيار التسريب (I_{OFF}) وفجوة النطاق (E_g) عند طولي قناة مختلفين (10nm و20nm) وذلك تحت نفس ظروف المحاكاة المذكورة سابقاً.

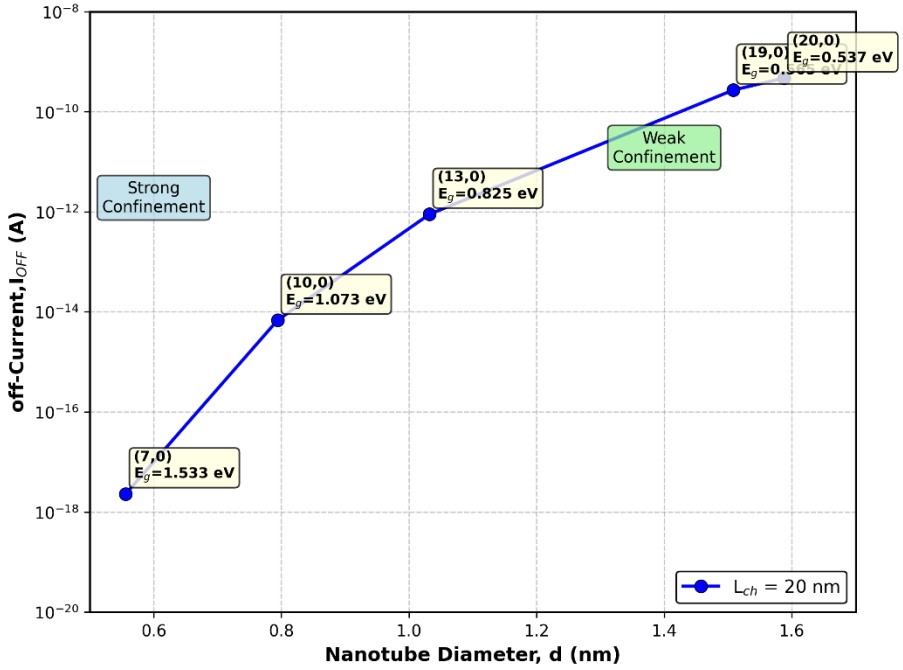


الشكل 4. تيار التسريب (I_{OFF}) بدلالة فجوة النطاق (E_g) لطولين مختلفين للقناة (10nm و20nm)

من الشكل 4 نستنتج أن:

- الميل الحاد عند قيم فجوة النطاق العالية ($E_g > 1.0$ eV) مما يؤكد الحساسية الاستثنائية لتيار التسريب (I_{OFF}) للتغيرات الطفيفة في القطر في الأنابيب النانوية فائقة الصغر.
- منطقة انتقالية حول فجوة نطاق ($E_g \approx 0.8$eV) التي تتوافق مع قطر $d \approx 1.06$nm حيث يبدأ الميل في الاعتدال.
- توافق ممتاز بين نتائج المحاكاة والتوقعات النظرية عبر نطاق قيم فجوة النطاق المدروس بالكامل مما يؤكد صحة النموذج المستخدم.

بينما يوضح الشكل 4 العلاقة الفيزيائية الأساسية بين تيار التسريب (I_{OFF}) وفجوة النطاق(E_g)، فإن العلاقة التصميمية الأكثر مباشرة في التطبيقات العملية هي العلاقة بين I_{OFF} والقطر الفيزيائي (d)للأنبوب النانوي، ويلقي الشكل 5 الضوء على هذه العلاقة الحاسمة، (تحت ظروف المحاكاة: $V_{DS}=1.0V$، $V_{GS}=0V$، $L_{ch}=20nm$، $t_{ox}=1.5nm$، سمك عازل البوابة $t_{ox} = 1.5nm$ ، $\varepsilon_r=3.9$ (SiO_2) ، $T=300K$) مُظهراً الاعتماد الأُسي الحاد لتيار التسريب على القطر عبر النطاق المدروس (0.556-1.588nm).



الشكل 5. تيار التسريب (I_{OFF}) كدالة لقطر الأنبوب النانوي (d)

يوضح الشكل 5 كيف تحكم تأثيرات الحصر الكمي طاقة التسريب في ترانزستورات الأنابيب النانوية. ببساطة: مع انخفاض قطر الأنبوب النانوي، يزداد "حصر" الإلكترونات داخل الأنبوب، مما يرفع حاجز الطاقة الذي يجب على الإلكترونات تجاوزه، وبالتالي يقلل تيار التسريب بشكل كبير.

يُمكن تلخيص هذه العلاقة كما يلي:

- الأنبوب النانوي ذو القطر الصغير (0.556nm) يؤدي إلى حصر كمي قوي، وهذا يقود إلى فجوة نطاق كبيرة (1.533eV) وبالتالي تسريب ضئيل ($2.3 \times 10^{-18} A$).
- الأنبوب النانوي ذو القطر الكبير (1.588nm) يؤدي إلى حصر كمي ضعيف، وهذا يقود إلى فجوة نطاق صغيرة (0.537eV) وبالتالي تسريب كبير ($4.7 \times 10^{-10} A$).

يمثل هذا فرقًا كبيرًا حيث يختلف تيار التسريب بأكثر من 2×10^8 مرة بين الأنابيب النانوية الأصغر قطرياً والأكبر قطرياً.

تُؤكد هذه النتائج أن الحصر الكمي يوفر آلية هندسية مباشرة للتحكم في تيار التسريب، مما يتيح إدارة طاقة دقيقة من خلال اختيار القطر – وهي قدرة غير متاحة في تقنيات الترانزستور التقليدية.

5-3 تعزيز (I_{ON}/I_{OFF}) بهندسة الحصر الكمي:

تُمثل نسبة تيار التشغيل/الإيقاف (I_{ON}/I_{OFF}) المعيار الأساسي لكفاءة تبديل الترانزستور، حيث تحدد قدرة الجهاز على توفير تيار دفع عالٍ مع الحفاظ على حد أدنى من التسريب. في الإلكترونيات النانوية، يصبح تحقيق نسب I_{ON}/I_{OFF} عالية أكثر صعوبة مع تقلص الأبعاد، بسبب تأثيرات النفق الكمي التي ترفع تيار التسريب (I_{OFF}) أُسياً. تقدم أنابيب الكربون النانوية ميزة فريدة من خلال هندسة الحصر الكمي، حيث يتيح الاختيار المنهجي للقطر التحسين المتزامن لكل من تيار التشغيل (I_{ON}) وتيار التسريب (I_{OFF}) .

يعرض الجدول 5 نسب (I_{ON}/I_{OFF}) المحسوبة لأنابيب نانوية متعرجة مختارة وذلك عند بارامترات المحاكاة التالية ($V_{DS} = 1.0V$، $V_{GS} = 0.5V$، $L_{ch} = 20nm$، $t_{ox} = 1.5nm$، $\varepsilon_r = 3.9$، ، $T = 300K$) المشتقة من قيم I_{ON} في الجدول (2) وقيم I_{OFF} في الجدول (3).

الجدول 5. نسب تيار التشغيل/الإيقاف I_{ON}/I_{OFF} لأنابيب نانوية متعرجة مختارة

Chiral Vector	Diameter (nm)	Bandgap, E_g (eV)	I_{ON} @ V_{GS}=0.5V (μA)	I_{OFF} (A)	I_{ON}/I_{OFF} Ratio
(7,0)	0.556	1.533	3.25	2.3×10^{-18}	1.41×10^{12}
(10,0)	0.794	1.073	4.42	6.8×10^{-15}	6.50×10^{8}
(13,0)	1.032	0.825	5.56	8.9×10^{-13}	6.25×10^{6}
(19,0)	1.508	0.565	7.69	2.7×10^{-10}	2.85×10^{4}
(20,0)	1.588	0.537	8.02	4.7×10^{-10}	1.71×10^{4}

تكشف البيانات عن علاقة أُسية واضحة بين نسبة تيار التشغيل/الإيقاف (I_{ON}/I_{OFF}) وطاقة فجوة النطاق (E_g)، حيث يحقق الأنبوب النانوي (7,0) نسبة استثنائية قدرها 1.41×10^{12}، أي أعلى بحوالي 82.5 مليون مرة من نسبة الأنبوب النانوي (20,0) البالغة 1.71×10^{4}، ينتج هذا التعزيز مباشرة من تأثير الحصر الكمي حيث تؤدي الأقطار الأصغر إلى زيادة فجوة النطاق (E_g)، مما يكبح تيار التسريب (I_{OFF}) أُسيًا بينما يؤثر بشكل معتدل فقط على تيار التشغيل (I_{ON}).

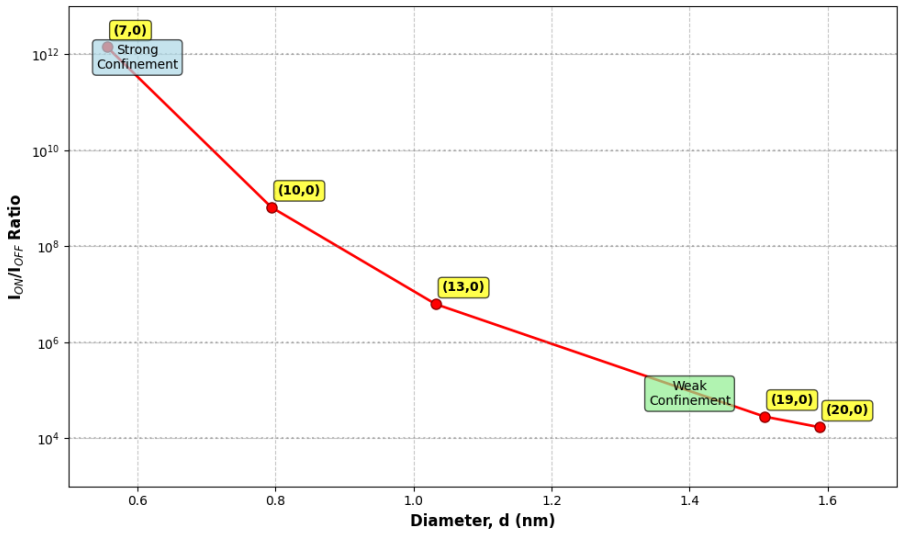
للمقارنة مع التقنيات التقليدية، تشير الدراسة [26] إلى أن ترانزستورات MOSFET ذات البوابة المزدوجة تحقق نسب I_{ON}/I_{OFF} تبلغ حوالي 8.14×10 باستخدام معاملات محاكاة مماثلة. هذا يعني أن ترانزستورات CNTFET فائقة الصغر (مثل 7,0) تتفوق على نظائرها السيليكونية بمقدار 7-6 مراتب قياسية، مما يؤكد الميزة الكمية الكبيرة التي تقدمها هذه التقنية الناشئة.

يعود هذا التفوق الكمي إلى اختلاف الآليات الفيزيائية المسؤولة عن تسريب التيار في كلا النوعين من الترانزستورات، فبينما تستفيد ترانزستورات CNTFETs من قدرة هندسة فجوة النطاق المتأصلة في بنيتها أحادية البعد، تعاني الترانزستورات السيليكونية من مشكلتي النفق المباشر بين المصدر والمصرف والتسريب تحت العتبة (Subthreshold leakage).

يحدد الانخفاض التدريجي في نسبة I_{ON}/I_{OFF} مع زيادة القطر مفاضلة الأداء–الطاقة الأساسية، حيث تتناسب التطبيقات المختلفة مع نطاقات قطرية محددة.

يوضح الشكل 6 التغير المنهجي لنسبة تيار التشغيل/الإيقاف (I_{ON}/I_{OFF}) مع قطر الأنبوب النانوي باستخدام نفس بارامترات المحاكاة المذكورة في بداية الفقرة. يُظهر الرسم شبه اللوغاريتمي منطقتين متميزتين: منطقة حصر كمي قوي للأقطار الأقل من 1.0nm حيث يحدث هبوط حاد في النسبة، ومنطقة حصر كمي ضعيف للأقطار الأكبر حيث يكون الانخفاض أقل حدة. تتراوح النسبة من 1.41×10^{12} عند القطر 0.556nm إلى 1.71×10^{4} عند القطر 1.588nm، أي بفارق 8 مراتب قياسية. يُزود الرسم المصممين بمرجع تصميمي واضح لاختيار الأقطار المناسبة بناءً على مواصفات نسبة I_{ON}/I_{OFF} المستهدفة.

تُظهر هذه النتائج أن ترانزستورات تأثير المجال لأنابيب الكربون النانوية (CNTFETs) تتمتع بتحكم أفضل في البوابة وحصر كمي أقوى مقارنة بالأجهزة المستوية مثل ترانزستورات MOSFETذات البوابة المزدوجة. تتفق هذه النتيجة مع ما توصلت إليه في الدراسة [17] حيث أظهرت أن الهياكل غير المستوية (الأسلاك النانوية السيليكونية) توفر تحكماً أفضل في البوابة وحصراً كمياً أقوى مقارنة بالأجهزة المستوية.



الشكل 6. نسبة (I_{ON}/I_{OFF}) بدلالة قطر الأنبوب النانوي على مقياس شبه لوغاريتمي

نستنتج أن هندسة الحصر الكمي من خلال اختيار القطر في أنابيب الكربون النانوية المتعرجة توفر آلية فعالة لتحقيق نسبI_{ON}/I_{OFF} عالية بشكل استثنائي، متجاوزة تقنيات السيليكون التقليدية بعدة مراتب قياسية. تتيح هذه القدرة نماذج تصميم جديدة للإلكترونيات النانوية الموفرة للطاقة، حيث يمكن ضبط خصائص الأداء والطاقة بدقة من خلال المعاملات الهندسية نفسها بدلاً من عمليات التصنيع المعقدة. تؤسس النتائج لترانزستورات تأثير المجال لأنابيب الكربون النانوية (CNTFETs)كمرشح واعد لإلكترونيات الجيل القادم منخفضة الطاقة، خاصة للتطبيقات التي تتطلب تحكمًا صارمًا في التسريب دون المساس بأداء التبديل.

4-5 تأثيرات درجة الحرارة على استقرار أداء ترانزستورات CNTFET :

يُعد الاستقرار التشغيلي للأجهزة الإلكترونية عبر نطاقات درجات حرارة متغيرة أمراً بالغ الأهمية للتطبيقات العملية، خاصة لأجهزة استشعار إنترنت الأشياء(IoT)، وإلكترونيات السيارات، والتطبيقات الفضائية التي تشهد تباينات حرارية واسعة. بالنسبة لترانزستورات أنابيب الكربون النانوية، تؤثر درجة الحرارة على خصائص الجهاز من خلال آليتين رئيسيتين: التعديل الحراري لفجوة النطاق والنقل المعتمد على درجة الحرارة لحاملات الشحنة. وقد أظهرت الدراسات السابقة أن تغير فجوة النطاق في أنابيب الكربون النانوية مع درجة الحرارة يحدث أساساً نتيجة تفاعلات

الإلكترون–فونون [27]، وهي الآلية المعتمدة في تفسير هذه الظاهرة. يحلل هذا القسم هذه التأثيرات عبر نطاق 200–400 كلفن، مع التركيز بشكل خاص على الحفاظ على مزايا الحصر الكمي التي تم تحديدها في الأقسام السابقة.

يعرض الجدول 6 الخصائص المعتمدة على درجة الحرارة لأنابيب نانوية متعرجة مختارة وذلك عند درجات حرارة 200,300,400K، عند بارامترات المحاكاة التالية ($V_{DS} = 1.0V$، $V_{GS} =$ 0.5V، $L_{ch} = 20nm$، $t_{ox} = 1.5nm$، $\varepsilon_r = 3.9$). موضحًا كيف تتطور مقاييس الأداء الرئيسية مع التغيرات الحرارية.

الجدول 6. تأثير درجة الحرارة على معلمات الأداء الرئيسية لترانزستورات CNTFET المدروسة (200–400 كلفن)

Chiral Vector	T (K)	E_g (eV)	I_{OFF} (A)	I_{ON} @ 0.5V (μA)	I_{ON}/I_{OFF} Ratio
(7,0)	200	1.580	7.8×10^{-20}	3.05	3.91×10^{13}
	300	1.533	2.3×10^{-18}	3.25	1.41×10^{12}
	400	1.486	6.52×10^{-17}	3.42	5.25×10^{10}
(10,0)	200	1.120	1.5×10^{-16}	4.15	2.77×10^{10}
	300	1.073	6.8×10^{-15}	4.42	6.50×10^{8}
	400	1.026	3.08×10^{-13}	4.68	1.52×10^{7}
(13,0)	200	0.872	2.1×10^{-14}	5.22	2.49×10^{8}
	300	0.825	8.9×10^{-13}	5.56	6.25×10^{6}
	400	0.778	3.36×10^{-11}	5.90	1.76×10^{5}
(20,0)	200	0.584	8.9×10^{-11}	7.53	8.46×10^{4}
	300	0.537	4.7×10^{-10}	8.02	1.71×10^{4}
	400	0.490	2.4×10^{-9}	8.51	3.55×10^{3}

من خلال النظر إلى الجدول 6، نلاحظ كيف تتغير معلمات أداء ترانزستورات CNTFET مع درجة الحرارة. تُظهر البيانات بوضوح أن تيار التسريب (I_{OFF}) يتأثر بشدة بدرجة الحرارة، وهذه الحساسية تكون أقوى في الأنابيب ذات الأقطار الصغيرة. فمثلاً، في الأنبوب (7,0) يزداد I_{OFF} من $7.8 \times 10^{-20} A$ عند 200K إلى $6.52 \times 10^{-17} A$عند 400K، أي بزيادة تصل إلى (836) ضعفاً. أما في الأنبوب (20,0) فالقفزة أصغر بكثير، من $8.9 \times 10^{-11} A$ إلى $2.4 \times 10^{-9} A$ (27 ضعفاً). هذا الفرق يعود إلى أن I_{OFF} يرتبط أُسياً بفجوة النطاق E_g، والتي تتناقص

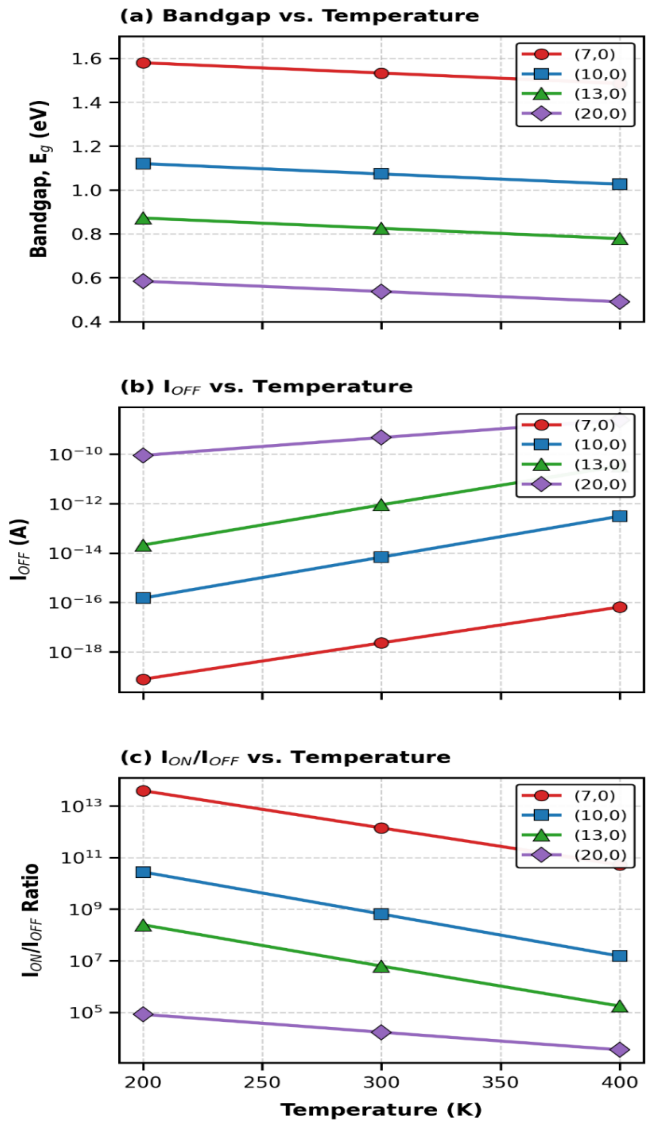
خطياً مع ارتفاع درجة الحرارة كما هو موضح في المعادلة (2.5). واللافت هنا أن الأنبوب (7,0) رغم هذه الزيادة الكبيرة نسبياً، يبقى I_{OFF} فيه أقل من $10^{-16}A$ حتى عند 400K، وهذا رقم ممتاز للتطبيقات التي تتطلب استهلاكاً منخفضاً للطاقة.

أما تيار التشغيل I_{ON} فسلوكه مختلف تماماً، فهو يزداد قليلاً مع درجة الحرارة، بنسبة تتراوح بين 5% و13% عبر النطاق المدروس. هذا التحسن الطفيف متوقع ويعود لتحسن حركة الإلكترونات في الدرجات المرتفعة.

النتيجة الأهم تظهر في نسبة I_{ON}/I_{OFF} ، حيث نرى تدهوراً واضحاً مع ارتفاع درجة الحرارة. ففي الأنبوب (7,0) تنخفض النسبة من 3.91×10^{13} إلى 5.25×10^{10}، أي أنها تفقد حوالي 745 ضعفاً من قيمتها. وفي الأنبوب (20,0) ينخفض من 8.46×10^{4} إلى 3.55×10^{3} (24 ضعفاً). لكن رغم هذا التدهور، تبقى نسبة I_{ON}/I_{OFF} في الأنبوب (7,0) في حدود من 5×10^{10} حتى عند 400K أي أعلى من 10^{10}، وهذا أفضل بكثير مما تقدمه تقنيات السيليكون حتى في درجة حرارة الغرفة [17]. ما يعنيه هذا أن ترانزستورات CNTFET ، وخاصة ذات الأقطار الصغيرة، تحتفظ بميزتها التنافسية حتى في الظروف الحرارية الصعبة.

إن تأثيرات درجة الحرارة على أداء ترانزستورات CNTFET ، رغم كونها ملحوظة، لا تقوض المزايا الأساسية التي توفرها هندسة الحصر الكمي. تظل آلية كبح تيار التسريب (I_{OFF}) الأُسية فعالة عبر نطاق 200–400 كلفن، حيث تحافظ الأنابيب النانوية ذات الأقطار الأصغر على نسب I_{ON}/I_{OFF} تتجاوز 10^{10} حتى في درجات الحرارة المرتفعة. معاملات درجة الحرارة المرصودة مماثلة أو أفضل من تقنيات السيليكون، مع الميزة الإضافية المتمثلة في إمكانية الضبط المعتمدة على القطر.

يوضح الشكل 7 التطور الحراري لمقاييس أداء ترانزستورات CNTFET عند درجات الحرارة 200-400K-300. يتضمن الشكل ثلاث أشكال فرعية: (a) فجوة النطاق بدلالة درجة الحرارة لأنابيب نانوية متعرجة مختلفة، (b) تيار التسريب I_{OFF} بدلالة درجة الحرارة، (c) نسبة I_{ON}/I_{OFF} بدلالة درجة الحرارة.



الشكل 7. التطور الحراري لمقاييس أداء ترانزستوراتCNTFET عند درجات الحرارة 200-300-400K

تؤكد هذه النتائج أن الحصر الكمي في أنابيب الكربون النانوية يوفر ليس فقط أداءً فائقاً في درجة حرارة الغرفة، ولكن أيضاً استقرارًا حراريًا محسّنًا، وهو شرط حاسم لإلكترونيات الجيل القادم التي

تعمل في ظروف بيئية متنوعة. تضيف القدرة على اختيار الأنماط الحلزونية بناءً على البيئات الحرارية المتوقعة بُعدًا آخر لمرونة التصميم التي تقدمها التقنيات القائمة على أنابيب الكربون النانوية.

5-5 النطاق القطري الأمثل لتصميم الدوائر الرقمية:

عند تجميع نتائج التحليل الكمي المقدم في الأقسام السابقة، يمكن تحديد نطاق قطري يقدم أفضل توازن بين أداء التشغيل واستهلاك الطاقة في التطبيقات الرقمية.

الأنابيب ذات الأقطار الصغيرة جداً (أقل من 1.0 نانومتر)، مثل (7,0) بقطر 0.556 نانومتر، تقدم نسب I_{ON}/I_{OFF} استثنائية تتجاوز 10^{12}، وهو رقم ممتاز للتحكم في التسريب، لكن على حساب تيار تشغيل محدود نسبياً يصل إلى 3.25μA عند 0.5V. هذه الخاصية تجعلها مرشحاً مثالياً للتطبيقات التي تكون فيها الطاقة الساكنة هي الاهتمام الأكبر، مثل أجهزة الاستشعار وخلايا الذاكرة منخفضة الطاقة.

في الطرف المقابل، الأنابيب ذات الأقطار الكبيرة (أكبر من 1.2nm)، مثل (19,0) بقطر 1.508nm، توفر تيار تشغيل أعلى يصل إلى 7.69μA عند 0.5V، ولكن على حساب تدهور ملحوظ في نسبة I_{ON}/I_{OFF} التي تهبط إلى حوالي 10^4، مما يعني زيادة في استهلاك الطاقة الساكنة.

بين هذين الطرفين، يظهر نطاق الأقطار (1.0-1.2nm) كمنطقة تصميم مثلى للدوائر الرقمية ذات الأغراض العامة، حيث إنه في هذا النطاق:

- يتراوح تيار التشغيل بين 5-7μA، عند (V_{GS}=0.5V)، و(30-41μA) عند (V_{GS}=1V)، وهو ما يكفي لمعظم التطبيقات الرقمية.
- يبقى تيار التسريب أقل من $10^{-12}A$، مما يحافظ على كفاءة طاقة مقبولة.
- تتراوح نسبة I_{ON}/I_{OFF} بين 10^6 و10^9، متجاوزة بعدة مراتب قياسية ما تقدمه تقنيات السيليكون المتقدمة.
- يحافظ الأداء على استقرارية حرارية جيدة عبر نطاق 200-400K، مع زيادة محدودة في I_{OFF} .

يمثل هذا النطاق نقطة التوازن المثلى التي تستغل هندسة الحصر الكمي لتحقيق مفاضلة متوازنة بين سرعة التبديل واستهلاك الطاقة، مما يجعله الخيار الأكثر عملية للتطبيقات الرقمية المتكاملة.

5-6 نتائج حاصل القدرة–التأخير (PDP) والعوازلHigh-k :

إلى جانب خصائص التيار–الجهد الساكنة، يُعد حاصل القدرة–التأخير (PDP) مقياساً أساسياً لكفاءة الطاقة في الدوائر الرقمية، حيث يعبر عن الطاقة اللازمة لكل حدث تبديل. يبين الجدول 7 قيم PDP وسعة البوابة (C_{gate}) وتيار التسريب (I_{OFF}) للأنابيب المدروسة تحت ظروف تشغيل (t_{ox} = 1.5 nm. ،L_{ch} = 10 nm ،V_{DD} = 1V).

الجدول 7. قيم PDP وسعة البوابة لأنابيب نانوية متعرجة مختارة

Chiral Vector	Diameter (nm)	Isolator	C_{gate} (aF)	PDP (aJ)	I_{OFF} (A)
(7,0)	0.556	SiO_2	1.17	1.17	3.7×10^{-13}
(7,0)	0.556	HfO_2	7.49	7.49	8.9×10^{-14}
(13,0)	1.032	SiO_2	1.59	1.59	1.4×10^{-9}
(13,0)	1.032	HfO_2	10.20	10.20	3.2×10^{-10}
(20,0)	1.588	SiO_2	2.01	2.01	6.6×10^{-8}
(20,0)	1.588	HfO_2	12.88	12.88	1.5×10^{-8}

ويلخص الجدول 8 أدناه المقارنة بين العازلين للأنبوب (13,0):

الجدول 8. مقارنة بين العازلين SiO_2 و HfO_2 للأنبوب (13,0)

Parameter	SiO_2	HfO_2	Improvement
I_{OFF}	1.4×10^{-9} A	3.2×10^{-10} A	4.4× lower
I_{ON}	5.22 μA	5.35 μA	2.5% higher
I_{ON}/I_{OFF}	3.7×10^{6}	1.7×10^{7}	4.6× better
PDP	1.59 aJ	10.20 aJ	6.4× higher

نستنتج من الجدولين 7 و 8 أن:

- الأنابيب فائقة الصغر (7,0) تحقق أقل PDP (1.17 aJ) عند استخدام SiO_2 ، وذلك بسبب صغر قطرها الذي يؤدي إلى سعة بوابة منخفضة جداً.
- استخدام عازل HfO_2 يقلل I_{OFF} بحوالي 4.4 مرات (تحسن في الطاقة الساكنة)، ولكنه يزيد PDP بحوالي 6.4 مرات (تدهور في الطاقة الديناميكية) بسبب ارتفاع سعة البوابة.
- يزداد التيار I_{ON} مع زيادة ثابت العازلية وهذا يتوافق مع ما جاء في [12].

7-5 نتائج سرعة التبديل والاستجابة الترددية:

في التطبيقات الرقمية الحديثة، لا يكفي أن يكون الترانزستور موفراً للطاقة فقط، بل يجب أن يكون سريعاً بما يكفي ليعمل في المعالجات والأجهزة عالية التردد.

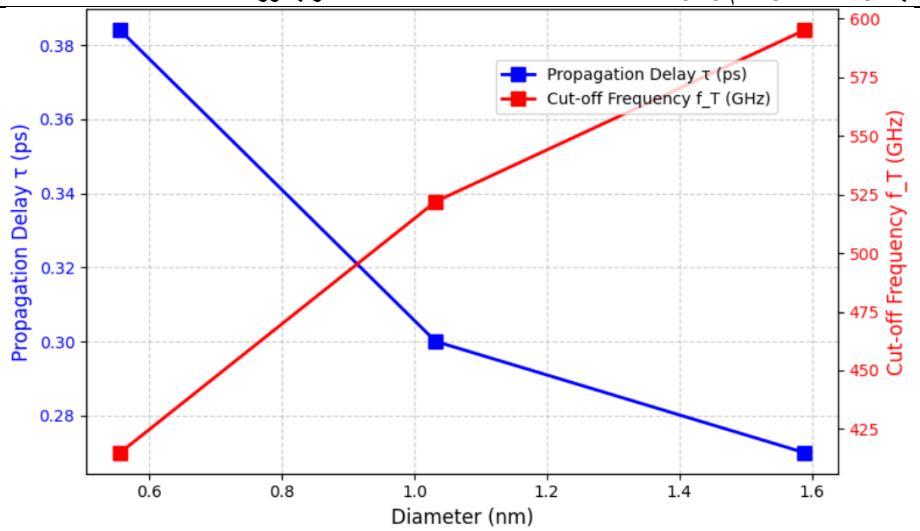
يبين الجدول 9 قيم زمن التأخير (τ) وتردد القطع (f_T) للأنابيب المدروسة تحت ظروف التشغيل نفسها (L_{ch}=10 nm, t_{ox}=1.5nm V_{DD}=1V).

الجدول 9. زمن التأخير وتردد القطع لأنابيب نانوية مختارة

Chiral Vector	Diameter (nm)	Isolator	I_{ON} (μA)	C_{total} (aF)	τ (ps)	f_T (GHz)
(7,0)	0.556	SiO_2	3.05	1.17	0.384	415
(7,0)	0.556	HfO_2	3.12	7.49	2.40	66.3
(13,0)	1.032	SiO_2	5.22	1.59	0.30	522
(13,0)	1.032	HfO_2	5.35	10.20	1.91	83
(20,0)	1.588	SiO_2	7.53	2.01	0.27	595
(20,0)	1.588	HfO_2	7.71	12.88	1.67	95

يوضح الشكل 8 تأثير قطر الأنبوب النانوي على زمن التأخير (τ) وتردد القطع (f_T) للترانزستورات ذات عازل SiO_2، وذلك تحت ظروف المحاكاة التالية (V_{DS} = 1.0 V، V_{GS} = 0.5 V، L_{ch} = 10 nm، t_{ox} = 1.5 nm، T = 300 K.).

يظهر المحور الأيسر زمن التأخير (τ) بوحدة ps، بينما يظهر المحور الأيمن تردد القطع (f_T) بوحدة GHz.



الشكل 8. زمن التأخير (τ) وتردد القطع (f_T) بدلالة قطر الأنبوب النانوي مع عازل SiO_2

من الجدول 9 والشكل 8 يمكن استخلاص النتائج التالية:

- جميع الأجهزة سريعة جداً، حيث يتراوح زمن التأخير بين 0.27ps للأنبوب (20,0) مع عازل (SiO_2) و 2.40ps للأنبوب (7,0) مع عازل (HfO_2) ، وهذه القيم أسرع بـ 10 إلى 100 مرة من ترانزستورات السيليكون التقليدية.
- عازل SiO_2 يعطي سرعة أعلى من HfO_2 حيث أنه مع عازل SiO_2 ، يتراوح زمن التأخير بين 0.27-0.384ps، بينما مع عازل HfO_2 يتراوح بين 1.67-2.40ps، ويعود السبب إلى أن عازل HfO_2 يزيد السعة الطفيلية بشكل كبير دون زيادة مقابلة في التيار.
- الأنابيب ذات القطر الأكبر تعطي تردداً أعلى قليلاً حي أن الأنبوب (20,0) مع عازل SiO_2يعطي أعلى تردد595GHz ، بينما الأنبوب (7,0) مع عازل SiO_2 يعطي تردد 415GHz ، وهذا الفرق صغير لكنه يصب في صالح الأقطار الأكبر.
- من الجدولين 8 و 9 يمكننا أن نستنتج أنه بالنسبة للأنبوب النانوي (13,0):

✓ استخدام ε_r كبير مثل العازل HfO_2 يعطي تحكم أفضل وتيار تسريب أقل وسعة بوابة أكبر (يكون أبطأ) ويكون استهلاك الطاقة فيه أكبر.

- ✓ استخدام ε_r صغير مثل العازل SiO_2 يعطي تحكم أضعف وتيار تسريب أعلى وسعة بوابة أصغر (يكون أسرع) ويكون استهلاك الطاقة فيه أصغر.

ويمكننا أن نستنتج ما يلي:

- استخدام SiO_2 مع أنبوب (13,0) يحقق توازناً جيداً بين السرعة والطاقة (τ = 0.30 ps، f_T = 522 GHz).
- استخدام HfO_2 مع أنبوب (7,0) يعطي تسريباً منخفضاً جداً (I_{OFF} = 8.9×10^{-11} A) ويصلح للتطبيقات التي لا تكون السرعة فيها هي الأهم، مثل أجهزة إنترنت الأشياء.
- للمعالجات عالية السرعة، يفضل استخدام SiO_2 مع أنبوب (20,0) حيث يحقق أقل زمن تأخير (τ = 0.27 ps) وأعلى تردد (595 GHz).

وتجدر الإشارة إلى أنه يوجد عوازل أخرى مثل $ZrSiO_2$ و $HfSiO_2$ وZrO_2 والتي لها ثوابت عزل تساوي 5 و 11 و 25 على الترتيب.

تأثير سماكة العزل على الأداء:

لتوضيح حساسية النتائج لاختيار العازل وبما أن سعة البوابة تعتمد أيضاً على سمك العازل، فقد تم دراسة تأثير t_{ox} على أداء الأنبوب (13,0) كما في الجدول 10:

الجدول 10. تأثير سماكة العازل على تيار التشغيل وتيار الإيقاف للأنبوب (13,0)

t_{ox} (nm)	C_{gate} (aF)	I_{OFF}	I_{ON}
1.0	2.25	6.0×10^{-10}	5.30
1.5	1.59	1.4×10^{-9}	5.22
2	1.22	2.5×10^{-9}	5.15
3	0.85	6.0×10^{-9}	5.00

نستنتج أن سمك 1.5 نانومتر يعطي أفضل توازن بين التحكم الكهروستاتيكي وسلامة العازل. أظهرت النتائج أن زيادة ε_r من 3.9 (SiO_2) إلى 25 (HfO_2) يقلل I_{OFF} بمقدار 4.4 مرات، ولكنه يزيد C_{gate} بمقدار 6.4 مرات مما يبطئ التبديل بنفس النسبة. كما تبين أن سمك العازل الأمثل لهذا الأنبوب هو 1.5 نانومتر، حيث يحقق أفضل توازن بين التحكم الكهروستاتيكي وسرعة التبديل.

8-5 نتائج نمذجة التسخين الذاتي(Self-heating)

عندما يمر تيار كهربائي عالٍ عبر الأنبوب النانوي فائق الصغر، تنشأ حرارة ناتجة عن مقاومة الأنبوب. هذه الحرارة لا تجد طريقاً سهلاً للخروج لأن الأنبوب النانوي معزول حرارياً بطبقة العازل المحيطة به. هذه الحرارة المتراكمة ترفع درجة حرارة الترانزستور، مما يؤدي إلى: تضييق فجوة النطاق(E_g) وزيادة تيار التسريب(I_{OFF}) وتدهور نسبة I_{ON}/I_{OFF} وكذلك احتمالية تلف الجهاز على المدى الطويل.

تجدر الإشارة إلى أن الأقطار الصغيرة للأنبوب النانوي تعني مقاومة أعلى لأن I_{ON} أصغر نسبياً، وهي تكون مساحة أقل لتصريف الحرارة، وبالتالي تكون كثافة التيار أعلى (تيار لكل وحدة مساحة).

يبين الجدول 11 تأثير التسخين الذاتي على أداء الترانزستورات للأنابيب المدروسة تحت ظروف تشغيل ($T_{ambient} = 300$ K. ،$L_{ch} = 10$ nm ، $V_{DS} = 1.0$ V)

الجدول 11. تأثير التسخين الذاتي على أداء الترانزستورات لأنابيب نانوية مختارة

Chiral Vector	Isolator	I_{ON} (μA)	P_{heat} (μW)	R_{th} (10^6K/W)	ΔT	T_{op} K	انخفاض Eg (meV)	I_{OFF} increase
(7,0)	SiO_2	3.05	3.05	10	30.5	330.5	14.3	× 3.2
(7,0)	HfO_2	3.12	3.12	8	25.0	325.0	11.8	× 2.7
(13,0)	SiO_2	5.22	5.22	5	26.1	326.1	12.3	× 2.8
(13,0)	HfO_2	5.35	5.35	4	21.4	321.4	10.1	× 2.4
(20,0)	SiO_2	7.53	7.53	3	22.6	322.6	10.6	× 2.5
(20,0)	HfO_2	7.71	7.71	2.5	19.3	319.3	9.1	× 2.2

نستنتج من الجدول 11 أن:

- الأنابيب فائقة الصغر تسخن أكثر، فالأنبوب (7,0) مع عازل □SiO يسخن أكثر (ΔT =30.5K) مقارنة بالأنبوب (20,0) مع □HfO (ΔT =19.3K)، وهذا يعود إلى أن المقاومة الحرارية أعلى بسبب صغر المساحة، وبالتالي تكون كثافة التيار أعلى.
- استخدام العازل □HfO يقلل ارتفاع درجة الحرارة بنسبة 15−20% مقارنة بالعازل □SiO والسبب أن العازل □HfO يوصل الحرارة بشكل أفضل من □SiO.
- انخفاض فجوة النطاق: بين 14-9 ميلي إلكترون فولت (تأثير بسيط)
- زيادة تيار التسريب: بين 2.2−3.2 مرة (تأثير ملحوظ لكن ليس قاتلاً للأداء)
- ارتفاع الحرارة يتراوح بين 19-30K أي أن الترانزستور يعمل بين 319-331K (46 إلى 58 درجة مئوية)، وهذه درجة حرارة مقبولة للأجهزة الإلكترونية.

على الرغم من التأثيرات الملحوظة للتسخين الذاتي، تبقى درجة حرارة التشغيل ضمن الحدود الآمنة للأجهزة الإلكترونية، مما يؤكد أن الترانزستورات المدروسة تحافظ على استقرار أدائها حتى تحت ضغط العمل المستمر.

9-5 نتائج كثافة الحالات (DOS) وتوزيع الشحنات

تعرف كثافة الحالات بأنها هي عدد "المسارات" أو "الحارات" المتاحة للإلكترونات للسير فيها عند كل سرعة (مستوى طاقة)، أي إذا كان عدد المسارات كبيراً يمر تيار أكبر وإذا إذا كان عدد المسارات قليلاً يمر تيار أقل.

أولاً: المستويات الفرعية للأنابيب المدروسة:

يبين الجدول 12 المستويات الفرعية (E□, E□, E□) للأنابيب النانوية المتعرجة المختلفة، حيث تحسب هذه المستويات بالعلاقة (15).

الجدول 12. المستويات الفرعية لأنابيب نانوية مختارة

Chiral Vector	Isolator	E_g (eV)	E_1 (eV)	E_2 (eV)	E_3 (eV)
(7,0)	SiO_2	1.533	0.766	1.533	2.299
(10,0)	SiO_2	1.073	0.536	1.073	1.609

(13,0)	SiO_2	0.825	0.412	0.825	1.237
(20,0)	SiO_2	0.537	0.268	0.537	0.805

ثانياً: توزيع الشحنات عند جهود بوابة مختلفة:

يبين الجدول 13 توزيع الإلكترونات في الأنبوب (13,0) عند جهود بوابة مختلفة، مما يوضح كيفية انتقال الترانزستور من حالة الإيقاف إلى حالة التشغيل.

الجدول 13. توزيع الشحنات للأنبوب (13,0) عند جهود بوابة مختلفة

V_{GS}	E_F(eV)	Level occupied	Number of electrons (nm^{-1})	Transistor state
0.0	0.000	Non	0	OFF
0.2	0.150	Non	0	OFF
0.4	0.300	Non	Very LOW	OFF
0.5	0.375	E1 only	3.2	ON
0.6	0.450	E1 only	3.8	ON
0.8	0.600	E1 and E2	4.0	Nearly Saturation
1.0	0.750	E1 and E2	4.0	Saturation

من الجدول 13 نلاحظ أن عند جهد التشغيل المستخدم(V_{GS} = 0.5 V) ، يكون مستوى فيرمي (E_F = 0.375 eV)قد تجاوز المستوى الفرعي الأول (E□ = 0.412 eV) مما يسمح بإشغاله بـ 3.2electron/nm، بينما لا يزال المستوى الفرعي الثاني (E□ = 0.825 eV) غير مشغول.

ثالثاً: تأثير كثافة الحالات على تيار التشغيل:

يبين الجدول 14 العلاقة بين قطر الأنبوب وعدد الإلكترونات وتيار التشغيل I_{ON} عند جهد V_{GS}=0.5V.

الجدول 14. العلاقة بين قطر الأنبوب وعدد الإلكترونات وتيار التشغيل I_{ON} عند جهد V_{GS}=0.5V

Chiral Vector	Diameter (nm)	E_g	Number of electrons (nm^{-1})	I_{ON}
(7,0)	0.556	1.533	2.1	3.05
(13,0)	1.032	0.825	3.2	5.22

(20,0)	1.588	0.537	3.9	7.53

تظهر النتائج أن الأنابيب ذات القطر الأكبر تمتلك فجوة نطاق أصغر، مما يسمح بدخول المزيد من الإلكترونات إلى القناة عند نفس جهد البوابة، وبالتالي تحقيق تيار تشغيل أكبر. هذا يؤكد العلاقة الطردية بين قطر الأنبوب وتيار التشغيل I_{ON}.

5-10نتائج حدود النموذج الباليستي (للقنوات الأطول)

كما تمت مناقشته في القسم 4–7، يفترض النموذج الباليستي غياب تأثيرات التشتت، وهو صالح عندما يكون طول القناة أقل من مسار الانتشار الحر ($MFP \approx 100$ nm) لتقييم مدى صلاحية هذا الافتراض لأطوال القناة الأكبر، يبين الجدول 15 تصحيح التيار للأنبوب (13,0) باستخدام العلاقة (15).

الجدول 15 تصحيح التيار لأطوال قناة مختلفة (للأنبوب 13,0) حيث أن MFP=100 nm و $I_{ballistic}$=5.22μA.

L_{ch} (nm)	L_{ch}/MFP	معامل التصحيح	I_{actual} (μA)	الانخفاض عن I_ballistic
10	0.10	0.91	4.75	9%
20	0.20	0.83	4.33	17%
30	0.30	0.77	4.02	23%
50	0.50	0.67	3.50	33%
100	1.00	0.50	2.61	50%

من الجدول 15 نستنتج أنه عندما يكون طول القناة من 10-20nm يصلح تطبيق النموذج الباليستي وعندما يكون طول القناة من 30-50nm عندها يكون النموذج بحاجة إلى إدخال التصحيح، وعندما يكون طول القناة أكبر من 50nm عندها يتطلب ذلك نموذجاً جديداً.

5-11 نتائج تأثير دالة الشغل للأقطاب

تؤثر دالة الشغل (Work Function) للمعدن المستخدم في أقطاب المصرف والمنبع بشكل كبير على أداء الترانزستور، حيث تحدد ارتفاع حاجز شوتكي (Schottky Barrier Height) عند واجهة التلامس بين المعدن والأنبوب النانوي.

يبين الجدول 16 تأثير اختيار معادن مختلفة على أداء الأنبوب (13,0) :

الجدول 16. تأثير اختلاف المعدن على الأداء (للأنبوب 13,0)

Metal	$\emptyset_{metal}$ (eV)	$\emptyset_B$ (eV)	Contact Type	I_{ON} (μA)	Contact Resistance (kΩ)	Drop from Ti
Al	4.1	-0.4	Ohmic	5.30	100	-1.5 better
Ti (used)	4.3	-0.2	Ohmic	5.22	80	0%
Cu	4.7	0.2	Low Schottky	4.80	250	8%
Au	5.1	0.6	Medium Schottky	4.20	500	20%
Ni	5.2	0.7	Medium Schottky	4.00	600	23%
Pd	5.6	1.1	High Schottky	3.20	1000	39%

من الجدول 16 يمكن استخلاص النتائج التالية:

- عندما تكون دالة شغل المعدن أقل من ألفة الإلكترون للأنبوب النانوي($\chi \approx 4.5$ eV)، يكون حاجز شوتكي سالباً ويتشكل تلامس أومي (Ohmic Contact) مثالي، كما في حالة الألومنيوم ($\emptyset_B = -0.4$eV) والتيتانيوم ($\emptyset_B = -0.2$eV) وهذا يؤدي إلى أعلى قيم لتيار التشغيل 5.30 μA و 5.22 μA على التوالي. أما عندما تزيد دالة شغل المعدن عن 4.5eV، فيصبح حاجز شوتكي موجباً ويتشكل تلامس شوتكي (Schottky Contact)، مما يقلل من تيار التشغيل بشكل ملحوظ، ففي حالة البلاديوم ($\emptyset_{metal} =$ 5.6eV)، يصل حاجز شوتكي إلى 1.1eV وينخفض I_{ON} إلى 3.20μA، أي بنسبة 39% مقارنة بالتيتانيوم.
- ترتبط مقاومة التلامس (Contact Resistance) ارتباطاً وثيقاً بارتفاع حاجز شوتكي. فكلما زاد الحاجز، زادت مقاومة التلامس، ففي حالة التلامس الأومي Ti و Al ، تتراوح

مقاومة التلامس بين 80-100kΩ، بينما في حالة التلامس الشوتكي العالي(Pd) ، تصل مقاومة التلامس إلى 1000kΩ.

- يعتبر التيتانيوم (Ti) والألومنيوم (Al) هما الأفضل للأنابيب النانوية الكربونية، حيث يشكلان تلامساً أومياً ويعطيان أعلى تيار تشغيل، بينما البلاديوم (Pd) والبلاتين (Pt) رغم شيوعهما في بعض الأبحاث، إلا أنهما يشكلان تلامساً شوتكياً عالياً ويخفضان I_{ON}بنسبة تصل إلى 40%.

5-12منهجية التكامل الهجين مع تقنيات CMOS الحالية:

بدلاً من طرح ترانزستورات CNTFET كبديل كامل لتقنيات السيليكون الحالية، وهو أمر غير عملي بسبب الاستثمارات الضخمة في صناعةCMOS ، نوصي بمنهجية التكامل الهجين (Hybrid Integration)حيث تدمج هذه الترانزستورات في التطبيقات التي تظهر فيها ميزة تنافسية واضحة، مثل وحدات الذاكرة SRAM وأجهزة إنترنت الأشياء .(IoT) ففي تطبيق SRAM، تعتبر خلايا الذاكرة مرشحاً واعداً لأنها تقضي معظم وقتها في حالة الإيقاف، مما يجعل تيار التسريب المنخفض ميزة حاسمة، حيث تحقق ترانزستورات CNTFET مثل الأنبوب (13,0) مع عازل HfO□ قيم I_{OFF} أقل بثلاث مرات وسرعة تبديل أسرع بـ 500 مرة مقارنة بـ CMOSأما في تطبيقات IoT التي تعمل معظم الوقت في وضع السكون، فإن ترانزستورات CNTFET ذات الأقطار الصغيرة جداً مثل الأنبوب (7,0) مع عازل HfO□ تحقق I_{OFF} = 8.9×10□¹□ A، وهو تسريب منخفض جداً يطيل عمر البطارية. يمكن تنفيذ التكامل الهجين بتصنيع ترانزستورات CNTFET في الطبقات العلوية للشريحة، فوق ترانزستورات CMOS السيليكونية في الطبقات السفلية، مع ربطهما عبر فتحات تمرير معدنية (Vias)، وهو تصميم هرمي يمكن تنفيذه بتعديلات طفيفة على خطوط الإنتاج الحالية. نوصي بالبدء بتطبيقات SRAM منخفضة الطاقة في الأجهزة المحمولة وأجهزة إنترنت الأشياء كخطوة أولى للتسويق التجاري، حيث تظهر ترانزستورات CNTFET ميزة تنافسية واضحة.

5-13 تحليل حساسية الأداء لانحرافات التصنيع:

كما تمت مناقشته في القسم 4–9، تعتبر انحرافات التصنيع من التحديات الرئيسية التي تواجه الإنتاج الضخم للأجهزة النانوية. بناءً على المنهجية الموصوفة هناك، تم في هذا القسم تحليل حساسية أداء الأنبوب (13,0) (d = 1.032 nm) لانحرافات القطر، باستخدام البيانات المحاكاة الواردة في الجداول 2 و 3. تم حساب متوسط التغير في I_{ON} و I_{OFF} لكل تغير بمقدار 0.1nm في قطر الأنبوب النانوي، وذلك في ثلاث مناطق مختلفة من نطاق الأقطار المدروس.

يوضح الجدول 17 نتائج هذا التحليل.

الجدول 17. متوسط التغير في I_{ON} و I_{OFF} لكل تغير بمقدار 0.1nm في قطر الأنبوب النانوي

Region	Diameter Range (nm)	Average ΔI_{ON} (μA/0.1nm)	Average ΔI_{OFF} (order of magnitude/0.1nm)
Small Diameters	0.556 - 0.794	0.39	≈ 3.5
Medium Diameters	0.794 - 1.032	0.32	≈ 2.0
Large Diameters	1.032 - 1.588	0.15	≈ 0.8

نستنتج من الجدول 17 مايلي:

- تبلغ حساسية I_{OFF} للتغيرات في القطر حوالي 3.5 مرتبة قياسية لكل 0.1 nm في منطقة الأقطار الصغيرة، وتنخفض إلى 0.8 مرتبة قياسية لكل 0.1 nm في منطقة الأقطار الكبيرة. هذا يعود إلى أن ميل العلاقة بين فجوة النطاق والقطر (dE_g/dd) يكون أكبر في الأقطار الصغيرة.
- تتغير I_{ON} بشكل شبه خطي مع القطر، حيث يبلغ متوسط التغير 0.39 μA لكل 0.1 nmفي الأقطار الصغيرة، وينخفض إلى 0.15 μA لكل 0.1 nm في الأقطار الكبيرة.
- الأقطار الأكبر من 1.2 nm هي الأقل حساسية لانحرافات التصنيع، مما يجعلها مناسبة للتطبيقات التي تتطلب صموداً عالياً للأداء.

تحدد هذه النتائج متطلبات دقة تصنيع قطر الأنبوب النانوي كما يلي:

- الأقطار الصغيرة (< 0.8 nm) :حساسية عالية جداً (3.5 مرتبة قياسية لكل 0.1nm)، وتتطلب هامش تصنيع أقل من ±0.05nm ، مما يزيد التكلفة ويحد من الإنتاجية. مناسبة فقط للتطبيقات المتخصصة.
- الأقطار المتوسطة(1.2 nm - 1.0) :حساسية متوسطة (2.0 مرتبة قياسية لكل 0.1nm) ، ويمكنها تحمل هامش ±0.1nm ، مناسبة للإنتاج التجاري المحدود.
- الأقطار الكبيرة (> 1.2 nm) : حساسية منخفضة (0.8 مرتبة قياسية لكل 0.1nm) ،وتتحمل هامش ±0.2nm ، مما يقلل التكلفة ويزيد الإنتاجية، ومناسبة للإنتاج الضخم (مثل SRAM والمعالجات).
- وبالتالي تكون الأقطار الأكبر من 1.2 nm هي الأقل حساسية لانحرافات التصنيع، مما يجعلها الخيار الأمثل للتطبيقات التي تتطلب صموداً عالياً للأداء وإنتاجاً بكميات كبيرة.

6- الاستنتاجات والتوصيات:

6-1 الاستنتاجات:

أثبتت هذه الدراسة أن هندسة الحصر الكمي من خلال ضبط قطر أنابيب الكربون النانوية المتعرجة تشكل منهج تصميم فعالة لتحسين أداء ترانزستورات CNTFET فقد أظهرت النتائج أن تقليل القطر إلى أقل من 1 nm يعزز فجوة النطاق بشكل كبير (حتى 1.533eV)، مما يؤدي إلى كبح أُسي لتيار التسريب (I_{OFF}) يتجاوز ثمانية مراتب قياسية، محققاً نسبة I_{ON}/I_{OFF} استثنائية تتجاوز 10^{12} للأنابيب فائقة الصغر.

كما تم إدراج مقاومة التلامس (100 kΩ) والسعات الطفيلية في النموذج، وتم حساب منتج القدرة–التأخير .(PDP) أظهرت النتائج أن الأنابيب فائقة الصغر (7,0) مع عازل □SiO تحقق أقل PDP (1.17 aJ)، بينما استخدام عازل □HfO يقلل I_{OFF} بمقدار 4.4 مرات، ولكنه يزيد PDP بمقدار 6.4 مرات. فيما يخص الأداء الديناميكي، تتراوح سرعة التبديل بين τ = 0.27-2.40 ps وتردد القطع بين f_T = 66-595 GHz ، مما يجعل هذه الأجهزة أسرع بـ 10-100 مرة من ترانزستورات السيليكون التقليدية. كما تم نمذجة التسخين الذاتي، حيث ارتفعت درجة حرارة التشغيل بمقدار 19-30K، وهي درجة حرارة مقبولة للأجهزة الإلكترونية.

تم تحديد النطاق القطري 1.0-1.2nm كمنطقة تصميم مثلى للدوائر الرقمية، حيث يحقق توازناً بين تيار تشغيل مناسب (5-7μA)، ونسبة I_{ON}/I_{OFF} تتراوح بين $10^6 - 10^9$. كما تم تحليل حساسية الأداء لانحرافات التصنيع، حيث تبين أن الأقطار الأكبر من 1.2nmهي الأقل حساسية، مما يجعلها مناسبة للإنتاج الضخم.

يربط هذا البحث بين ميكانيكا الكم الأساسية وهندسة الأجهزة العملية، مقدماً إطاراً كمياً لتحسين مفاضلة الأداء–الطاقة. كما تم اقتراح منهجية للتكامل الهجين مع تقنيات CMOS الحالية في تطبيقات SRAM وأجهزةIoT ، مما يعزز واقعية البحث وقابليته للتطبيق التجاري. ينقل هذا العمل

النموذج الفكري من اعتبار الحصر الكمي قيداً فيزيائياً إلى توظيفه كمتغير تصميم، مفتتحاً آفاقاً جديدة للحوسبة الموفرة للطاقة.

6-2 التوصيات:

على الرغم من الإطار النظري القوي الذي تقدمه هذه الدراسة، إلا أن هناك بعض التحديات التي تحد من التطبيق العملي المباشر لنتائجها. أبرز هذه التحديات يتمثل في صعوبة التصنيع الدقيق واسع النطاق لأنابيب الكربون النانوية المتعرجة ذات الأقطار المضبوطة، بالإضافة إلى تأثيرات السطح البيني ومقاومة التلامس في الأجهزة المصغرة.

بناءً على ذلك، يمكن اقتراح عدة اتجاهات للبحث المستقبلي:

- يوصى بتصنيع نماذج أولية من الترانزستورات المقترحة بأقطار مختارة (0.556nm و 1.032nm و 1.588nm) لمقارنتها مع نتائج المحاكاة الحالية وقياس خصائصها الكهربائية.
- يوصى بدراسة الأنماط الحلزونية(Chiral) و (Armchair) في أعمال مستقبلية، حيث قد توفر هذه الأنماط فجوات نطاق مختلفة أو خصائص معدنية مفيدة للتوصيلات داخل الدائرة الواحدة.
- : دراسة هياكل تجمع أقطاراً مختلفة لتحقيق متطلبات تطبيقية متنوعة.
- تطبيق مبادئ التصميم الهجين المقترحة على مستوى الدوائر المتكاملة منخفضة الطاقة، خاصة في وحدات SRAM وأجهزة IoT .
- دراسة عوازل High-k أخرى مثل(□ZrSiO ، □HfSiO، □ZrO) وسمكات مختلفة لتحقيق أفضل توازن بين الأداء والتسريب.
- محاكاة مونت كارلو: يوصى بإجراء محاكاة مونت كارلو شاملة في الدراسات المستقبلية لتقييم صمود الأداء بشكل كامل تحت تأثير انحرافات تصنيع متعددة.

المراجع العلمية:

[1] Park, J., Rosenblatt, S., Yaish, Y., et al. (2004). Electron–phonon scattering in metallic single-walled carbon nanotubes. *Nano Letters*, *4*, 517–520. https://doi.org/10.1021/nl035258c
[2] Wind, S. J., Appenzeller, J., & Avouris, P. (2003). Lateral scaling in carbon nanotube field-effect transistors. *Physical Review Letters*, *91*, 058301. https://doi.org/10.1103/PhysRevLett.91.058301
[3] Alvi, P. A., Lal, K. M., Siddiqui, M. J., & Naqvi, S. A. H. (2005). Carbon nanotubes field effect transistors: A review. *Indian Journal of Pure and Applied Physics*, *43*, 899–904.
[4] Jung, M., Schindele, J., Nau, S., et al. (2013). Ultraclean single, double, and triple carbon nanotube quantum dots with recessed Re bottom gates. *Nano Letters*, *13*, 4522–4526. https://doi.org/10.1021/nl402584p
[5] Schmid, D. R. (2014). *Suspended carbon nanotubes as electronical and nano-electro-mechanical hybrid systems in the quantum limit* (Doctoral dissertation). University of Regensburg.
[6] Bányai, L., & Koch, S. W. (1993). *Semiconductor quantum dots* (World Scientific Series on Atomic, Molecular and Optical Physics, Vol. 2). World Scientific.
[7] Brus, L. E. (1984). Electron–electron and electron–hole interactions in small semiconductor crystallites: The size dependence of the lowest excited electronic state. *The Journal of Chemical Physics*, *80*, 4403–4409. https://doi.org/10.1063/1.447218
[8] Verma, A., Gupta, H., & Chaturvedi, N. (2014). Analysis of the operational characteristics of CNTFET. *International Journal of Enhanced Research in Science Technology & Engineering*, *3*(3), 168–173.
[9] Singh, A. K. (2018). An analytical analysis of quantum capacitance in nano-scale single-wall carbon nano tube field effect transistor (CNTFET). *International Journal of Nanoelectronics and Materials*, *11*(3), 249–262

[10] Karamitaheri, H. (2013). *Thermal and thermoelectric properties of nanostructures* (Doctoral dissertation). Vienna University of Technology.

[11] Deng, J., & Wong, H.-S. P. (2007). A compact SPICE model for carbon-nanotube field-effect transistors including nonidealities and its application—Part II: Full device model and circuit performance benchmarking. *IEEE Transactions on Electron Devices*, *54*(12), 3195–3205. https://doi.org/10.1109/TED.2007.909043

[12] بشور، خ. (2021)، *تحسين أداء ترانزستورات الأثر الحقلي باستخدام الأنابيب النانوية لتطبيقات مختلفة* (أطروحة دكتوراه). جامعة حمص، كلية الهندسة الميكانيكية والكهربائية، قسم هندسة الالكترونيات والاتصالات.

[13] Lee, C.-S., & Wong, H.-S. P. (2015). Virtual-Source Carbon Nanotube Field-Effect Transistors Model Version 1.0.1: Technical User's Manual. *Stanford University, Dept. of Electrical Engineering*.

[14] Saito, R., Dresselhaus, G., & Dresselhaus, M. S. (1998). Physical properties of carbon nanotubes. *Imperial College Press*.

[15] Uyemura, J. (1988). Fundamentals of MOS digital integrated circuits. *Addison-Wesley*.

[16] John, D. L., & Pulfrey, D. L. (2006). Switching-speed calculations for Schottky-barrier carbon nanotube FETs. *Department of Electrical and Computer Engineering, University of British Columbia*.

[17] Pacheco-Sanchez, A., Bejenari, I., & Schröter, M. (2019). Self-heating characterization and thermal resistance modeling in multitube CNTFETs. *IEEE Transactions on Electron Devices, 66*(11), 4566-4571.

[18] Haroon, & Ahsan, M. A. H. (2023). Effect of van-Hove singularities in single-walled carbon nanotube leads on transport through T-shaped double quantum dot system. *Current Applied Physics*, 56, 100-110.

[19] Guo, J., & Lundstrom, M. (2005). Role of phonon scattering in carbon nanotube field-effect transistors. *Applied Physics Letters, 86*(19), 193103.

[20] Pourfath, M. (2007). *Numerical study of quantum transport in carbon nanotube-based transistors* [Doctoral dissertation, Technische Universität Wien].

[21] Burden, R. L., & Faires, J. D. (2011). *Numerical analysis* (9th ed.). Brooks/Cole, Cengage Learning.

[22] Ameen, S., Farhan, S., Monsur, F., & Nayan, F. (2024). Numerical analysis of diameter dependency of control coefficient of carbon nanotube field effect transistor. *Majlesi Journal of Electrical Engineering*, *18*(2), 1–7. Article 182435. https://dx.doi.org/10.57647/j.mjee.2024.1802.37

[23] Bachour, K., Ali, M., & Alabboud, I. (2020). Impact of changing channel length and band gaps of a carbon nanotube on the current of a Carbon Nanotube Field Effect Transistors (CNTFETs). *NJES*, *23*, 21–29. https://doi.org/10.29194/NJES.23010021

[24] Desgreys, P., Gomes Da Silva, J., & Robert, D. (2006). Dispersion impact on ballistic CNTFET n+-i-n+ performances. In *Proceedings of the European Nano Systems Workshop (ENS 2006)*. Paris, France.

[25] Sakib, M. N., Haque, A., & Nabi, M. (2016). Temperature effect of CNTFET under different dielectric materials. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, *4*(1), 1–4.

[26] Gupta, B. K., & Mishra, M. (2018). Comparative study of Ion/Ioff ratio of silicon nanowire transistor (SNWT) with planner MOSFET. *Bulletin of Pure and Applied Sciences Section D – Physics*, *37*, 1–6. https://doi.org/10.5958/2320-3218.2018.00001.5

[27] Capaz, R. B., Spataru, C. D., Tangney, P., Cohen, M. L., & Louie, S. G. (2008). Temperature dependence of the band gap of semiconducting carbon nanotubes. *Physical Review Letters*, *100*(4), 046801. https://doi.org/10.1103/PhysRevLett.100.046801

تقييم أداء شيفرة HAPPY/Steane الكمومية الهولوغرافية تحت أخطاء باولي النقية (X,Z,Y) وخطأ إزالة الاستقطاب (Depolarizing Error)

طالب الماجستير: مجد واصل سالم
جامعة حمص – كلية الهندسة الميكانيكية والكهربائية –قسم هندسة الالكترونيات والاتصالات
إشراف الدكتور: اسبر نسيم ابراهيم

الملخص

تعد الشيفرات الكمومية الهولوغرافية من أبرز المقاربات الحديثة لربط مفاهيم تصحيح الأخطاء الكمومية بمبدأ الهولوغرافيا ومراسلات الحافة والحجم (bulk-boundary) بوصفها نموذجاً بسيطاً وقابلاً للتحليل، من بين هذه الشيفرات، تبرز شيفرة Happy مبنية باستخدام تنسورات مثالية، مرتبة على شكل هرمي وقادرة على تمثيل الترميز الكمومي الهولوغرافي بطريقة هندسية واضحة. في هذا العمل، نقوم ببناء شيفرة HAPPY/Steane بمعدل صفري، وهي عبارة عن شيفرة يتم فيها دمج شيفرتين وهما شيفرة Happy وشيفرة Steane وتقييم أداءها تحت أخطاء باولي النقية (X,Y,Z) وخطأ إزالة الاستقطاب وذلك ضمن نموذج ضجيج مستقل على الكيوبتات الفيزيائية. نعتمد في تحليلنا علىInteger Optimisation Decoder. نقوم بدراسة معدلات الاسترجاع

بوصفها دالة في احتمال الخطأ الفيزيائي، مع مقارنة سلوك الشيفرة تحت أخطاء الانقلاب البتّي (X) وأخطاء الانقلاب الطوري (Z) وأخطاء من النوع (Y) وخطأ إزالة الاستقطاب (Depolarizing Error). تظهر النتائج ان شيفرة HAPPY/Steane تمتلك عتبة تصحيح تصل تقريباً الى 15% لكل من الأنواع الأربعة من الأخطاء مع اختلاف بمعدلات الاسترجاع لكل نوع من الأخطاء. كما تظهر النتائج ان معدلات الاسترجاع تزداد بزيادة عدد طبقات الشيفرة (R). تمت عملية المحاكاة باستخدام البيئة البرمجية Google Colabوتم من خلالها دراسة أداء الشيفرة.

كلمات مفتاحية: شيفرة Happy، شيفرة Steane، أخطاء باولي، خطأ إزالة الاستقطاب، Integer Optimisation Decoder

Performance Evaluation of the Holographic Quantum Happy/Steane Code under Pure Pauli Errors (X, Z, Y) and Depolarizing Error

Abstract

Holographic quantum codes represent one of the most prominent modern approaches for connecting the concepts of quantum error correction with the holographic principle and the bulk–boundary correspondence, offering a simple and analytically tractable model. Among these codes, the Happy code stands out as it is constructed using perfect tensors arranged in a hierarchical structure, enabling a clear geometric realization of holographic quantum encoding.In this work, we construct a zero-rate Happy/Steane code by integrating the Happy and Steane codes, and we evaluate its performance under pure Pauli errors of types (X, Y, Z) as well as depolarizing noise within an independent noise model acting on the physical qubits. Our analysis employs an Integer Optimisation Decoder. We study the logical recovery rates as a function of the physical error probability and compare the code's behavior under bit-flip (X), phase-flip (Z), Y-type Pauli errors, and depolarizing errors.The results show that the Happy/Steane code exhibits an error-correction threshold of approximately 15% for all four error types, with variations in recovery performance depending on the specific error model. Furthermore, the recovery rates improve as the number of code layers (R) increases. Simulations were performed using the Google Colab computational environment to evaluate the code's performance.

Keywords: Integer Optimisation Decoder, Happy Code Steane Code, Pauli

مقدمة

يُعد تصحيح الأخطاء الكمومية حجر الأساس لتحقيق الحوسبة الكمومية واسعة النطاق، نظرًا للهشاشة الشديدة للحالات الكمومية أمام الضجيج والتفاعلات البيئية. وقد طُوّرت خلال

العقود الماضية عائلات متعددة من الشيفرات الكمومية، مثل شيفرات السطح Surface) (Codesوالشيفرات التوافقية(Concatenated Codes) ، والتي أثبتت كفاءة عالية في نماذج ضجيج معينة، ولكن تتطلب مخططات تصحيح الأخطاء الكمومية لهذه الشيفرات آلاف الكيوبتات الفيزيائية لكل كيوبت منطقي. على سبيل المثال، تعد الشفرات السطحية (surface codes)من أكثر الأساليب الواعدة، حيث تتطلب($O(d^2)$ كيوبتات فيزيائية لمسافة الشيفرة d، مع عبء يتوسع بمقدار $O(n)$ كيوبت فيزيائي لكل كيوبت منطقي لتحقيق معدلات خطأ منطقية بمقدار ($O(e^{-a\sqrt{n}})$. ومع ذلك، على الرغم من التقدم الكبير في السنوات الأخيرة، بما في ذلك العروض التجريبية لتحقيق ميزة حسابية كمومية، لا تزال العديد من التحديات الأساسية تعيق تطوير حواسيب كمومية عملية. تدفع هذه القيود المستمرة إلى استكشاف نموذج الحوسبة الكمومية الهولوغرافية، وهو منظور جديد يهدف إلى مواجهة هذه التحديات من خلال الاستفادة من الرؤى المستقاة من المبدأ الهولوغرافي ومطابقة ADS/CFT. يقترح المبدأ الهولوغرافي، الذي اقترحه هوفْت وساسكيند أن محتوى المعلومات في منطقة من الفضاء يمكن وصفه بواسطة نظرية تعمل على حدود هذه المنطقة. يتحدى هذا المبدأ فهمنا التقليدي لكيفية تخزين المعلومات في الأنظمة الفيزيائية، ويقترح وجود صلة عميقة بين المعلومات وهندسة الزمكان. مراسلة (ADS/CFT)، التي اقترحها مالداسينا، توفر تحقيقًا ملموسًا لمبدأ الهولوغرافيا. إنها تفترض وجود ازدواجية (تكافؤ) بين نظرية الجاذبية في فضاء Anti-de Sitter ذي الأبعاد (d + 1) ونظرية المجال المتوافق (CFT)الموجودة على الحدود ذات الأبعاد(d) لهذا الفضاء. وقد أظهرت أعمال حديثة أن هذا المبدأ يمكن تفسيره بلغة تصحيح الأخطاء الكمومية، حيث يمكن النظر إلى الحجم (bulk)بوصفه معلومات منطقية مشفّرة بطريقة غير محلية في درجات حرية الحافة. أدى هذا الإدراك إلى ظهور فئة جديدة تُعرف بالشيفرات الكمومية الهولوغرافية. تُعد شيفرةHappy التي قُدمت لأول مرة بواسطةPastawski وآخرين، نموذجًا أوليًا لهذه الشيفرات، حيث تُبنى باستخدام تنسورات مثالية(Perfect Tensors) مرتبة في شبكة هرمية تمثل هندسة فراغية

سالبة الانحناء. إلى جانب شيفرةHappy، تم تطوير العديد من شيفرات تصحيح الأخطاء الكمومية الهولوغرافية الأخرى، حيث يُعد اختيار بُنى التنسورات عاملاً حاسماً في هذه التصاميم. ومن بين هذه البُنى شيفرة ستين الهولوغرافية، التي تعتمد على تنسور شيفرة ستين. وتُعد شيفرة ستين نفسها من شيفرات تصحيح الأخطاء الكمومية الراسخة والمعروفة بامتلاكها خصائص مميزة عديدة، مما يجعلها مرشحاً طبيعياً لبناء نسخة هولوغرافية منها. وبالنظر إلى متانتها وقدرتها على تصحيح الأخطاء بكفاءة، يمكن توقّع أن ترث شيفرة ستين الهولوغرافية مزايا مهمة تعزز من أدائها في تطبيقات الجاذبية الكمومية والمعلومات الكمومية. لقد تمت دراسة خصائص شيفرة HAPPY و شيفرة ستين وقدرتها على تصحيح أخطاء محو الكيوبتات بالإضافة الى أخطاء باولي النقية و Depolarizing Error كما تمت دراسة خصائص شيفرة Happy/Steane تحت أخطاء محو الكيوبتات، ولكن في دراستنا سنركز على بناء وتقييم أداء شيفرة Happy/Steane تحت أخطاء باولي النقية وخطأ إزالة الاستقطاب (Depolarizing Error).

تنظيم البحث:

تم تنظيم البحث وفق الترتيب التالي: تم البدء بمقدمة عن العمل المنجز ثم تم ذكر الهدف من البحث وطرائقه، وبعدها تم إجراء دراسة مرجعية في الفقرة 1، ثم تم في الفقرة2 شرح وتصميم الشيفرة وأخيراً تم مناقشة النتائج التي تم الحصول عليها في الفقرة3.

هدف البحث:

يهدف البحث الى بناء وتقييم أداء شيفرة Happy/Steane الكمومية الهولوغرافية واختبار قدرتها على تصحيح أخطاء باولي النقية وهي أخطاء الانقلاب البتّي (X) وأخطاء الانقلاب الطوري (Z) وأخطاء باولي من النوع (Y) وخطأ إزالة الاستقطاب.

مواد وطرق البحث:

يعتمد البحث على الطرق التجريبية والتسنيد المرجعي وتمت المحاكاة باستخدام برنامج Google Colab لبناء واختبار الشيفرة وذلك على حاسب بمواصفات RAM 12 GB CPU 2.00 GH ونظام تشغيل Windows10.

1 دراسة مرجعية

تمت الدراسات الشيفرات الهولوغرافية في عدد من الأبحاث، حيث قام (2015, Fernando Pastawski et al) في [9] بتقديم عائلة من النماذج الاختبارية القابلة للحل بدقة تامة لتمثيل مراسلة (ADS/CFT)، وذلك بالاستناد إلى بناء مبتكر لشفرة تصحيح أخطاء كمومية ذات بنية قائمة على شبكات الموترات. حيث وجدو أن شيفرة البنتاغون تفشل في تحقيق شرط امتلاك عتبة محو، ولكن قاموا بتعديل الشيفرة عن طريق وضع بنتاغون في المركز، ثم إضافة طبقات متناوبة من سداسيات من دون أدلة bulk متدلّية وبنتاغونات لكل منها دليل bulk واحد ووجدوا أن المحو يمكن تصحيحه بواسطة الخوارزمية الشجعة greedy) algorithm) باحتمال نجاح عالٍ عندما يكون $p \leq p_c^{greedy} \approx 0.26$. وتمت دراسة الشيفرة المعدلة عن طريق وضع دليل منطقي واحد في المركز ووجدوا أن عتبة المحو قريبة جدًا من $p_c=1/2$ وقد تم التحليل باستخدام greedy algorithm. وفي [6] قام (2020,Robert J. Harris et al) بدراسة شيفرةHappy بمعدل اعظمي ومعدل صفري وشيفرة ستين وشيفرة SCF تحت قناة إزالة الاستقطاب وقد تمت الدراسة باستخدام Integer Optimization Decode ووجدوا ان كل من شيفرة Happy وSteane بمعدل صفري تملك عتبة تصل الى 15%. وفي الدراسة [4] قام (2024, J. Fan et al) بدراسة مجموعة من الشيفرات الهولوغرافية مثل شيفرةSteane, SCF HAPPY و[6, 1, 3] ووجدوا أنها

تملك عتبات مختلفة تحت أخطاء باولي المختلفة وتمت الدراسة باستخدام tensor-network decoder ويمكن تلخيص النتائج في الجدول التالي.

الجدول1: نقاط بيانات عتبة الاسترجاع p_{th} لعددٍ من حالات الضجيج النقي (pure noise)

وضجيج باولي المنحاز (2-Pauli biases)

Noise Bias (r_X, r_Y, r_Z)	Hashing (%)	HaPPY (%)	Steane (%)	SCF (%)	$[\![6,1,3]\!]$ (%)
Depolarizing $(1/3, 1/3, 1/3)$	18.929	17.9 ± 0.81	18.98 ± 0.36	18.83 ± 0.13	18.46 ± 0.36
Pure X $(1, 0, 0)$	50.00	49.92 ± 0.096	10.79 ± 0.18	11.42 ± 0.52	33.56 ± 1.08
Pure Y $(0, 1, 0)$	50.00	49.91 ± 0.11	10.8 ± 0.19	44.66 ± 2.82	33.56 ± 1.08
Pure Z $(0, 0, 1)$	50.00	49.91 ± 0.11	10.8 ± 0.19	11.42 ± 0.52	22.99 ± 1.39
Pure XZ $(1/2, 0, 1/2)$	22.709	21.45 ± 0.68	22.1 ± 0.16	24.027 ± 1.08	22.08 ± 0.04
Pure XY $(1/2, 1/2, 0)$	22.709	21.45 ± 0.69	22.1 ± 0.17	21.77 ± 0.51	22.58 ± 0.17
Pure YZ $(0, 1/2, 1/2)$	22.709	21.45 ± 0.68	22.1 ± 0.17	21.77 ± 0.51	22.08 ± 0.04

وفي [10] قام (2025,M. Steinberg et al) بدراسة كل من شيفرات QRMو Steane و Happy.أيضاً تمت دراسة شيفرة (Happy Plus Steane) وذلك تحت قناة المحو ووجدوا أن العتبة العددية لكل من (Steane, Happy)هي 50% اما بالنسبة لشيفرة QRM قيمتها 19.2% أما عتبة Happy Plus Steane تصل الى 50%.

2 الشيفرة الكمومية الهولوغرافية [Holographic Quantum Code]

تُعرف الشيفرة الكمومية الهولوغرافية عبر خريطة ترميز ايزومترية

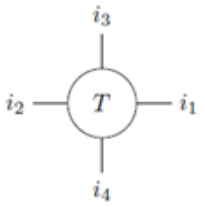
$$V: H_{bulk} \rightarrow H_{boundary} \quad (1)$$

التي تقوم بربط فضاء هيلبرت في الحجم (bulk) بفضاء هيلبرت على الحد.

يتم بناء الشيفرات الكمومية الهولوغرافية باستخدام شبكات التنسور التي تحاكي هندسة فضاء ADS. المكونات الأساسية لهذا البناء هي التنسورات المثالية (Perfect Tensors) والتبليطات الزائدية (Hyperbolic Tessellations).

2.1 التنسور [Tensor]

التنسور (Tensor): هو كائن رياضي ذو عدة أدلة (indices)، يمثل كل دليل درجة حرية محددة.



شكل1: التمثيل الرسومي للتنسور

في التمثيل الرسومي:

العقد (Nodes) تمثل التنسورات (tensors).

الحواف (Edges) تمثل عمليات الربط (Contractions) بين العقد، وهي عمليات جمع على الأدلة المشتركة بين التنسورات.

الرتبة (Rank): عدد الحواف المتصلة بالعقدة يمثل رتبة الموتر (عدد الأدلة). تمثل الحواف المفتوحة الأدلة الحرة.

2.2 التنسور المثالي [Perfect Tensor]

2.2.1 التنسور الايزومتري [Isometries Tensor]

إذا كانHA و HB فضاءان من فضاءات هيلبرت(Hilbert spaces) ، ليس بالضرورة أن يمتلكا البُعد نفسه. يُقال إن التحويل

$$T: HA \rightarrow HB \quad (2)$$

هو تحويل متساوي القياس (isometry) إذا كان تحويلًا خطيًا (linear map) يحافظ على الجداء الداخلي.

إذا كان HA و HB محدودي البُعد، فإن ذلك يعني مباشرةً أن مثل هذا التحويل T لا يمكن أن يوجد إلا إذا كان بُعداهما يحققان:

$$\dim(A) \leq \dim(B) \qquad (3)$$

وفي الحالة الخاصة التي يكون فيها:

$$\dim(A) = \dim(B) \qquad (4)$$

فإن T يكون ببساطة تحويلًا وحدوياً. ومن الواضح أن تركيب تحويلين متساويي القياس يعطي أيضًا تحويلًا متساوي القياس.

إذا كان

$$T: HA \rightarrow HB \qquad (5)$$

يمكننا تمثيل التحويل T على شكل تنسور ذي دليلين (two-index tensor) يعمل على الشكل التالي:

$$T: |a\rangle \longrightarrow \sum_b |b\rangle T_{ba} \qquad (6)$$

حيث أن المجموعة $\{|ai\rangle\}$ تمثّل أساسًا متعامدًا متكاملاً لفضاء HA، وتمثّل المجموعة $\{|bi\rangle\}$ أساسًا متعامدًا متكاملاً لفضاء HB عندئذٍ يكون T تحويلًا متساوي القياس (isometry) إذا وفقط إذا تحقق الشرط:

$$\sum T^{\dagger}_{a'b} T_{ba} = \delta a' a \qquad (7)$$

حيث δ هي دلتا كرونيكر و $\dagger$ هو المنقول.

2.2.2 التنسور المثالي

يكون التنسور ذو 2n فهرسًا

Ta1a2…a2n

تنسور مثالي (perfect tensor) إذا كان، لأي تقسيم ثنائي (bipartition) لأدلته إلى مجموعة A ومجموعة مكمّلة A^c بحيث

$$|A| \leq |A^c|$$

فإنT يكون متناسبًا مع تنسور ايزومتري (isometric tensor) من A إلى A^c.

2.3 التبليطات الزائدية(Hyperbolic Tilings)

سيُعطى نموذج الشيفرة الكمومية لتصحيح الأخطاء لدينا بواسطة شبكة تنسورات (tensor network) نضعها داخل قرص بوانكاريه (Poincaré disc) وسنفعل ذلك من خلال اختيار تبليط زائدي منتظم للمستوي.

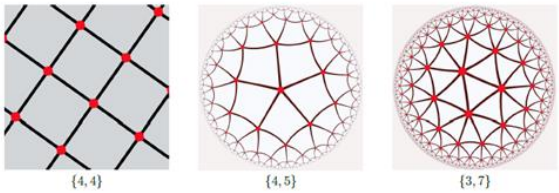
يُعرَّف التبليط الزائدي المنتظم على أنّه تغطية (covering) للمستوي الزائدي (hyperbolic plane) بواسطة مضلعات منتظمة (regular polygons) بحيث لا تتقاطع فيما بينها ولا تترك فجوات. إحدى الطرق الممكنة لوصف هذه التبليطات أو الفسيفساء (tessellations) هي باستخدام رمز شلافلي

{p;q}، حيث:

- p : عدد الحواف (أو الأضلاع) التي تلتقي عند كل رأس.
- q : عدد الحواف لكل مضلع في التبليط.

يكون التبليط المنتظم ثنائي ذي رمز شلافلي زائدياً(hyperbolic) إذا وفقط إذا كان:

$$q > \frac{2p}{p-2} \qquad (8)$$



شكل 2: بعض التبليطات المنتظمة ثنائية الأبعاد

مع رمز شلافلي (Schläfli symbols). لاحظ أنّ الأول منها ينتمي إلى الهندسة الإقليدية(Euclidean) ، بينما الاثنان الآخران ينتميان إلى الهندسة الزائدية (hyperbolic).

2.4 شيفرة Happy

يتم بناء شيفرة Happy على النحو التالي:

1. البناء الأساسي:(Base Construction)

- يبدأ البناء بشيفرة مثالية[[5, 1, 3]] ، تقوم بترميز كيوبت منطقي واحد في خمسة كيوبتات فيزيائية، مع مسافة شيفرة d=3.
- تتكون مجموعة المثبّتات (Stabilizer Group) لهذه الشيفرة المثالية من 4 مولدات، يُرمز إليها بـ ⟨S1,S2,S3,S4⟩=S والمولدات هي كما يلي:

$$S_1 = X \otimes Z \otimes Z \otimes X \otimes I$$
$$S_2 = I \otimes X \otimes Z \otimes Z \otimes X$$
$$S_1 = X \otimes I \otimes X \otimes Z \otimes Z$$
$$S_1 = Z \otimes X \otimes I \otimes X \otimes Z$$

شكل 3: مجموعة المثبّتات لشيفرةHappy حيث تشير ⊗ الى الضرب التنسوري.

- المؤثرات المنطقيةX_L وZ_L لهذه الشيفرة المثالية هي:

$$X_L = X \otimes X \otimes X \otimes X \otimes I$$
$$Z_L = Z \otimes Z \otimes Z \otimes Z \otimes Z$$

شكل 4: المؤثرات المنطقية لشيفرة Happy

2.تغطية قرص بوانكاريه:(Tiling the Poincaré Disk)

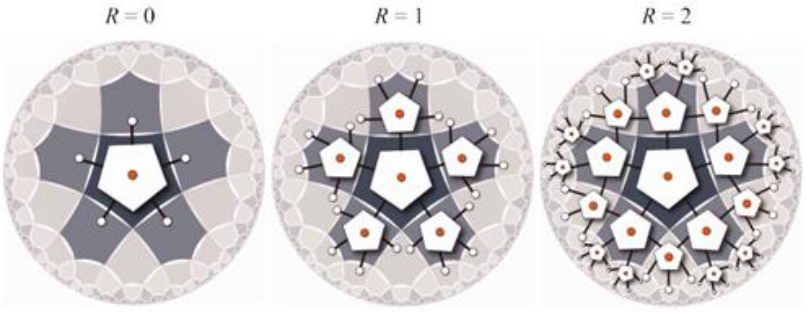
- تُوضع شيفرة مثالية [[5, 1, 3]] في مركز قرص بوانكاريه ذو انحناء سلبي.
- يتم بعد ذلك بناء شيفرةHappy باستخدام طريقة التجزئة

[5,4] = [p,q] مع استخدام تضخيم الحواف لتوسيع الشيفرة من R=0إلى R=2.

i. R=0 : توضع الشيفرة المثالية [[5, 1, 3]] في مركز قرص بوانكاريه.

ii. R=1 : تُكدّس طبقة جديدة من شيفرات [[5, 1, 3]] المثالية هولوغرافيًا فوق طبقة R=0، وفق قواعد تضخيم الحواف(Edge Inflation) .

iii. R=2 : تُكدّس طبقة أخرى، مما يوسّع شيفرة Happy بشكل أكبر.



شكل 5: عملية النمو لشيفرة Happy من R = 0 إلى R = 2، باتباع قواعد تضخم الحواف. ضع الشيفرة المثالية [[5، 1، 3]] على قرص بوانكاريه ذي انحناء سالب. هذا يشكل شيفرة Happy بنصف قطر R = 0. ثم، باستخدام وضع تضخم الحواف، قم بتكديس هولوغرافي لطبقة جديدة من الشيفرة المثالية

[[5، 1، 3]] على شيفرة Happy بنصف قطر R = 0 مع [p,q] = [5,4] مما ينتج عنه شيفرة Happy بنصف قطر R = 1. من خلال الاستمرار في التكديس الهولوغرافي لطبقة من الشيفرة المثالية [[5، 1، 3]] على شيفرة Happy بنصف قطر R = 1 مع [p,q] = [5,4] يمكن الحصول على شيفرة Happy بنصف قطر R = 2.

2.4.1 تضخيم الرؤوس (Vertex Inflation):

عند إضافة طبقة جديدة من التنسورات، يجب أن يشترك كل موتر في حافة مع موتر من الطبقة السابقة، كما يجب أن يتصل بالموترات الأخرى ضمن نفس الطبقة.

2.4.2 تضخيم الحواف (Edge Inflation):

بالمقابل، يتطلب تضخيم الحواف فقط أن تُضاف الموترات الجديدة عن طريق التمدد من حافة موتر في الطبقة السابقة. على عكس تضخيم الرؤوس، لا يوجد شرط صارم بأن تكون جميع الرؤوس مشتركة بين عدد محدد من المضلعات q.

2.5 معدل شيفرةHappy

يُعطى معدل الشيفرة الكمومية [[n, k, d]] بالعبارة k/n، حيث k هو عدد الكيوبتات المنطقية و nهو عدد الكيوبتات الفيزيائية و d مسافة الشيفرة. يمكن بناء شيفرة Happy بمعدلات مختلفة، اعتمادًا على كثافة ترميز الكيوبتات المنطقية داخل شبكة التنسورات.

2.6 شيفرة Steane

تستخدم شيفرة ستين الهولوغرافية تنسور شيفرة ستين. كنوع من شيفرات تصحيح الأخطاء الكمومية، تحتوي شيفرة ستين على العديد من الميزات البارزة، مما يعطينا سببًا لنتوقع أن شيفرة ستين الهولوغرافية ستحمل خصائص مرغوبة.

عند فحص تنسور شيفرة ستين، نأمل بشكل طبيعي أن يكون تنسوراً مثاليًا، مشابهًا للتنسور [[5، 1، 3]] المستخدم في شيفرة Happy، لكن هذا ليس الحال. ومع ذلك، الخبر السار هو أن التنسور المقابل لشيفرة ستين يمكن أن يشكل تطبيقاً إيزومترياً تحت بعض التقسيمات الثنائية المحدودة، والتي نشير إليها باسم تنسورات مثالية كتلية. حيث تُعرَّف التنسورات المثالية الكتلية بأنها التنسورات التي تحقق شرط الإيزومترية، ولكن ليس لجميع ترتيبات الأدلة، بل فقط بالنسبة لفئة محددة منها، وهي التبديلات الدورية لمجموعة الأدلة.

وبصورة أدق، يُقال إن التنسور يكون كتليًا مثاليًا إذا ظل يمثل تحويلًا متساوي القياس تحت جميع الإزاحات الدورية σ المطبَّقة على مجموعة الأدلة J . بحيث:

2.6.1 بناء شيفرة ستين الهولوغرافية

لبناء شيفرة ستين الهولوغرافية، دعونا أولاً ننظر إلى التنسور الأساسي الذي نستخدمه: تنسور ستين.

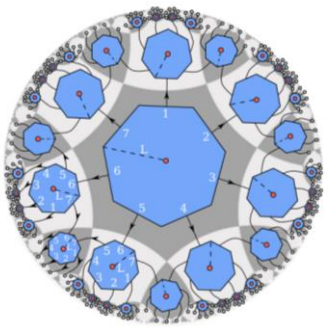
أدناه مولدات مجموعة المثبتات والمشغلات المنطقية لتنسور ستين:

Index label:	1	2	3	4	5	6	L	7
S_1	X	X	I	I	I	X	I	X
S_2	I	X	X	X	I	I	I	X
S_3	I	I	I	X	X	X	I	X
S_4	Z	Z	I	I	I	Z	I	Z
S_5	I	Z	Z	Z	I	I	I	Z
S_6	I	I	I	Z	Z	Z	I	Z
$S_{\bar{X}}$	X	X	X	X	X	X	X	X
$S_{\bar{Z}}$	Z	Z	Z	Z	Z	Z	Z	Z

شكل 6: مجموعة المثبتات والمشغلات المنطقية لتنسور ستين

تنسور ستين ليس تنسوراً مثاليًا لأنه لا يمكن أن يضمن تشكيل تطبيق إيزومتري تحت جميع التبديلات. ومع ذلك، طالما أن أدلة "المدخل" لتنسور ستين متجاورة تسلسليًا، على سبيل المثال، 7، L، 6، مع اعتبار الأدلة المتبقية ك ."مخرج"، يمكن التحقق من أن تنسور ستين سيوفر تطبيقاً إيزومترياً تحت هذه الشروط.

باستخدام تنسور ستين كتنسور بذري واستخدام طريقة تبليط [7,4] = [p,q] على قرص بوانكاريه، يمكن أن يشكل تبليط هولوغرافي لتنسورات ستين شيفرة ستين الهولوغرافية. كما هو موضح في الشكل.



شكل7: تبليط البذرة السباعية (heptagonal seed) ضمن تمثيل شبكة تنسورية أكبر ذات نصف قطر R=3 لشيفرة سباعية(heptagon code) .

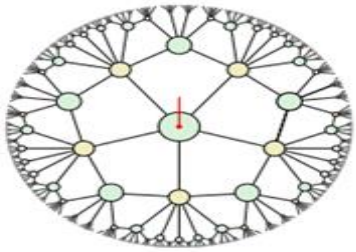
يمكن بناء شيفرة Steane بمعدلات مختلفة، اعتمادًا على كثافة ترميز الكيوبتات المنطقية داخل شبكة التنسورات.

2.7. شيفرة HAPPY/ Steane

شكل8:Happy+Steane

في هذه الشيفرة الهولوغرافية غير المتجانسة، قمنا بوضع شيفرة Happy في الطبقات الزوجية، وشيفرة ستين في الطبقات الفردية، مُشكِّلين بذلك شيفرة هولوغرافية غير متجانسة من النوعHappy+Steane. استخدام شيفرة هولوغرافية غير متجانسةHappy/Steane بدلًا من الشيفرة المتجانسة يؤدي الى مساحة أقل مقارنة بالشيفرة الهولوغرافية المتجانسةSteane، كما يتضح من طريقة بناء الشيفرة.

يمكن بناء شيفرة Happy+Steane بمعدلات مختلفة، اعتمادًا على كثافة ترميز الكيوبتات المنطقية داخل شبكة التنسورات وفي هذا البحث سنعمل على شيفرة Happy+Steane بمعدل صفري أي ترميز كيوبت منطقي واحد كما هو موضح بالشكل.



شكل9: شيفرة Happy/Steane حيث تتناوب طبقات الشبكة بين تنسوراتHappy وتنسوراتSteane، مع تمثيل تنسورات Steane في الشكل باللون الأصفر الفاتح لتمييزها بصريًا داخل البنية التنسورية.

3.النتائج

تمت محاكاة الشيفرة الموضحة بالشكل حيث تم، حيث تم بناء الشيفرة باستخدام لغة البرمجة بايثون وتم بناء طبقتين فقط من الشيفرة R0,R1 وتمت المحاكاة على البيئة البرمجية Google Colab وتم استخدام Decoder من نوع Integer Optimization Decode.

تم اختبار الشيفرة تحت أخطاء باولي النقية وخطأ إزالة الاستقطاب.

أخذنا مصفوفات باولي الشهيرة التالية، والتي تمثل عمليات الخطأ على الكيوبتات:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (9)$$

$$Bit\ Flip:\ X|0\rangle = |1\rangle, X|1\rangle = |0\rangle \quad (10)$$

$$Phase\ Flip:\ Z|0\rangle = +1|0\rangle, Z|1\rangle = -1|1\rangle \quad (11)$$

$$Y = X + Z \quad (12)$$

3.1 خطأ باولي النقي من النوع X.

حيث

$$X = 1 , Y = 0 , Z = 0$$

أي أن الكيوبتات تتعرض فقط لخطأ انقلاب البت.

يوضح الجدول التالي نتائج المحاكاة التي تم فيها اختبار الشيفرة تحت خطأ X بالنسبة للطبقة R0.

الجدول 2: احتمالية الخطأ مقابل معدل الاسترجاع لخطأX للطبقة R0

Recovery Rate	Probability
0.998	0.01
0.994	0.02
0.992	0.03
0.994	0.04
0.978	0.05
0.988	0.060000000000000005
0.944	0.06999999999999999
0.936	0.08
0.948	0.09
0.932	0.09999999999999999
0.896	0.11
0.906	0.12
0.832	0.13
0.868	0.14
0.86	0.15000000000000002
0.826	0.16
0.822	0.17
0.768	0.18000000000000002
0.75	0.19
0.75	0.2
0.674	0.21000000000000002

0.712	0.22
0.69	0.23
0.676	0.24000000000000002
0.634	0.25
0.63	0.26
0.536	0.27
0.596	0.28
0.564	0.29000000000000004
0.538	0.3

الأرقام الموضحة في العمود الأول تدل على احتمالية الخطأ اما العمود فهو معدل الاسترجاع المنطقي المقابل لكل احتمالية خطأ.

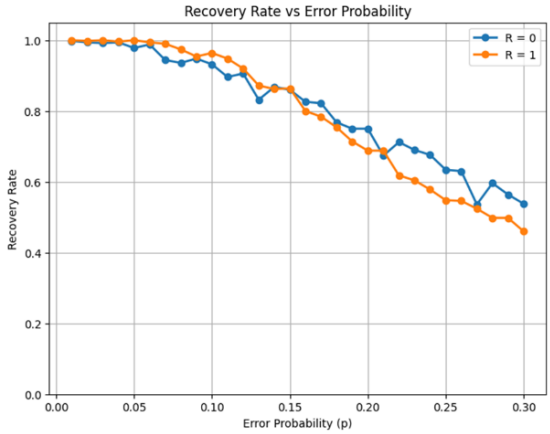
أما بالنسبة للطبقة R1 فالنتائج موضحة بالجدول التالي.

الجدول 3: احتمالية الخطأ مقابل معدل الاسترجاع لخطأX للطبقة R1

Recovery Rate	Probability
1.0	0.01
0.998	0.02
1.0	0.03
0.996	0.04
1.0	0.05
0.994	0.060000000000000005
0.99	0.06999999999999999
0.974	0.08

0.954	0.09
0.964	0.09999999999999999
0.948	0.11
0.92	0.12
0.872	0.13
0.862	0.14
0.864	0.15000000000000002
0.8	0.16
0.784	0.17
0.754	0.18000000000000002
0.714	0.19
0.688	0.2
0.688	0.21000000000000002
0.618	0.22
0.604	0.23
0.578	0.24000000000000002
0.548	0.25
0.546	0.26
0.524	0.27
0.498	0.28
0.498	0.29000000000000004
0.46	0.3

نقوم بتلخيص النتائج بالشكل التالي للطبقتين R0,R1.



شكل 10: منحنى معدل الاسترجاع (Recovery Rate Curve) للشيفرة Happy/Steane ذات المعدّل الصفري (Zero-Rate) تحت ضجيج باولي النقي من النوع X، مع عتبة تصحيح تقارب 15%~.

3.2 خطأ باولي النقي من النوع Y.

حيث

$$X = 0\,, Y = 1\,, Z = 0$$

يوضح الجدول التالي نتائج المحاكاة التي تم فيها اختبار الشيفرة تحت خطأ Y بالنسبة للطبقة R0.

الجدول 4: احتمالية الخطأ مقابل معدل الاسترجاع لخطأY للطبقة R0

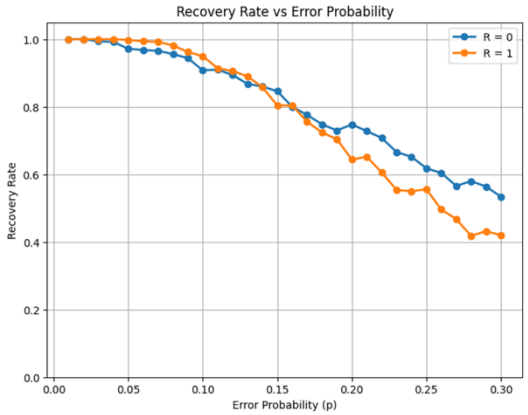
Recovery Rate	Probability
1.0	0.01
1.0	0.02
0.994	0.03
0.992	0.04
0.972	0.05
0.968	0.060000000000000005
0.966	0.06999999999999999
0.956	0.08
0.944	0.09
0.908	0.09999999999999999
0.91	0.11
0.894	0.12
0.868	0.13
0.86	0.14
0.846	0.15000000000000002
0.8	0.16
0.776	0.17
0.748	0.18000000000000002
0.73	0.19
0.748	0.2
0.728	0.21000000000000002
0.708	0.22
0.666	0.23
0.652	0.24000000000000002
0.618	0.25
0.604	0.26
0.566	0.27
0.58	0.28
0.564	0.29000000000000004
0.534	0.3

أما بالنسبة للطبقة R1 فالنتائج موضحة بالجدول التالي.

الجدول 5: احتمالية الخطأ مقابل معدل الاسترجاع لخطأY للطبقة R1

Recovery Rate	Probability
1.0	0.01
1.0	0.02
1.0	0.03
1.0	0.04
0.998	0.05
0.994	0.060000000000000005
0.992	0.06999999999999999
0.982	0.08
0.962	0.09
0.95	0.09999999999999999
0.914	0.11
0.906	0.12
0.89	0.13
0.858	0.14
0.804	0.15000000000000002
0.804	0.16
0.756	0.17
0.724	0.18000000000000002
0.704	0.19
0.644	0.2
0.652	0.21000000000000002
0.606	0.22
0.554	0.23
0.55	0.24000000000000002
0.556	0.25
0.496	0.26
0.468	0.27
0.418	0.28
0.432	0.29000000000000004

0.42	0.3



شكل 11: منحنى معدل الاسترجاع (Recovery Rate Curve)

للشيفرةHappy/Steane ذات المعدّل الصفري (Zero-Rate) تحت ضجيج باولي النقي من النوع Y، مع عتبة تصحيح تقارب 15%~.

3.3 خطأ باولي النقي من النوع Z.

حيث

$$X = 0\,, Y = 0\,, Z = 1$$

يوضح الجدول التالي نتائج المحاكاة التي تم فيها اختبار الشيفرة تحت خطأ Z بالنسبة للطبقة R0.

الجدول 6: احتمالية الخطأ مقابل معدل الاسترجاع لخطأZ للطبقة R0

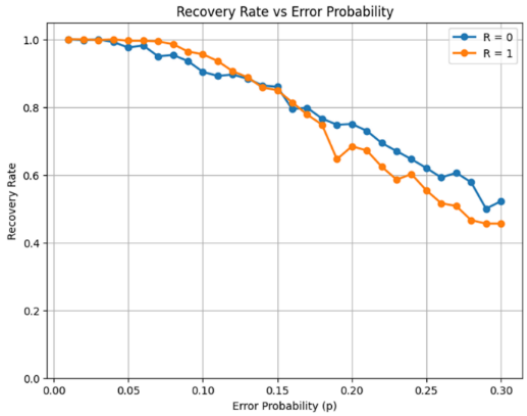
Recovery Rate	Probability
1.0	0.01
0.998	0.02
1.0	0.03
0.992	0.04
0.976	0.05
0.982	0.060000000000000005
0.95	0.06999999999999999
0.954	0.08
0.936	0.09
0.904	0.09999999999999999
0.892	0.11
0.896	0.12
0.884	0.13
0.864	0.14
0.86	0.15000000000000002
0.794	0.16
0.798	0.17
0.766	0.18000000000000002
0.748	0.19
0.75	0.2
0.73	0.21000000000000002
0.694	0.22
0.67	0.23
0.646	0.24000000000000002
0.62	0.25
0.592	0.26
0.606	0.27
0.578	0.28
0.5	0.29000000000000004
0.522	0.3

أما بالنسبة للطبقة R1 فالنتائج موضحة بالجدول التالي.

الجدول 7: احتمالية الخطأ مقابل معدل الاسترجاع لخطأZ للطبقة R1

Recovery Rate	Probability
1.0	0.01
1.0	0.02
0.998	0.03
1.0	0.04
0.996	0.05
0.996	0.060000000000000005
0.994	0.06999999999999999
0.986	0.08
0.964	0.09
0.956	0.09999999999999999
0.936	0.11
0.906	0.12
0.888	0.13
0.858	0.14
0.85	0.15000000000000002
0.814	0.16
0.778	0.17
0.748	0.18000000000000002
0.646	0.19
0.684	0.2
0.672	0.21000000000000002
0.624	0.22
0.586	0.23
0.602	0.24000000000000002
0.554	0.25

0.516	0.26
0.508	0.27
0.466	0.28
0.456	0.29000000000000004
0.456	0.3



شكل12: منحنى معدل الاسترجاع

(Recovery Rate Curve) للشيفرة ذات المعدّل الصفري (Zero-Rate) تحت ضجيج باولي النقي من النوع Z، مع عتبة تصحيح تقارب 15%~.

3.4 خطأ إزالة الاستقطاب (Depolarizing Error)

في هذا النوع من الأخطاء تؤثر أخطاء باولي جميعها بنسب متساوية حيث.

$$X = Y = Z = \frac{1}{3}$$

كما يمكننا تغيير النسبة 1/3 ولكن نحصل على معدلات خطا بنسب مختلفة. الجدول التالي نتائج المحاكاة التي تم فيها اختبار الشيفرة تحت خطأ إزالة الاستقطاب بالنسبة للطبقة R0.

الجدول 8: احتمالية الخطأ مقابل معدل الاسترجاع لخطأ Depolarizing للطبقة R0

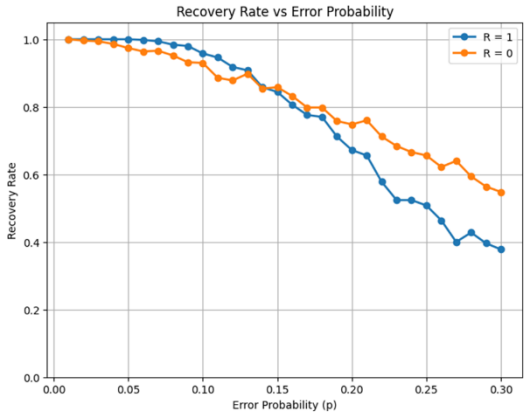
Recovery Rate	Probability
1.0	0.01
0.996	0.02
0.994	0.03
0.986	0.04
0.974	0.05
0.964	0.060000000000000005
0.966	0.06999999999999999
0.952	0.08
0.932	0.09
0.93	0.09999999999999999
0.886	0.11
0.878	0.12
0.898	0.13
0.854	0.14
0.858	0.15000000000000002
0.832	0.16
0.798	0.17
0.798	0.18000000000000002
0.758	0.19
0.748	0.2
0.76	0.21000000000000002
0.712	0.22
0.684	0.23
0.666	0.24000000000000002
0.656	0.25
0.622	0.26
0.64	0.27
0.594	0.28
0.564	0.29000000000000004
0.548	0.3

أما بالنسبة للطبقة R1 فالنتائج موضحة بالجدول التالي

الجدول 9: احتمالية الخطأ مقابل معدل الاسترجاع لخطأ Depolarizing للطبقة R1

Recovery Rate	Probability
1.0	0.01
1.0	0.02
1.0	0.03
1.0	0.04
1.0	0.05
0.998	0.060000000000000005
0.994	0.06999999999999999
0.984	0.08
0.98	0.09
0.958	0.09999999999999999
0.946	0.11
0.918	0.12
0.908	0.13
0.858	0.14
0.844	0.15000000000000002
0.806	0.16
0.776	0.17
0.77	0.18000000000000002
0.712	0.19
0.672	0.2
0.656	0.21000000000000002
0.578	0.22
0.524	0.23
0.524	0.24000000000000002
0.508	0.25
0.464	0.26

0.4	0.27
0.428	0.28
0.396	0.29000000000000004
0.378	0.3



شكل13: منحنى معدل الاسترجاع (Recovery Rate Curve)
للشيفرة Happy/Steane ذات المعدّل الصفري (Zero-Rate) تحت ضجيج depolarizingمع عتبة تصحيح تقارب 15%~.

3.5 ملخص النتائج

من خلال استخدام Integer Optimisation Decoder وتطبيقه على الشيفرة نلاحظ ان أخطاء باولي النقية وخطا إزالة الاستقطاب تختلف فيما بينها اختلاف بسيط بالنسبة لمعدلات الاسترجاع المنطقي المقابلة لاحتمالية الخطأ، وكذلك نلاحظ انه بالنسبة لجميع الأخطاء تملك الشيفرة عتبة قدرها *15%*، وتم ملاحظة ان معدلات الاسترجاع المنطقي تكون أكبر بالنسبة للطبقة R1 وذلك قبل ان تنهار الشيفرة عند العتبة 15%.

لقد تم اختبار الشيفرة تحت أخطاء محو الكيوبتات في الدراسة [10] وتم الحصول على عتبة قدرها 50%.

4. خاتمة

تم تصميم شيفرة Happy Plus Steane الكمومية الهلوغرافية، حيث تم دمج كل من الشفرتين في شيفرة واحدة حيث وضعنا التنسور المركزي من نوع شيفرة Happy أما الطبقة الأولى من التنسورات فكانت من نوع شيفرة Steane، ثم قمنا باختبار الشيفرة تحت أخطاء باولي النقية (X,Y,Z)، وخطأ إزالة الاستقطاب وذلك باستخدام Integer Optimization Decode، وذلك بالنسبة للطبقتين R0,R1 وحصلنا على عتبة خطأ تساوي تقريباً 15% وحصلنا على معدل الاسترجاع المنطقي مقابل كل احتمال خطأ معين. حيث أن عملية الدمج هذه بين الشيفرات لها فوائد عديدة في دعم الحوسبة الكمومية المتحملة للأخطاء وكذلك عملية الدمج تحقق امكانية الاستفادة من ميزات كل شيفرة ضمن بنية موحدة، بالإضافة الى تقليل عدد الكيوبتات الفيزيائية وبالتالي توفير كبير في الموارد، وأيضاً يمكّن الدمج من ضبط معدل الشيفرة والمسافة. ففي الشيفرة المدروسة في البحث حققنا انخفاضاً بعدد الكيوبتات بنسبة 28.6% تقريباً وذلك بالنسبة لشيفرة ستين مع الحفاظ على كفاءة التصحيح بالنسبة لخطأ إزالة الاستقطاب وزيادتها بالنسبة لأخطاء باولي النقية.

5. المراجع

[1] El Morsalani, M. (2025). *Encoding spacetime: A comprehensive review of holographic quantum codes* (Preprint). https://doi.org/10.13140/RG.2.2.34220.99205

[2] Fan, J. (2024). *Biased-noise threshold studies for holographic quantum error-correcting codes* (Master's thesis, Delft University of Technology, QuTech, Netherlands).

[3] Fan, J., Steinberg, M., Jahn, A., Cao, C., Feld, S., & Sarkar, A. (2024). *LEGO HQEC: A software tool for analyzing holographic quantum codes*. arXiv. https://arxiv.org/abs/2410.22861

[4] Fan, J., Steinberg, M., Jahn, A., Cao, C., & Feld, S. (2024). *Overcoming the zero-rate hashing bound with holographic quantum error correction*. arXiv. https://arxiv.org/abs/2408.06232

[5] Franco Rubio, A. (2016). *Holographic quantum error correcting codes* (Bachelor's thesis, Universidad de Zaragoza, Spain).

[6] Harris, R. J., Coupe, E., McMahon, N. A., Brennen, G. K., & Stace, T. M. (2020). *Decoding holographic codes with an integer optimisation decoder*. arXiv. https://arxiv.org/abs/2008.10206

[7] Nye, L. (2024). *Holographic quantum computation*. Carnegie Mellon University, School of Computer Science. https://doi.org/10.13140/RG.2.2.12587.66083

[8] Nye, L. (2025). Foundations of holographic quantum computation. *Journal of Applied Mathematics and Physics, 13*(1), 11–60.

[9] Pastawski, F., Yoshida, B., Harlow, D., & Preskill, J. (2015). Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. *Journal of High Energy Physics, 2015*(6), 149. https://doi.org/10.1007/JHEP06(2015)149

[10] Steinberg, M., Fan, J., Eisert, J., Feld, S., Jahn, A., & Cao, C. (2025). *Universal fault-tolerant logic with heterogeneous holographic codes*. arXiv. https://arxiv.org/abs/2504.10386

تحسين ترتيب الربط باستخدام خوارزمية MCTS الموجَهة بمعلومات الاستعلام في قواعد البيانات

محمد حاج حسن* أ.د. يُسر السيد سليمان الأتاسي**

الملخص:

يُعدّ تحسين الاستعلامات مكوّناً أساسياً في نظم إدارة قواعد البيانات، إذ يرتبط ارتباطاً مباشراً بكفاءة تنفيذ الاستعلامات. ويُعدّ اختيار ترتيب عمليات الربط عاملاً حاسماً في أداء استعلامات قواعد البيانات، نظراً لتأثيره المباشر على زمن تنفيذ الاستعلامات. تكمن صعوبة هذه المهمة في اتساع فضاء البحث مع زيادة عدد جداول الاستعلام، مما يجعل مسألة اختيار ترتيب الربط من المسائل المعقدة حسابياً NP-hard.

في هذه الدراسة، نقدم نهجاً لتحسين أداء قواعد البيانات من خلال اختيار ترتيب عمليات الربط باستخدام تقنيات التعلم المعزّز. تعتمد طريقتنا المقترحة على خوارزمية Monte Carlo Tree Search (MCTS). ترتكز الفكرة الرئيسية على محاكاة العديد من ترتيبات الربط الممكنة في بنية شجرية واحدة وتطبيق MCTS الموجهة بمعلومات خاصة بالاستعلام (حجم الجدول) لاختيار ترتيب التنفيذ الذي يحقق أعلى أداء ممكن.

نقوم بإجراء مجموعة من التجارب لتقييم فعالية طريقتنا. تظهر النتائج أداء أكثر استقراراً وتحسيناً ملحوظاً عبر مختلف أحجام الاستعلامات والقدرة على التكيف مع تعقيد البيانات بشكل عام. على وجه التحديد تُظهر النتائج أن الطريقة المقترحة تحقق تسريع قدره 1.14x عند المقارنة مع Alphajoin1.0 على مقياس WRL في حين حققت تسريعاً قدره 1.3× مقارنة بـ Alphajoin1.0 على مقياس GMRL.

تساهم هذه الدراسة في المجال من خلال تقديم منهجية اختيار ترتيب الربط دون معرفة واسعة بمجال إدارة قواعد البيانات ويؤكد عملنا على إمكانيات دمج تعلم الآلة مع أنظمة إدارة قواعد البيانات لتحسين الأداء.

كلمات مفتاحية: عملية الربط، تحسين ترتيب الربط، تعلم الآلة، التعلم المعزز، خوارزمية MCTS.

*) طالب ماجستير: قسم هندسة البرمجيات ونظم المعلومات – كلية الهندسة المعلوماتية – جامعة حمص – حمص – سوريا

**) أستاذ في هندسة البرمجيات ونظم المعلومات: قسم هندسة البرمجيات ونظم المعلومات – كلية الهندسة المعلوماتية – جامعة حمص – حمص – سوريا

Join Order Optimization using query-information-driven MCTS algorithm in Databases

* M.H.Hasan ** Y.Atassi

Abstract:

Query optimization is a fundamental component of database management systems, as it is directly related to the efficiency of query execution. Selecting the join order is a critical factor in database query performance due to its direct impact on query execution time. The difficulty of this task lies in the rapid expansion of the search space as the number of tables in a

query increases, making join order selection a computationally complex NP-hard problem.

In this study, we propose an approach to improve database performance by selecting join orders using reinforcement-learning techniques. Our proposed method is based on the Monte Carlo Tree Search (MCTS) algorithm. The core idea is to simulate a large number of possible join orders within a single tree structure and apply MCTS guided by query-specific information (table size) to select the execution order that achieves the best possible performance.

We conduct a series of experiments to evaluate the effectiveness of our approach. The results demonstrate more stable performance and noticeable improvements across different query sizes, as well as an overall ability to adapt to data complexity. Specifically, the proposed method achieves a speedup of 1.14× compared to AlphaJoin1.0 under the WRL metric, while achieving a speedup of 1.3× compared to AlphaJoin1.0 under the GMRL metric.

This study contributes to the field by presenting a methodology for join order selection that does not require extensive domain knowledge in database management systems. Our work highlights the potential of integrating machine learning with database management systems to improve performance.

Keywords: join operation, join order optimization, machine learning, reinforcement learning, MCTS algorithm.

---

*) Master Student: Department of Software Engineering and Information Systems – Faculty of Informatics Engineering – Homs University – Homs – Syria

**) Professor of Software Engineering and Information Systems: Department of Software Engineering and Information Systems – Faculty of Informatics Engineering – Homs University – Homs – Syria

1. مقدمة:

تشكل قواعد البيانات حجر الأساس لأنظمة المعلومات والبرمجيات الحديثة. وتتجلى أهميتها ليس فقط في قدرتها على معالجة كميات ضخمة من البيانات، بل أيضاً في دورها المحوري في دعم عمليات اتخاذ القرار. توفّر أنظمة إدارة قواعد البيانات Database Management Systems (DBMS) مجموعة واسعة من الإعدادات الضرورية لضمان الأداء الأمثل لقواعد البيانات، مما يُسهِم في سرعة وفعالية استرجاع المعلومات ذات الصلة. ويُعتبر تحسين الاستعلامات أحد المكونات الجوهرية في هذه الأنظمة، ويشكّل تحسين ترتيب عمليات الربط (Join Ordering) –وهو مشكلة فرعية ضمن تحسين الاستعلامات –أحد أبرز التحديات البحثية في هذا المجال. وتتمثل هذه المهمة في إيجاد الترتيب الأمثل للعلاقات (الجداول) المشاركة في عملية الربط، بهدف تقليل الكلفة المقدّرة لعملية التنفيذ إلى أقل ما يمكن [2].

تعتمد الممارسات الحالية في تحسين أداء قواعد البيانات بصورة رئيسة على عمليات ضبط يدوية ومعقدة ومنهجيات وتوصيفات تجريبية، ينفذها مسؤولي قواعد البيانات (DBAs) ممن يمتلكون خبرة متقدمة ومعرفة متراكمة. ومع التوسع المستمر في أحجام قواعد البيانات وتزايد درجة تعقيدها، يغدو هذا النهج اليدوي أقل جدوى وأكثر إشكالية وغير قادر على تلبية متطلبات الأداء العالي لقواعد البيانات والتطبيقات المتنوعة، والمستخدمين المتنوعين.

يمكن توظيف تقنيات الذكاء الاصطناعي للتغلب على القيود المرتبطة بالأساليب التقليدية في تحسين أداء قواعد البيانات، وذلك عبر قدرتها على استكشاف فضاء تصميم أوسع مما هو ممكن للبشر، واستبدال النهج التجريبي في معالجة المشاكل المعقدة بآليات أكثر منهجية وفعالية [3]. قدم الباحثون في [9] نموذجاً يعتمد على تقنيات التعلّم المعزز العميق لتحسين ترتيب عمليات الربط في قواعد البيانات، وفي [2] فقد طُرح إطار عمل يصوغ مشكلة تحسين ترتيب عمليات الربط كمشكلة تعلم معزز مُلاحظة بالكامل وعلى الرغم من النتائج الجيدة التي تم تحقيقها في

الأعمال السابقة إلا أنها كانت تعتمد على نماذج كلفة تفتقر إلى الدقة مما يؤثر في كثير من الأحيان على جودة الخطة النهائية.

أظهرت الدراسات الحديثة في مجال التعلّم المعزّز قدرة ملحوظة على معالجة المشكلات التي تتطلب مستوى عالٍ من الحدس، حيث حققت نتائج مشجعة مدعومة بخبرة مكتسبة من الملاحظة والتجربة. وقد برز هذا بوضوح في تطبيقات ناجحة لتعلّم استراتيجيات لعب الألعاب المعقدة ذات المساحات البحثية الواسعة. وانطلاقاً من هذه الإنجازات، **يهدف هذا العمل إلى استكشاف إمكانية توظيف التعلّم المعزّز –وبالتحديد خوارزمية Monte Carlo Tree Search (MCTS)–في تحسين أداء أنظمة قواعد البيانات، وذلك من خلال تحسين اختيار عملية الربط بشكل أكثر كفاءة.**

تتجسّد مساهمات هذه البحث في:

- ✓ توجيه خوارزمية UCT بمعلومات بنيوية عن الاستعلام: إذ أتاح دمج المعرفة ببنية الاستعلام تعزيز فهم الخوارزمية لسياق الاستعلام المعالج، والحدّ من مستويات عدم اليقين الملازمة للخوارزمية.
- ✓ اعتماد سياسة موجهة بطبيعة الاستعلام بدلاً من السياسات العشوائية: لأن استكشاف فضاء البحث الضخم بشكل عشوائي غير فعال يؤدي إلى ترتيبات ربط سيئة وبالتالي الحصول على خطط استعلام كارثية، بينما يضمن التوجيه المستند إلى خصائص الاستعلام مساراً أكثر فاعلية.

2. أهداف البحث:

تهدف دراستنا إلى توظيف تقنيات تعلم الآلة، وبشكل خاص التعلم المعزز، في تحسين عملية اختيار ترتيب الربط داخل قواعد البيانات. نسعى من خلال ذلك إلى **الوصول إلى خطة ربط مثلى للاستعلام**، بما يضمن **تقليص زمن التنفيذ** والحصول على النتائج المطلوبة في أقصر وقت ممكن،

وهو ما ينعكس إيجاباً على أداء قاعدة البيانات ككل. ويركّز الأسلوب المقترح على تحقيق الأهداف التالية:

- ✓ تعزيز كفاءة وفعالية استعلامات الربط عبر تطبيق خوارزميات التعلم المعزز.
- ✓ ايجاد حلول فعالة من ناحية زمن تنفيذ الاستعلامات.
- ✓ الحدّ من مستويات عدم اليقين الملازمة لخوارزمية Upper Confidence Bounds applied to Trees.
- ✓ توجيه عملية البحث ضمن فضاء البحث عن خطط ربط مثلى للاستعلامات بمعلومات خاصة بالاستعلام.

3. الدراسة المرجعية:

يقترح المؤلفون في [9] ReJOIN بوصفه نموذجاً يعتمد على تقنيات التعلّم المعزز العميق لتحسين ترتيب عمليات الربط في قواعد البيانات. يعتمد على النهج التقليدي القائم على الكلفة في نظم إدارة قواعد البيانات الحديثة. يُصاغ ترتيب الربط كمسألة تعلّم معزز. يعامل كل استعلام كحلقة تتضمن حالات وإجراءات؛ تمثل الحالات الأشجار الفرعية لشجرة الربط مع شروط الربط والاختيار،و تُعرّف الإجراءات بدمج هذه الأشجار وصولاً إلى الحالة النهائية، أما المكافأة تعادل مقلوب الكلفة. يعتمد على خوارزمية تدرج السياسة باستخدام شبكة عصبية تولّد احتمالات الإجراءات الممكنة، وتُضبط أوزانها تدريجياً استناداً إلى المكافآت لتحسين الأداء. يستخدم النموذج متجهاً خطياً لتمثيل الأشجار الفرعية، ومصفوفة ثنائية لتمثيل شروط الربط، إضافة إلى متجه أخر لتمثيل شروط الاختيار. عند تقييمه باستخدام JOB أظهر تفوقاً على PostgreSQL وQuickpick، حيث خفّض متوسط الكلفة إلى 80% من كلفة PostgreSQL وأنتج خططاً أقل كلفة بنسبة 20%، مع أداء أفضل أو مماثل في زمن التنفيذ وزمن تحسين شبه ثابت خطي مقارنة بالزيادة غير الخطية في PostgreSQL. ومع ذلك، فإن أبرز القيود تكمن في اعتماده على نماذج كلفة قد تفتقر إلى الدقة، واقتصاره على جانب ترتيب الربط دون معالجة اختيار المشغلات أو الفهارس، فضلاً عن حاجته إلى فترة تدريب طويلة، إضافة إلى بساطة أسلوب الترميز الذي قد لا يستوعب الخصائص الهيكلية أو الدلالية للاستعلام.

طرحت الدراسة [2] FOOP كإطار عمل يصوغ مشكلة تحسين ترتيب عمليات الربط كمشكلة تعلم معزز مُلاحظة بالكامل يتم تمثيلها وفق Markov Decision Process (MDP)، حيث تُعرِّف الحالات خطط الاستعلام الجزئية والأفعال عمليات الربط المحتملة بين الجداول ممثلة كمتجهات ميزات بينما المكافآت هي الكلفة السلبية لخطة الاستعلام النهائية. يتتبع FOOP جميع العلاقات والنتائج الوسيطة أثناء عملية التحسين و يدمج خوارزميات تعلم معزز متعددة: DQN, DDQN , PPO . استندت عملية التقييم إلى معيار JOB، فأظهرت DQN متوسط كلفة استعلام يقارب 0.2×10^{10} مع تباين كبير في النتائج أدى إلى خطط مرتفعة الكلفة لنصف الاستعلامات. بينما DDQN خفّضت المدى الربيعي والحد الأقصى للكلفة بنسبة 50% وتفوّقت PPO حيث قلّصت المدى الربيعي وأقصى كلفة خطة الاستعلام بأكثر من النصف مقارنةً بـ DDQN. أدى إعادة ترتيب مجموعات التدريب والاختبار وتطبيق التعلم الجماعي إلى تحسين أداء النماذج، إذ انخفضت التكاليف القصوى والمدى الربيعي والمتوسط. كما أظهرت النتائج أن DP تتفوق عادةً من حيث جودة الخطط، لكنها تعاني من تعقيد زمني $O(n!)$ في المقابل، تتميز PPO بتعقيد $O(n)$ وقد أظهر التحليل أن PPO يمكن أن تتفوق على DP في الاستعلامات التي تتراوح بين 3 و16 ربطاً. أما من حيث زمن التدريب، استغرقت DQN وDDQN حوالي 7 دقائق، بينما احتاجت PPO إلى 25 دقيقة. ومع التعلم الجماعي، ارتفع زمن التدريب ليصل إلى 45 دقيقة لـ DQN و125 دقيقة لـ PPO. وللمقارنة، يتطلب ReJoin نحو 3 ساعات. يواجه FOOP قيوداً تتعلق بالاعتماد على تقديرات غير دقيقة للكلفة، إضافةً إلى عبء حسابي مرتفع عند تطبيق التعلم الجماعي، مما يضعف قابليته للتوسع في بيئات قواعد البيانات الكبيرة.

اقترح الباحثون في [10] RTOS كمُحسِّن متعلّم يدمج التعلّم المعزز العميق مع بنية Tree-LSTM لمعالجة مسألة ترتيب الربط. تُصاغ كمسألة تعلّم معزز. تمثل الحالات غابات ربط وسيطة، والاجراءات شروط الربط الممكنة، والمكافآت تُستمد من الكلفة المقدّرة أو زمن الاستجابة الفعلي. يعتمد التدريب على التعلّم متعدد المهام يبدأ باستخدام قيم الكلفة للتعلّم الأولي ثم يُضبط

وفق زمن الاستجابة لتحسين النموذج. يستخدم شبكة DQN لتقريب قيم Q واختيار الاجراء الأمثل، مع الاحتفاظ بالخبرات السابقة في Memory Pool لدعم عملية التدريب. يقدّم تمثيلاً متعدد المستويات يشمل مصفوفة ثنائية للعلاقات، متجهات لترميز شروط الربط والاختيار، ومصفوفات لتضمينات الأعمدة لتمثيل الجداول. أما شجرة الربط فتعالج باستخدام Tree-LSTM عبر دمج نموذج N-ary لالتقاط ترتيب الأبناء وChild-Sum للأشجار غير المرتبة، مع بناء البيان ديناميكياً باستخدام البحث بالعمق أولاً لإنتاج تمثيلات تضمينية لكل شجرة ربط، ليُشكّل متجه الحالة النهائي دمجاً بين تمثيل الاستعلام والغابة. أظهر تقييم RTOS باستخدام PostgreSQL على معياري JOB وTPC-H أنه يحقق كلفة شبه مثالية مماثلة للبرمجة الديناميكية، متفوقاً على جميع الأساليب الأخرى وخاصة على مجموعة JOB بمتوسط كلفة نسبية (MRC) قدره 1.01. كما حقق أفضل أداء في زمن الاستجابة بقيم Geometric Mean Relevant Latency (GMRL) بلغت 0.67 على JOB و0.92 على TPC-H، أي تحسناً يصل إلى 37٪ مقارنة بالبرمجة الديناميكية، مع قدرة عالية على التكيّف مع تغييرات المخطط. يتطلب تدريب RTOS وقتاً طويلاً، إذ تستغرق مرحلة تدريب الكلفة نحو ساعة ونصف، يعقبها ضبط زمن الاستجابة لما يقارب عشر ساعات.فضلاً عن الحاجة إلى بيانات تدريب كبيرة وتغذية راجعة مكثفة، إضافة إلى كلفة زمنية عالية للتدريب المعتمد على زمن الاستجابة، كما أن زمن التخطيط أطول نسبياً بسبب حسابات Tree-LSTM.

تُقدِّم الدراسة GTDD [13] كنهج يجمع التعلّم المعزز العميق وتقنيات تعلّم التمثيل المتقدّم لتحسين اختيار ترتيب عمليات الربط.يعتمد على وكيل قائم على Dueling-DQN يختار إجراءات الربط تدريجياً حتى اكتمال الخطة، ثم يُنفَّذ الترتيب عبر نظام إدارة قواعد البيانات الذي يوفّر تغذية راجعة تُخزَّن وتُستخدم لتحديث الوكيل بشكل تكراري. يمثل الأعمدة كمتجه يصف مشاركته في عمليات الربط وشروط الاختيار المرتبطة به. أما لتمثيل الجداول يدمج جميع تمثيلات الأعمدة الخاصة بالجدول الواحد، مع تضمين العلاقات المستخرجة من المخطط. كما يمثل الاستعلام باستخدام الشبكات العصبية الرسومية حيث تعبر الجداول عقداً وتمثل الحواف شروط الربط. ولتمثيل الحالة يلتقط تطور خطة الربط حيث تُعالج الحالات الوسيطة باستخدام Tree-LSTM، ويُدمج التمثيل

النهائي عبر شبكة عصبية متعددة الطبقات قبل اتخاذ القرار. يعتمد الإطار على التعلّم المنهجي بتقسيم الاستعلامات وفق درجة التعقيد لضمان استقرار التقارب، بينما صُمِّمت دالة المكافأة كنسبة تقارن كلفة الخطة الأساسية بالخطة المتعلمة لتوحيد التغذية الراجعة عبر مختلف الاستعلامات. أظهرت التجارب أن GTDD يتفوّق باستمرار على أحدث المُحسِّنات المتعلَّمة مثل DQ، ReJOIN، RTOS، JOGGER وعلى المناهج الجينية. فقد حقق تقارباً أسرع وأدنى معدل خطأ متوسط علىJOB. كما حقق أفضل Geometric Mean Relevant Latency (GMRL) على JOB وTPC-H. على مستوى القوالب، تفوق GTDD على RTOS في 87% وعلى JOGGER في 78%. وبلغ GMRL قيمة 0.60381 على JOB، محققاً تحسناً مقارنةً بـ GTD (+20%)، RTOS (+23%)، JOGGER (+56%). يتطلب GTDD وقت تدريب أطول من الأساليب التقليدية بسبب العبء الحسابي لبنية Dueling-DQN، كما إن الكلفة المرتفعة للتدريب تحد من إمكانية تطبيقه عملياً في البيئات ذات المتطلبات الزمنية الصارمة. فضلاً عن أن دالة المكافأة صُمِّمت كنسبة بين كلفة الخطط الأساسية والمُتعلمة لتوفير تغذية راجعة موحّدة، إلا أن هذا التجريد قد يكون مبسطاً بشكل مفرط، مما قد يؤثر على جودة الخطة النهائية.

انجز الباحثون في [8] AlphaJoin وهو مُحسّن للاستعلامات مستوحى من AlphaGo، يعتمد على خوارزمية MCTS لتحسين ترتيب الربط. يتألف الإصدار الأول AlphaJoin 1.0 من شبكة OVN للتنبؤ بدرجة زمن التنفيذ لخطط الاستعلام الجزئية وخوارزمية MCTS لاستكشاف فضاء الترتيبات الربط باستخدام خوارزمية UCT لتحقيق التوازن بين الاستكشاف والاستغلال. أما الإصدار الثاني AlphaJoin 2.0 فأضاف شبكة ADN لتحديد ما إذا كان الأنسب استخدام AlphaJoin أو الاعتماد على PostgreSQL. يعتمد النظام على ترميزين متكاملين: ترميز SQL لتمثيل بنية الاستعلام عبر مصفوفة تجاور وترميز السمات المستخدمة في الشروط، وترميز خطة التنفيذ لتمثيل الخطط الجزئية أو الكاملة باستخدام أرقام ترتيبية تشير إلى أولوية عمليات الربط. أظهر تقييم AlphaJoin باستخدام معيار JOB تفوقاً واضحاً على PostgreSQL وNEO حتى مع احتساب زمن البحث. كما حقق AlphaJoin 2.0 انخفاضاً في زمن التنفيذ بنسبة 60.1% مقارنةً بـPostgreSQL، و38.7% مقارنةً بـNEO، و31.3% مقارنةً

بـAlphaJoin1.0،، تقليص نسبة الاستعلامات التي تفوّق فيها PostgreSQL من 48% في الإصدار الأول إلى 9% فقط في الإصدار الثاني.تُظهر نتائج AlphaJoin بعض القيود، أبرزها أن زيادة عامل البحث تُحسّن زمن التنفيذ لكنها ترفع زمن البحث بسبب كثافة محاكاة MCTS. كما أن شبكة OVN تحقق دقة محدودة (نحو 60.9% على JOB)، مما يجعل النظام يعتمد على المحاكاة للتعويض، إضافةً إلى ضعف أداء AlphaJoin 1.0 في الاستعلامات السريعة وعدم اليقين في خوارزمية UCT الذي قد يؤدي إلى ترتيبات غير مثالية.

يتضح من خلال مراجعة الدراسات المرجعية أن معظم الطرق السابقة كانت تعتمد في مناهجها على نماذج كلفة تقديرية تفتقر إلى الدقة مثل [2,9,10] وهو ما انعكس سلباً على جودة خطط الاستعلامات الناتجة. كما لجأت بعض الدراسات إلى أساليب محاكاة عشوائية وخوارزميات تتسم بدرجة من عدم اليقين مثل [8] الأمر الذي يؤدي إلى ترتيبات ربط دون المستوى الأمثل. وانطلاقاً من هذه الملاحظات، يهدف هذا العمل –على خلاف [2,9,10]– إلى تحسين اختيار ترتيب الربط بالاعتماد على أزمنة تنفيذ الاستعلامات بدلاً من النماذج التقديرية غير الدقيقة. كما نسعى إلى معالجة القيود المرتبطة في [8]، وذلك عبر تقليل مستويات عدم اليقين واستبدال المكوّنات العشوائية بأخرى موجّهة بمعلومات الاستعلام، بما يحدّ من احتمالية اختيار ترتيبات ربط ضعيفة ويعزز الوصول إلى حلول أكثر قرباً من المثالية.

4. طرائق البحث:

سنبدأ بطرح مجموعة من المفاهيم الأساسية. نبدأ بتقديم تعريف شامل لعملية الربط في استعلامات SQL. يلي ذلك التطرق إلى مفهوم تحسين الاستعلامات، من ثم نستعرض خوارزمية MCTS بإيجاز وفي الختام، ننتقل إلى شرح الطريقة المقترحة التي تستند إلى مجموعة من الأفكار المستوحاة من العمل المقدم في [8].

1.4عملية الربط:

تُعد عملية الربط إحدى الركائز الجوهرية في استعلامات قواعد البيانات العلائقية، إذ تُستخدم لاسترجاع البيانات من علاقات متعددة بالاعتماد على السمات المشتركة فيما بينها. يمكن أن يُنفّذ الربط بصيغة ثنائية الاتجاه أو متعددة الاتجاهات؛ ففي الحالة الأولى يُطبّق الربط بين علاقتين فقط، بينما في الحالة الثانية يُستخدم لدمج أكثر من علاقتين. وعادةً ما يُنفّذ الربط بين n علاقة عبر سلسلة من $(n-1)$ عملية ربط ثنائية الاتجاه.

2.4 تحسين الاستعلامات:

يلعب تحسين الاستعلامات دوراً محورياً في أداء نظم إدارة قواعد البيانات الحديثة، لا سيما في ظل ظروف البيانات الضخمة. في الأنظمة الواقعية، يمكن أن تتسبب الاستعلامات غير الفعالة في تأخر كبير، واستهلاك مفرط للموارد، وتدهور الخدمة ومع ازدياد اعتماد المؤسسات على الاستعلامات المعقدة، يصبح تحسين الاستعلامات الفعال ضرورياً لضمان الاستجابة، وقابلية التوسع، وكفاءة الكلفة [5]. إن الهدف الأساسي من تحسين الاستعلامات هو الوصول إلى خطة تنفيذية ذات أقل كلفة ممكنة لا سيما من حيث زمن التنفيذ [6].

3.4 خوارزمية Monte Carlo Tree Search (MCTS) العامة:

هي خوارزمية لاتخاذ القرارات تتبع الإجراء العام للبحث في شجرة Monte Carlo [23]. تُوسّع الشجرة من خلال عمليات محاكاة أثناء البحث في الفضاء حتى اتخاذ قرار، وتُعيد النتيجة النهائية للقرار ("فوز" أو "خسارة") إلى عُقد الشجرة لتحديثها. بعد عدد كبير من عمليات المحاكاة، تُمثّل معلومات كل عُقدة نسبة عدد القرارات الصحيحة إلى إجمالي عدد عمليات المحاكاة عند هذه العُقدة، مما يُشير إلى قيمة هذه العُقدة [8]. وقد كان لها بالفعل تأثير عميق على مناهج الذكاء الاصطناعي في المجالات التي يمكن تمثيلها بأشجار من القرارات المتسلسلة، وخاصة الألعاب ومسائل التخطيط.

تعتمد MCTS على مفهومين أساسيين: أن القيمة الحقيقية لأي فعل يمكن تقريبها باستخدام المحاكاة العشوائية؛ وأن هذه القيم يمكن استخدامها بكفاءة لتعديل السياسة باتجاه استراتيجية

الأفضل أولاً. تقوم الخوارزمية تدريجياً ببناء شجرة جزئية، مسترشدة بنتائج الاستكشاف السابق لتلك الشجرة. تُستخدم الشجرة لتقدير قيم الحركات، وتصبح هذه التقديرات (خاصة تقديرات الحركات الواعدة) أكثر دقة مع تقدم بناء الشجرة. تتضمن الخوارزمية الأساسية بناء شجرة بحث بشكل تكراري حتى الوصول إلى حد معين من الموارد الحسابية المحددة مسبقاً –عادةً ما يكون قيداً زمنياً أو متعلقاً بالذاكرة أو بعدد التكرارات –وعند هذه النقطة، يتوقف البحث ويتم إرجاع أفضل إجراء. تمثل كل عقدة في شجرة البحث حالة من حالات المجال، وتمثل الروابط الموجهة إلى العقد الفرعية إجراءات تؤدي إلى الحالات اللاحقة [23].

يتم تطبيق أربع خطوات لكل تكرار بحث:

الاختيار (Selection)، التوسيع (Expansion)، المحاكاة (Roll-out)، والانتشار العكسي (Backup). إلا أنه بدلاً من تقييم العقد الطرفية في الشجرة عبر خطوة المحاكاة، يتم استخدام شبكة عصبية لتوفير سياسة وتقديرات لقيمة الحالة للعقدة الجديدة في الشجرة [7].

i. الاختيار: بدءاً من العقدة الجذر، يتم تطبيق سياسة اختيار بشكل متكرر للنزول عبر الشجرة حتى الوصول إلى العقدة القابلة للتوسيع الأكثر إلحاحاً. تكون العقدة قابلة للتوسيع إذا كانت تمثل حالة غير نهائية ولها عقد فرعية غير مُزارة (أي غير مُوسعة)
ii. التوسيع: تتم إضافة عقدة فرعية واحدة (أو أكثر) لتوسيع الشجرة، وفقاً للإجراءات المتاحة.
iii. المحاكاة: يتم تشغيل محاكاة من العقدة (العقد) الجديدة وفقاً للسياسة لإنتاج نتيجة.
iv. الانتشار العكسي: يتم نسخ نتيجة المحاكاة أي نشرها عكسياً عبر العقد المختارة لتحديث إحصاءاتها.

4.4 الطريقة المقترحة:

تتمثل الفكرة الرئيسية لمنهجيتنا المقترحة على تمثيل جميع ترتيبات الربط الممكنة ضمن بنية شجرية واحدة، بما يتيح تسهيل عملية البحث والاستكشاف. عقب ذلك، نطبق خوارزمية MCTS موجهة معلومات ذات صلة بالاستعلام وقاعدة البيانات (مثل حجم الجدول – من الجدير بالذكر بأننا أجرينا دراسة مطولة باستخدام معلومات أخرى غير حجم الجدول مثل انتقائية عملية الربط

وانتقائية الشروط لكن لم نتطرق لها هنا حرصاً على عدم إطالة المقالة)، بغرض تعلم وتوليد ترتيبات ربط عالية الكفاءة داخل مُحسّن الاستعلام. نوظف شبكة عصبية [8] للتنبؤ بزمن تنفيذ الاستعلام لخطة معينة، بحيث ندمج هذه الشبكة ضمن إطار عمل MCTS لتقييم خطط الاستعلام المُرشّحة.

1.4.4 قاعدة البيانات و عبء العمل:

قاعدة البيانات: تم اختيار قاعدة بيانات Internet Movie Database (IMDB). تتميز هذه القاعدة بغزارة بياناتها وتنوعها، حيث تحتوي على معلومات موسَّعة تتعلق بالأفلام، والممثلين، والمخرجين، وشركات الإنتاج، إضافةً إلى تفاصيل دقيقة حول تواريخ الإصدار وغيرها من المعلومات ذات الصلة [15].

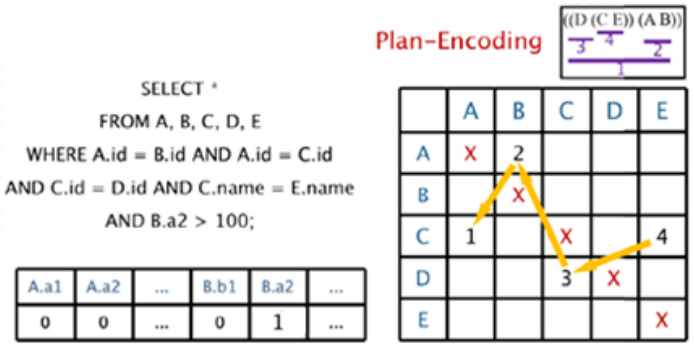
عبء العمل: يُعَدّ معيار ترتيب الربط (Join Order Benchmark – JOB) أحد المعايير المرجعية المهمة في تقييم أداء مُحسِّنات الاستعلامات. يتكوّن هذا المعيار من 113 استعلاماً تحليلياً صُمِّمت لاختبار كفاءة محسنات الاستعلامات على قاعدة بيانات IMDB المجمَّعة من مصادر واقعية [11]. تتراوح هذه الاستعلامات بين 3 و 16عملية ربط، بمتوسط يقارب 8 عمليات ربط لكل استعلام.

2.4.4 ترميز البيانات:

نقوم بتمثيل خطة التنفيذ، سواء كانت جزئية أو كاملة. ويُشترط أن يكون هناك تطابق صارم بين كل ترميز وبين ترتيب الربط المقابل له، بحيث تكون آلية الترميز قابلة للترميز وفك الترميز في آن واحد. يتكون ترميز الخطة أيضاً من عنصرين:

- ❖ مخطط الربط (Join Graph): يُمثَّل باستخدام مصفوفة تجاور تُظهر ترتيب عمليات الربط باستخدام أرقام مختلفة. كلما ارتفع الرقم في المصفوفة، دلّ ذلك على أولوية أعلى لعملية الربط بين الجدولين المعنيين.

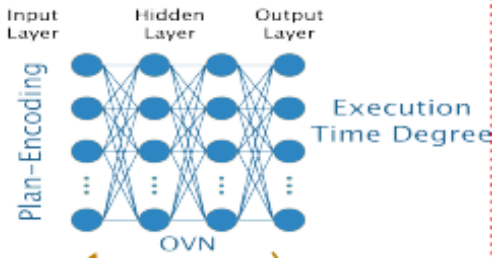
- ترميز الخصائص (Attributes Encoding): يمثل باستخدام الترميز الواحد الساخن One-Hot Encoding بحيث تأخذ الخاصية المشاركة في شرط الاستعلام القيمة 1، بينما تأخذ الخصائص غير المشاركة القيمة 0.



الشكل 4-1: طريقة الترميز Plan Encoding [8]

3.4.4 بناء نموذج الشبكة العصبية Order Value Network (OVN):

تُعد شبكة قيمة الترتيب (Order Value Network -OVN) نموذجاً عصبياً صُمِّم للتنبؤ بدرجة زمن التنفيذ المثلى لاستعلام جزئي ضمن خطة تنفيذية. كما هو موضح في الشكل 4-2، تتألف البنية المعمارية للشبكة من طبقة إدخال (I)، يليها ثلاث طبقات مخفية بأحجام (2048، 512، 128 وحدة) مفعّلة باستخدام دالة ReLU، ثم طبقة إخراج (O).



الشكل 4-2: البنية المعمارية للشبكة العصبية OVN [8]

يهدف هذا النموذج إلى تقدير ما إذا كانت خطط الاستعلام سريعة أو بطيئة، حيث تُغذّى الشبكة بترميز خطة الاستعلام، ويكون الإخراج على شكل تصنيف متعدد الفئات. في هذا السياق، تُستخدم أربع فئات محتملة (K = 4) لتمثيل درجات زمن التنفيذ (من 0 إلى 4)، بحيث تعكس الدرجة الأدنى زمن تنفيذ أقل. لتوليد توزيع احتمالي مناسب على درجات زمن التنفيذ، تعتمد الشبكة على دالة Softmax في طبقة الإخراج، كما يتم تطبيق تقنية Dropout Regularization للحد من مشكلة فرط التعلّم (Overfitting).

4.4.4 توليد بيانات مجموعة تدريب الشبكة العصبية OVN:

نجمع بيانات التدريب الخاصة بشبكة القيمة (OVN) من خلال تنفيذ عدد كبير من خطط الاستعلام على قاعدة بيانات IMDB باستخدام نظام PostgreSQL، وذلك بهدف توليد ترتيبات ربط متعددة، تنفيذها، وتسجيل كل من أزمنة التنفيذ وخطط الاستعلام الناتجة.

نبدأ بتهيئة مجموعة البيانات، ثم قراءة ملفات الاستعلام المتاحة. ولكل استعلام، نولد مجموعة ترتيبات الربط الممكنة عبر استكشاف تركيبات الربط المتوافقة مع شروط الربط المحددة ضمن الاستعلام بشكل تكراري.

ننفّذ كل ترتيب ربط صالح على PostgreSQL مع تفعيل إضافة pg_hint_plan [19]، التي تتيح فرض ترتيبات ربط محددة على النظام بدلاً من الاعتماد على الترتيب الافتراضي الذي يقترحه المُحسِّن الأصلي، وذلك باستخدام ما يُعرف بتلميحات hints.

ولضمان قياس أزمنة تنفيذ الاستعلامات في ظروف متسقة وخالية من تأثيرات التخزين المؤقت، نقوم بإعادة تشغيل PostgreSQL ومسح ذاكرة التخزين المؤقت للنظام قبل تنفيذ الاستعلامات، مما يزيل أي تحيز محتمل في الأداء ناجم عن البيانات المخزنة مسبقاً.

بعد ذلك، نسجِّل نتائج التنفيذ التي نستخرجها باستخدام الأمر EXPLAIN ANALYZE في PostgreSQL، بما يشمل زمن التنفيذ، خطة الاستعلام، وترتيب الربط المستخدم.

تتكرر هذه العملية بصورة مستمرة، بما يضمن كافة الاستعلامات وتغطية فضاء واسع من ترتيبات الربط، وبالتالي توفير بيانات تدريبية غنية ومتنوعة.

ملاحظة: نقوم بفرض قيود على عدد ترتيبات الربط في الاستعلامات التي تتضمن عدداً كبيراً من الجداول (خمسة جداول فأكثر)، وذلك بهدف تحقيق توازن بين تنوع البيانات والجدوى الحسابية.

تنبع هذه الحاجة من أننا لم نفرض أي قيود على بنية شجرة الاستعلام الناتجة، سواء كانت ذات عمق يساري أو يميني أو ذات بنية كثيفة، ونتيجة لذلك يصبح فضاء البدائل الممكنة غير قابل للإدارة بسرعة كبيرة عند أخذ جميع أشكال الأشجار الكثيفة بعين الاعتبار؛ فعلى سبيل المثال، ينتج عن استعلام يتكون من خمسة جداول فقط 1680 ترتيباً محتملاً.

5.4.4 المعالجة المسبقة لمجموعة بيانات التدريب المولَّدة:

من أجل تدريب الشبكة العصبية OVN للتنبؤ بزمن تنفيذ الاستعلامات على مجموعة البيانات المولَّدة، كان من الضروري إجراء معالجة مسبقة تتضمن الخطوات التالية:

- استبعاد بعض الاستعلامات ذات الحجم الكبير: هناك مجموعة من الاستعلامات ضمن نماذج مختلفة من JOB تحتوي على عدد ضخم من الجداول (12 جدولاً فأكثر). لم تتمكن الإضافة pg_hint_plan من فرض ترتيبات ربط بديلة على نظام PostgreSQL، إذ بقي النظام يعتمد على الترتيب الذي يوصي به المحسّن الداخلي. ونتيجة لذلك، لم يكن بالإمكان الحصول على زمن التنفيذ لهذه الفئة من الاستعلامات، مما استدعى استبعادها من مجموعة البيانات. وتشمل هذه النماذج: 24، 26، 27، 28، 29، 30، 33.

- تحديد عتبة زمنية لتنفيذ الخطط: بعض الخطط ضمن مجموعة البيانات المولَّدة استغرقت زمناً طويلاً جداً للتنفيذ، وصل في بعض الحالات إلى أكثر من ساعة. ولتفادي الاستهلاك المفرط للوقت والموارد الحاسوبية، كان من الضروري وضع عتبة زمنية قصوى. ولتحديد هذه القيمة، أجرينا سلسلة من التجارب باستخدام عتبات زمنية مختلفة: 5 دقائق، 10 دقائق، 15 دقيقة، 20 دقيقة، 30 دقيقة، 60 دقيقة. وقد أظهرت النتائج أن قيمة 5 دقائق هي الأكثر كفاءة.

- ترميز الخطط المولَّدة: من أجل تدريب الشبكة العصبية OVN على التنبؤ بزمن تنفيذ الاستعلامات، كان من الضروري تحويل مجموعة الخطط الناتجة إلى تمثيل قابل للمعالجة من قبل النموذج. ولتحقيق ذلك، قمنا بترميز جميع الخطط وفقاً لترميز الخطة المشروح سابقاً.

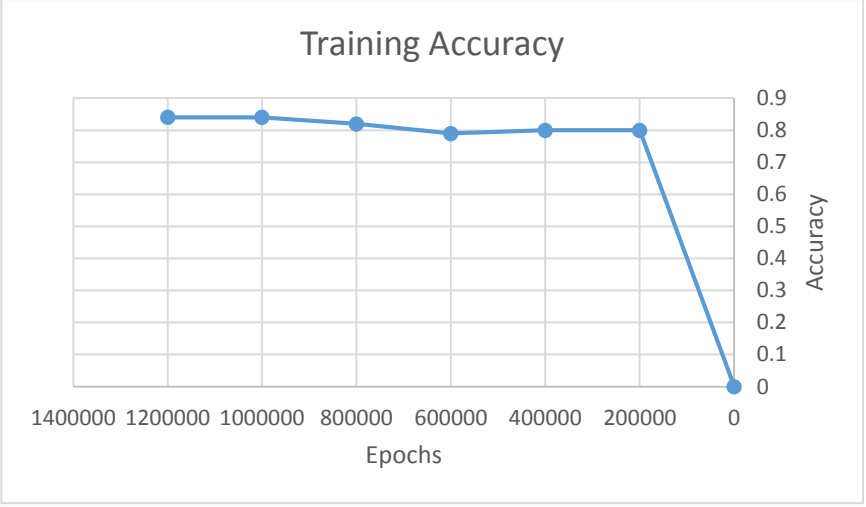
- إضافة ترميز الخصائص: يتم الدمج بين ترميز الخطة وشعاع ترميز الخصائص لنشكل ترميز SQL.

- إضافة الفئات: نقوم بترتيب مجموعة البيانات تصاعدياً اعتماداً على أزمنة التنفيذ ومن ثم نقسم مجموعة البيانات المرتبة إلى k مجموعة بحيث كل مجموعة لها نفس الفئة (تستخدم لتدريب الشبكة باستخدام التعلم الخاضع للإشراف) والتي تمثل درجة سرعة التنفيذ.

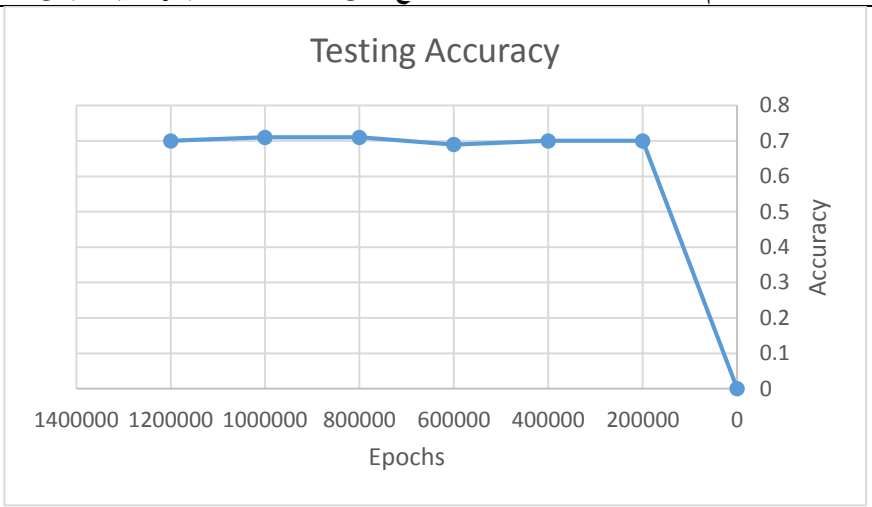
وفي نهاية المطاف نقسم مجموعة البيانات إلى 70% بيانات تدريب و30% بيانات اختبار.

6.4.4 تدريب الشبكة العصبية OVN:

يجري تدريب نموذج الشبكة العصبية OVN بالاعتماد على مجموعة بيانات التدريب، وذلك بهدف تقليل قيمة دالة الخسارة Cross Entropy [16] بين خطط الاستعلام التاريخية وزمن التنفيذ المتوقع لها. وتُعد دالة Cross Entropy المعيار القياسي في مشاكل التصنيف متعدد الفئات، إذ يعد مقياساً لمدى التقارب التوزيعات الاحتمالية بين القيم المتوقعة والقيم الفعلية[8]. استغرقت عملية التدريب ما يقارب خمس ساعات، وأسفرت النتائج عن تحقيق دقة بلغت 84% على مجموعة بيانات التدريب، مقابل 70% على مجموعة بيانات الاختبار. يوضح الشكل 4-3 والشكل 4-4 المنحنى البياني لمستوى الدقة المحقق على كل من مجموعة التدريب ومجموعة الاختبار على الترتيب.

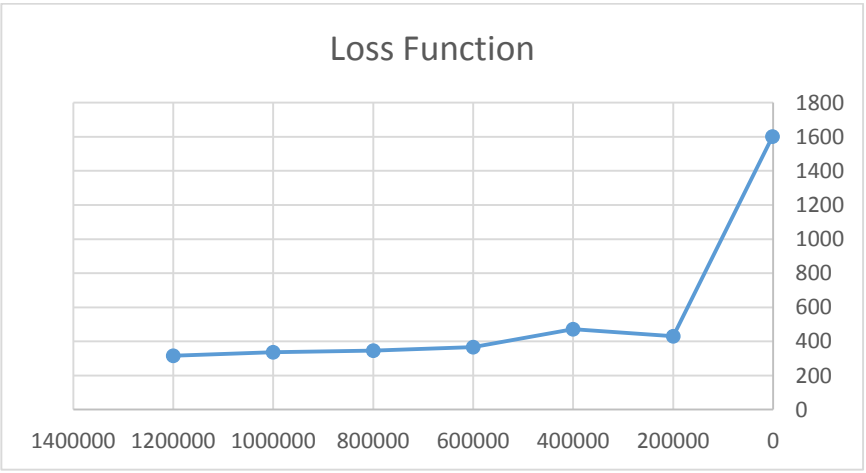


الشكل 4-3: المنحني البياني لدقة تدريب الشبكة OVN على مجموعة بيانات التدريب



الشكل 4-4: المنحني البياني لدقة تدريب الشبكة OVN على مجموعة بيانات الاختبار

يوضح الشكل 4-5 المنحني البياني لدالة الخسارة أثناء عملية التدريب:



الشكل 4-5: المنحني البياني لدالة الخسارة أثناء عملية تدريب الشبكة

7.4.4 خوارزمية Monte Carlo Tree Search (MCTS) المعدلة:

قمنا بإعادة تنفيذ خوارزمية Monte Carlo Tree Search (MCTS) كما وردت في [8]، مع إدخال مجموعة من التحسينات البنيوية والمنهجية على مراحلها المختلفة، وذلك على النحو الآتي:

- ✓ تطوير نسخة معدّلة من خوارزمية UCT: قمنا بتعديل خوارزمية Upper Confidence Bounds applied to Trees (UCT) لتعتمد على معلومات موجّهة بالاستعلام (حجم الجدول)، إضافةً إلى مجموعة من القواعد الاستدلالية التي تُستخدم في مرحلة اختيار العقدة التالية، بما يعزز جودة عملية التوسّع داخل الشجرة.
- ✓ استبدال سياسة المحاكاة العشوائية: قمنا باستخدام سياسة محاكاة موجهة بالمعلومات ذات الصلة بالاستعلام بدلاً من السياسة العشوائية.

بداية سنقدم خوارزمية UCT ثم نسلط الضوء على النسخة المعدلة منها يلي ذلك مناقشة تابع الجائزة وفي الختام، نستعرض خوارزمية MCTS بمراحلها الأربعة.

1.7.4.4 خوارزمية Upper Confidence Bounds applied to Trees (UCT):

تُعد خوارزمية حدود الثقة العليا المطبقة على الأشجار (UCT) إحدى خوارزميات البحث في أشجار الألعاب، وتهدف إلى معالجة مسألة اختيار العقدة المثلى ضمن فضاء البحث. وتعتمد هذه الخوارزمية على مبدأ الاستكشاف مقابل الاستغلال، بحيث توازن بين الاستفادة من المعرفة المتراكمة حول العقد التي تم اختبارها سابقاً، وبين استكشاف عقد جديدة لم تُجرَّب بعد، بما يحدّ من احتمالية الوقوع في الحلول المثلى المحلية. صيغة حساب UCT هي:

$$\mathbf{UCT}(v_i, v) = \frac{Q(v_i)}{N(v_i)} + C\sqrt{\frac{\log N(v)}{N(v_i)}}$$

حيث أن:

- تُمثل v_i العقدة الحالية التي يتم دراستها
- v العقدة الأبوية للعقدة v_i.
- $Q(v_i)$ عدد مرات اكتساب ميزة في العقدة الحالية.
- $N(v_i)$ $(N(v)$ إجمالي عدد مرات الوصول إلى العقد الحالية (العقد الأبوية).
- C مُعامل لضبط حساسية الاستكشاف.

يُشير الجزء الأول من الصيغة إلى الاستغلال الذي يُحدد احتمالية "الفوز" بعد اختيار هذه العقدة. أما الجزء الثاني من الصيغة فيُشير إلى الاستكشاف، حيث تزداد احتمالية استكشاف العقد الأخرى غير الممسوحة (بدلاً من v_i) إذا كانت قيمة $N(v_i)$ كبيرة. أن المسار من كل عقدة فرعية إلى العقدة الجذر في الشجرة يُمثل ترتيب ربط كامل [8].

بالاستفادة من [14] التي قامت بتطبيق خوارزمية PUCT (Predictor + UCT) في مرحلة الاختيار من MCTS كالتالي:

$$a = \arg\max_{a} Q(s,a) + c\ P(s,a)\ \frac{\sqrt{N(s)}}{1+N(s,a)}$$

حيث أن:

- $P(s, a)$ يرمز إلى الاحتمال المسبق لاختيار الإجراء a في الحالة s.
- $Q(s,a)$ هي متوسط قيمة الإجراء (الاستغلال).
- c هو مُعامل فائق يُعدِّل التوازن بين الاستكشاف والاستغلال.

- N(s) هو إجمالي عدد الزيارات للحالة s الحالية.
- N(s,a) هو عدد الزيارات لاتخاذ الإجراء a من الحالة s.

قمنا بتعديل خوارزمية UCT لتقوم بتقدير جودة الحالة (العقدة) من خلال تحليل بنية الربط الحالية وتطبيق قاعدة استدلالية تُفضل الربط بين الجداول الأصغر حجماً (وهو مبدأ شائع في تحسين الاستعلامات). غالباً ما يُقلل ربط الجداول الأصغر حجماً أولاً من أحجام النتائج الوسيطة مما يُساعد هذا في توجيه البحث نحو خطط استعلام أكثر كفاءة ويُحسّن الأداء.

صيغة حساب UCT المعدلة:

$$UCT = Exploitation + Exploration + Heuristic_value$$

حيث أن:

Exploitation تشير إلى جزء الاستغلال من صيغة UCT الاساسية

Exploration تشير إلى جزء الاكتشاف من صيغة UCT الاساسية

Heuristic_value قيمة استدلالية تعتمد على معلومات الاستعلامات (مثل حجم الجدول).

2.7.4.4 تابع الجائزة:

تمت إعادة بناء تابع المكافأة وفقاً لما ورد في المرجع [8]. في سياق تحسين الاستعلام، يتمثل الهدف الرئيسي في توليد خطة استعلام تحقق أدنى زمن تنفيذ ممكن؛ إذ إن انخفاض زمن التنفيذ يعكس زيادة احتمالية تحقيق الأداء الأمثل. وقد صُمم تابع المكافأة على النحو التالي:

$$R_j = (K - k_j) / K$$

حيث تمثل R_j قيمة المكافأة الخاصة بترتيب الربط j الذي يتم تقييمه باستخدام خوارزمية MCTS. تشير k إلى القيمة المرجعية لزمن التنفيذ المحددة مسبقاً ضمن شبكة OVN، بينما يرمز K_j (يتراوح بين 0 و k) إلى زمن التنفيذ المتوقع لخطة الربط j. تصل المكافأة إلى قيمتها العظمى، أي 1، عندما يُتنبأ بأن زمن التنفيذ للخطة يساوي الحد الأدنى الممكن، وهو 0.

3.7.4.4 خوارزمية MCTS لمشكلة تحسين ترتيب عمليات الربط:

تم تنفيذ خوارزمية MCTS وفقاً للمنهجية الواردة في المرجع [8]، مع إدخال مجموعة من التعديلات في مرحلتي الاختيار والمحاكاة. تُعد خوارزمية MCTS إحدى خوارزميات اتخاذ القرار التي تعتمد على توسيع الشجرة عبر عمليات محاكاة متكررة ضمن فضاء البحث حتى الوصول إلى قرار نهائي. يتم بعد ذلك إرجاع نتيجة القرار (فوز أو خسارة) إلى العقد الممثلة في الشجرة بغرض تحديثها. ومع تزايد عدد عمليات المحاكاة، تصبح قيمة كل عقدة معبّرة عن نسبة القرارات الصحيحة إلى إجمالي عدد المحاكاة المنفذة عند تلك العقدة، وهو ما يُحدد تقدير قيمتها. ولتقليص مساحة البحث الممتدة من العقدة الجذر إلى العقد الطرفية، تم تعريف عدد عمليات المحاكاة المطلوبة لكل عقدة على النحو التالي:

$$S_n = N_C * F_s$$

حيث يشير N_C إلى عدد العقد الفرعية المنبثقة من العقدة الحالية، بينما يمثل F_s عامل البحث الذي يتحكم في إجمالي زمن المحاكاة. وكلما ارتفعت قيمة F_s، ازداد زمن البحث الذي تستغرقه الخوارزمية. أما إجمالي عدد عمليات المحاكاة لكل استعلام SQL فيُعطى بالعلاقة:

$$\sum_1^J S_n$$

حيث يرمز J إلى عدد عمليات الربط في الاستعلام. تستمر هذه العملية حتى يتم استكشاف جميع ترتيبات الربط الممكنة أو بلوغ الحد الأقصى المحدد مسبقاً لعدد عمليات المحاكاة.

يمكن تقسيم كل عملية محاكاة في خوارزمية MCTS إلى أربع مراحل أساسية [8]:

- الاختيار: نطبيق خوارزمية UCT المعدلة لاختيار ترتيب الربط بدءاً من العقدة الجذرية وصولاً إلى العقدة الطرفية. عند مواجهة عقدة فرعية طرفية أثناء التنقل، تنتقل الخوارزمية إلى مرحلة التوسيع.
- التوسيع: إضافة عقدة فرعية جديدة (تمثل ترتيب ربط جديد) إلى الشجرة عند العقدة التي تم الوصول إليها بشكل أمثل خلال عملية الاختيار [8].
- المحاكاة: بدلاً من إجراء محاكاة عشوائية كاملة حتى نهاية خطة الاستعلام، نستخدم التوجيه الاستدلالي لتسريع عمليات المحاكاة وتحسين كفاءتها. بدلاً من اختيار إجراء عشوائي، نختار أفضل إجراء استدلالي حيث تُعطى الأولوية لعمليات الربط بين الجداول الأصغر لأن عمليات الربط الأصغر عادةً ما تُقلل وقت التشغيل. تجعل هذه السياسة التي قمنا بتنفيذها خوارزمية MCTS للبحث عن الجداول أكثر كفاءة في تحسين الاستعلام من حيث:
  - ✓ عمليات المحاكاة العشوائية البحتة تُهدر الوقت في استكشاف ترتيبات الربط غير المناسبة.
  - ✓ عبر إيلاء الأولوية لعمليات الربط ذات الحجم الأصغر، فإننا نوجَّه عملية المحاكاة نحو خطط أكثر كفاءة وقابلية للتنفيذ.
- الانتشار العكسي: إعادة تمرير نتائج المحاكاة من العقدة الجديدة إلى العقدة الجذرية. خلال هذه العملية، يُزاد العدد الإجمالي لعمليات المحاكاة المسجلة في كل عقدة بمقدار واحد. وإذا أسفرت المحاكاة عن زمن تنفيذ أقل، تُمنح العقدة مكافأة إضافية [8].

8.4.4 أهم الأدوات و المكتبات المستخدمة في البحث:

- ✓ PostgreSQL هو نظام قواعد بيانات علائقية كائنية قوي ومفتوح المصدر، يستخدم لغة SQL ، بالإضافة إلى العديد من الميزات التي تُخزّن وتُوسّع أحجام البيانات الأكثر تعقيدًا بأمان يتميز بموثوقيته، وسلامة بياناته، ومجموعة ميزاته القوية، وقابليته للتوسيع. يعمل PostgreSQL على جميع أنظمة التشغيل الرئيسية، وهو متوافق مع معايير ACID [17].

- ✓ PostgreSQL JDBC 42.2.6: برنامج تشغيل JDBC مفتوح المصدر مكتوب بلغة جافا، ويتصل باستخدام بروتوكول الشبكة الأصلي native network protocol لـ PostgreSQL. يُمكّن برنامج تشغيل JDBC لـ PostgreSQL برامج جافا من الاتصال بقاعدة بيانات PostgreSQL باستخدام شفرة جافا قياسية ومستقلة عن قواعد البيانات. لهذا السبب، فإن برنامج التشغيل مستقل عن النظام الأساسي يمكن استخدامه على أي نظام [18].

- ✓ pg_hint_plan : تتيح الإضافة pg_hint_plan تعديل خطط تنفيذ PostgreSQL باستخدام تلميحات hints في تعليقات SQL. تقرأ عبارات التلميح في تعليق ذي شكل خاص مُعطى لعبارة SQL. يمكن تحديد التلميح بإضافة البادئة "/*+" في بدايته ونهاية العبارة "*/". تتكون عبارات التلميح من أسماء التلميحات والمعلمات محاطة بأقواس ومفصولة بمسافات بيضاء. يمكن استخدام أسطر جديدة في عبارات التلميح لتحسين سهولة القراءة [19].

- ✓ PyTorch: إطار عمل مفتوح المصدر للتعلم الآلي يستخدم على نطاق واسع لبناء وتدريب نماذج التعلم العميق [20].

9.4.4 مقارنة بين الطريقة المقترحة و Alphajoin1.0:

يسلط الجدول 4-1 الضوء على أوجه التشابه والاختلاف بين منهجيتنا المقترحة والعمل المنجز في [8]:

الجدول 4-1: أوجه التشابه والاختلاف بين الطريقة المقترحة و Alphajoin1.0 [8]

المقارنة	أوجه التشابه	أوجه الاختلاف
طريقة الترميز	كلا الطريقتين تستخدمان الترميز plan encoding	لا يوجد
تقييم زمن خطط الاستعلامات	كلا الطريقتين تستخدمان OVN	لا يوجد
MCTS من أجل JOS	كلا الطريقتين تستخدمان MCTS لتوليد ترتيبات ربط للاستعلامات	[8] تستخدم خوارزمية UCT لاختيار العقد وتعتمد على اختيارات عشوائية في مرحلة المحاكاة بينما طريقتنا المقترحة تستخدم نسخة معدلة من خوارزمية UCT في مرحلة اختيار العقد وتطبق سياسات مقادة بمعلومات ذات صلة بالاستعلام لتنفيذ مرحلة المحاكاة

5. التجارب والنتائج:

في إطار التجارب المنفّذة، تم الاعتماد على بيئة Apache NetBeans IDE 14 لتوليد مجموعة البيانات المستخدمة في تدريب النماذج. كما استُخدم نظام إدارة قواعد البيانات PostgreSQL 16، وتم تحقيق الاتصال بين PostgreSQL 16 و Apache NetBeans IDE 14 عبر المشغّل PostgreSQL JDBC 42.2.6 [18]، وهو مشغّل JDBC مفتوح المصدر مكتوب بلغة جافا، يتيح للبرامج المطوّرة بلغة جافا الاتصال بقاعدة بيانات PostgreSQL باستخدام شيفرة قياسية ومستقلة عن نوع قاعدة البيانات.أما فيما يتعلّق ببناء النماذج، فقد تم استخدام PyTorch

[21]، في حين تم تطبيق خوارزمية التعلّم المعزّز MCTS بالاعتماد على التنفيذ المتاح في [8]، وذلك باستخدام الإصدار 3.10.12 من لغة البرمجة Python.

أجريت التجارب باستخدام جهاز حاسوب محمول يعمل بنظام تشغيل Linux Mint 21.3 Cinnamon والجدول 5-1 يظهر مواصفات الجهاز.

الجدول 5-1: مواصفات الجهاز الذي أجريت عليه التجارب

الجهاز	المواصفات
نظام التشغيل	Linux Mint 21.3 Cinnamon
إصدار Cinnamon	6.0.4
المعالج	Intel© Core™ i5-5200U CPU @ 2.20GHz × 2
ذاكرة الوصول العشوائي	12 GB
التخزين	SSD 256 GB HDD 1TB
بطاقة الرسوميات	Intel Corporation HD Graphics 5500

1.5 إعداد التجارب:

مجموعة البيانات وأعباء العمل: تم الاعتماد على مجموعة بيانات مفتوحة المصدر وشائعة الاستخدام على نطاق واسع، وهي IMDB، إلى جانب عبء العمل المقابل لها وهو JOB.

إعدادات قاعدة البيانات وبيئة التنفيذ: أجرينا التجارب على نظام إدارة قواعد البيانات PostgreSQL 16، حيث استخدمنا الإضافة pg_hint_plan [19] لتمكين التلميحات (hints) لمُحسِّن الاستعلامات، بهدف تنفيذ خطط استعلام بديلة عن تلك التي يولّدها PostgreSQL

بشكل افتراضي. كما قمنا بتحديد عتبة زمنية قدرها 5 دقائق للاستعلامات التي تستغرق وقتاً طويلاً للتنفيذ و أخيراً قمنا بتقسيم مجموعة استعلامات حمل العمل لدينا إلى ثلاثة مجموعات كالتالي:

- ✓ استعلامات صغيرة: يبلغ عددها 23 استعلام يتراوح عدد جداولها بين 4 و5 جدول.
- ✓ استعلامات متوسطة: يبلغ عددها 39 استعلام يتراوح عدد جداولها بين 6 و7 و8 جدول.
- ✓ استعلامات ضخمة: يبلغ عددها 31 استعلام يتراوح عدد جداولها بين 9 و10 و11 جدول

المقاييس: نُطبّق عدة مقاييس لتقييم الأداء وهي Workload Relative Latency (WRL) و Geometric mean Relative Latency (GMRL) وإجمالي زمن التنفيذ (يشمل على أزمنة توليد الخطط أيضاً) وعدد الخطط ذات أقل زمن تنفيذ.

يمثل زمن الاستجابة النسبي لحمل العمل WRL إجمالي زمن التنفيذ على كامل حمل العمل. في حمل العمل Q، لتكن S_{bi} وS_i الخُطّتين المولّدتين بواسطة مُحسِّن نظام إدارة قواعد البيانات وطريقتنا المقترحة للاستعلام q_i على التوالي، يُحسب WRL كما يلي:

$$\mathbf{WRL}(Q) = \frac{\sum_{q_i \in Q} T_{\mathbf{exe}}(S_i)}{\sum_{q_i \in Q} T_{\mathbf{exe}}(S_{bi})}$$

تمثل القيمة WRL/1 تسريع زمن استجابة حمل العمل بينما يمثل متوسط زمن الاستجابة النسبي الهندسي (GMRL) المتوسط الهندسي لزمن الاستجابة النسبي ضمن حمل العمل، ويُحسب على النحو التالي:

$$\mathbf{GMRL}(Q) = \sqrt[|Q|]{\prod_{q_i \in Q} \frac{T_{\mathbf{exe}}(S_i)}{T_{\mathbf{exe}}(S_{bi})}}$$

تشير الصيغ السابقة إلى أن القيم الدنيا لكل من WRL وGMRL تعكس أداءً أفضل في التحسين. يقيم كل من WRL وGMRL الأداء من منظور مختلف: حيث يبيّن WRL زمن التنفيذ الكلي، مع تركيز أكبر على الاستعلامات البطيئة، بينما يركّز GMRL على تسريع زمن

الاستجابة النسبي على مستوى كل استعلام، موفراً تقييماً عادلاً لكل من الاستعلامات السريعة والبطيئة، إلا أنه غير قادر على قياس التسريع الكلي على مستوى حمل العمل [12].

المقارنة: نقارن نموذجنا المقترح بالطرق التقليدية وأحدث الطرق المتقدمة على النحو التالي:

- ✓ PostgreSQL: نستخدم مُحسِّن الاستعلامات الخاص بـ PostgreSQL (الإصدار 16) بالإعدادات الافتراضية.
- ✓ PG-NoMerge يحسّن أداء PostgreSQL من خلال إلغاء تفعيل المعامل merge join على كامل حمل العمل [12].
- ✓ alphajoin1.0: منهج يستخدم MCTS لتعلّم وتوليد ترتيب عملية الربط بكفاءة عالية داخل مُحسِّن الاستعلامات [8].

طريقة التقييم: نكرر كل تجربة 10 مرات ثم نعرض متوسط أداء المقاييس على مجموعة النتائج التي حصلنا عليها.

2.5 التجربة الأولى:

في هذه التجربة، قمنا بإعادة تحقيق خوارزمية MCTS كما وردت في [8] مع إدخال مجموعة من التعديلات على النحو التالي:

- ✓ تم تعديل خوارزمية UCT في مرحلة اختيار العقدة لتأخذ بعين الاعتبار معلومات عن الاستعلام، حيث أصبح توجيه اختيارات الخوارزمية يتم من خلال إضافة وزن يعتمد على حجم الجداول الموجودة في الاستعلام المعالج، بدلاً من الاعتماد فقط على عدد زيارات العقدة.
- ✓ تم استبدال السياسات العشوائية في مرحلة المحاكاة بسياسات موجهة بمعلومات عن الاستعلام، بحيث أصبحت اختيارات السياسة تُدار وفق حجم الجداول بدلاً من الاعتماد على الاختيارات العشوائية.

بعد الانتهاء من تنفيذ هذه التجربة، قمنا بمقارنة النتائج التي حصلنا عليها مع كل من:

PostgreSQL , PG-NoMerge , Alphajoin1.0

وذلك باستخدام مجموعة المقاييس التالية:

- ✓ عدد الخطط الناتجة من المُحسِّن التي حققت أقل زمن تنفيذ مقارنةً ببقية المُحسّنات

- ✓ WRL و GMRL
- ✓ إجمالي زمن التنفيذ (بما في ذلك أزمنة توليد الخطط)

وأظهرت النتائج ما يلي:

أولاً عدد الخطط المولدة ذات أقل زمن تنفيذ:

الجدول5-2 : متوسط عدد الخطط المولدة ذات أقل زمن تنفيذ عبر عشرة تكرارات عبر مجموعات الاستعلامات

Query Group \ Optimizer	Alphajoin1.0 Vs PostgreSQL	Alphajoin Table Size Vs PostgreSQL	Alphajoin1.0 Vs PG-No Merge	Alphajoin Table Size Vs PG-No Merge
Small Queries	50	52.61	53.04	57.83
Medium Queries	52.31	61.79	62.82	74.1
Large Queries	25.16	38.71	29.03	44.19

يوضح الجدول 5-2 متوسط عدد الخطط المولدة ذات أقل زمن تنفيذ عبر عشرة تكرارات كنسبة مئوية على مجموعات الاستعلامات الصغيرة والمتوسطة والكبيرة وكانت النتائج كالتالي:
نلاحظ أن طريقتنا المقترحة Alphajoin Table Size تحقق أداءً أفضل بشكل عام مقارنةً بـ AlphaJoin 1.0 في جميع السيناريوهات، سواء عند المقارنة مع PostgreSQL أو مع PG-NoMerge.

بالنسبة إلى الاستعلامات الصغيرة، تسجل الطريقة المقترحة تحسناً بنسبة 2.6% مقارنةً بـ AlphaJoin 1.0 على PostgreSQL وبنسبة 4.78% على PG-NoMerge. أما في الاستعلامات المتوسطة، يكون التحسّن أكثر وضوحاً إذ بلغت نسبة التحسين 9.49% مقارنة مع Alphajoin1.0 على PostgreSQL وبمقدار 11.28% على PG-No Merge وبالنسبة إلى الاستعلامات الكبيرة فتستمر الطريقة المقترحة في تحقيق أفضلية واضحة فبلغت نسبة التحسين 13.54% مقارنة مع Alphajoin1.0 على PostgreSQL و15.16% على PG-No Merge. بشكل عام، تؤكد النتائج أن الطريقة المقترحة تؤدي إلى أداء أكثر استقراراً وتحسيناً ملحوظاً عبر مختلف أحجام الاستعلامات، مع أفضلية واضحة عند المقارنة بخطوط الأساس التقليدية (PostgreSQL و PG-NoMerge).

الجدول 5-3: متوسط عدد الخطط ذات أقل زمن تنفيذ عبر عشرة تكرارات

المقارنة	المتوسط
Alphajoin1.0 Vs PostgreSQL	39.7
Alphajoin Table Size Vs PostgreSQL	48.2
Alphajoin Table Size Vs Alphajoin1.0	57.8
PG-No Merge Vs PostgreSQL	7.7
Alphajoin1.0 Vs PG-No Merge	45.7
Alphajoin Table Size Vs PG-No Merge	55.9

يوضح الجدول 5-3 متوسط عدد الخطط الناتجة ذات أقل زمن تنفيذ عبر التكرارات العشرة ونلاحظ التالي:

الطريقة المقترحة هي الأفضل من حيث عدد الخطط المثلى (ذات أقل زمن تنفيذ)، حيث تتفوق على جميع المحسنات الأخرى، خصوصاً عند مقارنتها بـ Alphajoin1.0 و PG-No Merge.

كما أن الطريقة المقترحة تتميز بقدرتها على توليد عدد أكبر من الخطط ذات زمن تنفيذ منخفض، مما يدل على كفاءة عالية في تحسين الأداء بينما Alphajoin1.0 يقدم أداء أفضل من PG-

No Merge لكنه لا يرقى إلى مستوى الطريقة المقترحة أو PostgreSQL. بشكل عام تُظهر النتائج أن الطريقة المقترحة تتفوق بوضوح في اختيار الخطط المثلى، مما يجعلها خياراً واعداً لتحسين أداء أنظمة قواعد البيانات.

ثانياً WRL وGMRL:

الجدول5-4 : متوسط قيم المقياس WRL على مستوى عشرة تكرارات

المقارنة	المتوسط
WRL (.Alphajoin Table Size, PostgreSQL)	0.9
WRL (.Alphajoin Table Size, PG- No Merge)	0.79
WRL (.Alphajoin Table Size, Alphajoin1.0)	0.87

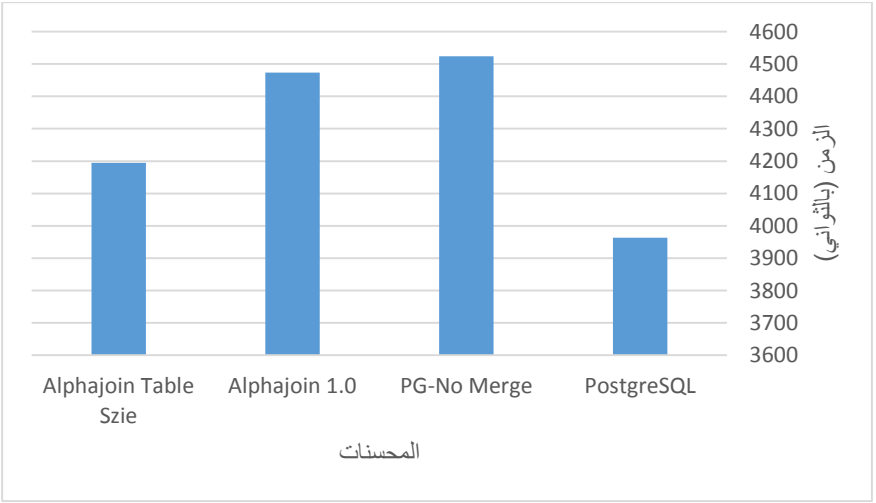
يوضح الجدول 5-4 متوسط قيم المقياس WRL على فترة 10 تكرارات كنسبة مئوية ونلاحظ أن الطريقة المقترحة Alphajoin Table Size تحقق تسريعاً قدره 1.11x مقارنة مع محسن PostgreSQL و1.26x مقارنة مع PG-No Merge و1.14x مقارنة مع alphajoin1.0. بشكل عام تؤكد النتائج السابقة أن الطريقة المقترحة هي الأفضل من حيث تخفيض زمن الاستجابة النسبي.

الجدول 5-5: متوسط قيم المقياس GMRL على مستوى عشرة تكرارات

المقارنة	المتوسط
GMRL (.Alphajoin Table Size, PostgreSQL)	1.15
GMRL (.Alphajoin Table Size, PG- No Merge)	0.96
GMRL (.Alphajoin Table Size, Alphajoin1.0)	0.78

يوضح الجدول 5-5 متوسط قيم المقياس GMRL على فترة 10 تكرارات كنسبة مئوية ونلاحظ أن الطريقة المقترحة Alphajoin Table Size تحقق تسريعاً قدره 0.9x مقارنة مع محسن PostgreSQL (أداء مقارب لـ PostgreSQL) و1.04x مقارنة مع PG-No Merge و1.3x مقارنة مع alphajoin1.0. ويعزى سبب التراجع الطفيف عند المقارنة مع PostgreSQL أن المتوسط النسبي الهندسي حساس لكافة الاستعلامات السريعة والبطيئة ومن الجدول 5-3 فإن PostgreSQL يتفوق بنسبة 44.8 بالمتوسط بعدد الخطط المثلى الناتجة على الطريقة المقترحة (تبقى الطريقة المقترحة أفضل بنسبة 48.2).

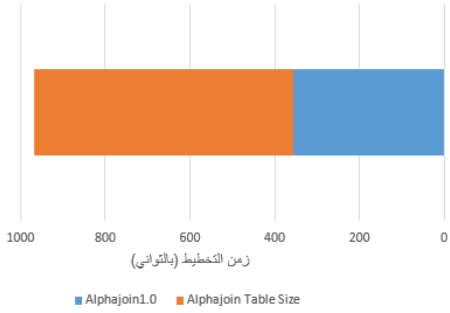
ثالثاً إجمالي زمن التنفيذ:



الشكل 5-1: متوسط إجمالي أزمنة تنفيذ المحسنات عبر عشرة تكرارات (أزمنة التخطيط مضمنة)

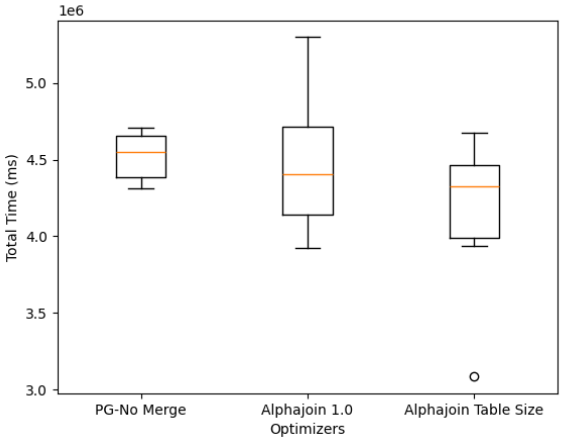
يوضح الشكل 5-1 متوسط إجمالي أزمنة تنفيذ المحسنات بما يشمل أزمنة التخطيط أيضاً عبر عشرة تكرارات ونلاحظ التالي: PostgreSQL يتفوق من حيث سرعة التنفيذ، بينما تقدم الطريقة المقترحة أداء جيد، أما Alphajoin 1.0 سجل متوسط زمن تنفيذ أبطأ من الطريقة المقترحة،

مما يبرز أثر التحسينات الجديدة واخيراً يعاني PG-No Merge من أعلى زمن تنفيذ، ما يعكس ضعفاً واضحاً في الأداء.



الشكل 5-2: متوسط أزمنة التخطيط عبر عشرة تكرارات

يوضح الشكل 5-2 متوسط أزمنة التخطيط عبر عشرة تكرارات ونلاحظ ارتفاع متوسط زمن تخطيط الطريقة المقترحة ويعزى ذلك بسبب أنها تجري حسابات أكثر في عملية اختيار العقد وتنفيذ السياسة الموجهة بمعلومات الاستعلام وبالتالي تجبر عملية البحث على تقييم قرارات (عقد) بشكل أكثر تنظيماً بدلاً من الاعتماد على عمليات المحاكاة العشوائية السريعة ومنه صحيح أننا ندفع كلفة حسابية إضافية لكننا نحصل على توجيه أذكى وبالتالي خطط أسرع.



الشكل 5-3: تقييم استقرار أداء المحسنات

يوضح الشكل 5-3 مخطط Box Plot يقارن إجمالي زمن التنفيذ عبر توزيع القيم وليس فقط المتوسط ونلاحظ التالي: يسجل Alphajoin 1.0 أعلى زمن تنفيذ وأكبر تشتت، مما يدل على عدم استقرار الأداء ووجود استعلامات بطيئة جداً بينما يقدم PG-No Merge أداءً أفضل من Alphajoin 1.0 بينما Alphajoin Table Size هي الأكثر استقراراً والأقل زمناً مما يعكس كفاءة عالية في اختيار الخطط وتنفيذها. الوسيط المنخفض في Alphajoin Table Size يعني أن معظم الاستعلامات تنفذ بسرعة. بشكل عام تؤكد النتائج أن الطريقة المقترحة تتفوق من حيث استقرار الأداء، مقارنة بالمحسنات الأخرى، مما يجعلها خياراً مثالياً في بيئات قواعد البيانات التي تتطلب استجابة سريعة ومتسقة.

3.5 التجربة الثانية:

تتيح دراسات الإزالة Ablation Study فهماً دقيقاً لدور كل مكوّن داخل النظام ومدى مساهمته في الأداء الكلي. ومن خلال إزالة عنصر واحد في كل مرة أو تعطيل وظيفة محددة، يمكن قياس التأثير المباشر لهذا المكوّن على جودة النتائج أو كفاءة التنفيذ مما يعزز موثوقية النموذج عبر سيناريوهات مختلفة. وفي هذا السياق، تهدف التجارب التالية إلى تحليل مكونات الطريقة المقترحة. قمنا في هذه التجربة بدراسة تأثير ما يلي:

- خوارزمية UCT الموجهة بمعلومات خاصة بالاستعلام.
- السياسة الموجهة بمعلومات الاستعلام.

تم إجراء التجربة لكل مكون من المكونات السابقة لخمسة تكرارات مع تحديد قيمة العتبة الزمنية للاستعلامات التي لها زمن تنفيذ طويل بمقدار 3 دقائق.

بعد الانتهاء من تنفيذ التجارب، قمنا بمقارنة النتائج التي حصلنا عليها مع كل من:

PostgreSQL , Alphajoin 1.0

وذلك باستخدام المقاييس التالية:

- عدد الخطط الناتجة من المُحسِّن التي حققت أقل زمن تنفيذ مقارنةً ببقية المُحسّنات.
- إجمالي زمن التنفيذ (بما في ذلك أزمنة توليد الخطط).

دراسة تأثير خوارزمية UCT الموجهة بمعلومات خاصة بالاستعلام:

ملاحظة: سوف نستخدم الاسم Alphajoin UCT-D للإشارة إلى هذه التجربة في النتائج.

في هذه التجربة قمنا بإعادة تحقيق MCTS كما ورد في [8] لكن مع استبدال خوارزمية UCT بالخوارزمية الموجهة بمعلومات الاستعلام. وأظهرت النتائج ما يلي:

أولاً عدد الخطط المولدة ذات أقل زمن تنفيذ:

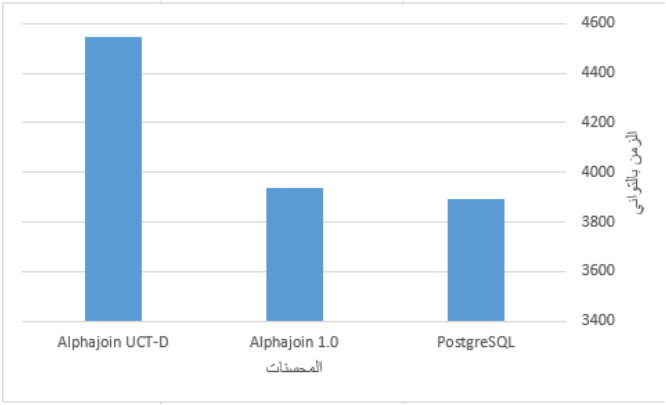
الجدول 5-6: متوسط عدد الخطط ذات أقل زمن تنفيذ عبر خمسة تكرارات

Query Group \ Optimizer	Alphajoin1.0 Vs PostgreSQL	Alphajoin UCT-D Vs PostgreSQL
Small Queries	49	50
Medium Queries	53	51
Large Queries	25	25

يشير الجدول 5-6 إلى متوسط عدد خطط الاستعلامات التي حققت أقل زمن تنفيذ عبر خمسة تكرارات، ومنه نلاحظ التالي: تشير النتائج إلى أن Alphajoin UCT-D قريبة جداً من النسخة الأساسية. بالنسبة للاستعلامات الصغيرة (فضاء البحث لا يزال قابلاً للإدارة)، نلاحظ تحسناً

طفيفاً، يعزى ذلك إلى أن معلومات الاستعلام تساعد عملية البحث على التركيز مبكراً على ترتيبات الربط الواعدة. أما بالنسبة للاستعلامات المتوسطة، نلاحظ انخفاضاً طفيفاً والسبب هو أنه مع ازدياد عدد الجداول، يتسع نطاق الخطة بسرعة كبيرة، وقد تصبح التوجيهات أقل دقة، ففي نطاق بحث أكبر، تبدأ Alphajoin UCT-D في تفويت بعض المناطق المثلى من فضاء البحث، في حين قد تستكشف Alphajoin1.0 ذات العشوائية الأكبر أحياناً فروعاً أكثر تنوعاً وتصل بالصدفة إلى خطط أكثر مثالية.

ثانياً إجمالي زمن التنفيذ:



الشكل 5-4: متوسط إجمالي أزمنة تنفيذ المحسنات عبر خمسة تكرارات (أزمنة التخطيط مضمنة)

يوضح الشكل 5-4 متوسط أزمنة تنفيذ المحسنات عبر خمسة تكرارات ونلاحظ أن الطريقة المقترحة تسجل أعلى متوسط زمن تنفيذ ويعزى ذلك إلى الاعتماد على العشوائية في عمليات المحاكاة مما يؤدي إلى التأثير على اختيارات الطريقة المقترحة وبالتالي توليد خطط سيئة.

دراسة تأثير السياسة الموجهة بمعلومات الاستعلام:

ملاحظة: سوف نستخدم الاسم Alphajoin P-D للإشارة إلى هذه التجربة في النتائج.

في هذه التجربة قمنا بإعادة تحقيق MCTS كما ورد في [8] لكن مع استبدال السياسة العشوائية بأخرى موجهة بمعلومات الاستعلام. وأظهرت النتائج ما يلي:

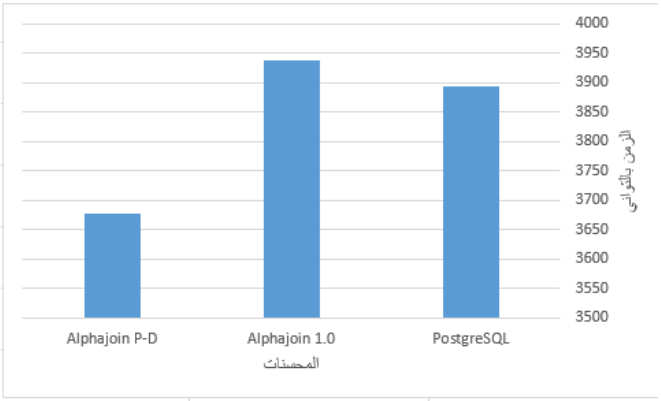
أولاً عدد الخطط المولدة ذات أقل زمن تنفيذ:

الجدول 5-7: متوسط عدد الخطط ذات أقل زمن تنفيذ عبر خمسة تكرارات

Query Group \ Optimizer	Alphajoin1.0 Vs PostgreSQL	Alphajoin P-D Vs PostgreSQL
Small Queries	49	53
Medium Queries	53	56
Large Queries	27	32

يشير الجدول 5-7 إلى متوسط عدد خطط الاستعلامات التي حققت أقل زمن تنفيذ عبر خمسة تكرارات، وذلك في سياقات مقارنة متعددة بين المحسنات المختلفة وتشير النتائج إلى تفوق Alphajoin P-D على Alphajoin1.0 في جميع المقارنات، مما يثبت أن السياسة الموجهة بمعلومات الاستعلام أفضل من السياسة العشوائية كما أنها توجه عملية البحث نحو أجزاء واعدة من فضاء البحث. ومنه نجد أن الطريقة المقترحة قادرة على التكيف مع التوزيع المتفاوت لحجم البيانات وتعقيد الربط.

ثانياً إجمالي زمن التنفيذ:



الشكل 5-5: متوسط إجمالي أزمنة تنفيذ المحسنات عبر خمسة تكرارات (أزمنة التخطيط مضمنة)

يوضح الشكل 5-5 متوسط أزمنة تنفيذ المحسنات عبر خمسة تكرارات ونلاحظ أن الطريقة المقترحة تسجل أقل متوسط زمن تنفيذ بين المحسنات محل المقارنة مما يعكس مدى تفوق السياسة الموجهة بمعلومات الاستعلام على توليد خطط استعلامات فعالة ذات أقل زمن تنفيذ. بشكل عام نستنتج أن الطريقة المقترحة أسرع، وتحقق أداءً أفضل من كل البدائل عبر جميع المقارنات.

6. الاستنتاجات والتوصيات:

تُبرز هذه الدراسة، استناداً إلى مراجعة منهجية وشاملة للأدبيات ذات الصلة، التوجّه المتنامي نحو توظيف تقنيات التعلّم الآلي في تحسين كفاءة وأداء أنظمة قواعد البيانات. وقد برزت تقنيات التعلم المعزز بوصفه أحد الأساليب الواعدة في تحسين ترتيب عمليات الربط، في ظل تزايد الاهتمام بتطبيق هذه التقنيات لمعالجة التحديات المتصاعدة التي تواجهها أنظمة إدارة قواعد البيانات الحديثة.

اعتمدت هذه الدراسة على توظيف خوارزمية MCTS بالتكامل مع معلومات خاصة بالاستعلام، بما يمثّل توجهاً منهجياً نحو تحسين ترتيب عمليات الربط وتعزيز دقّة اتخاذ القرار في توليد خطط تنفيذ فعّالة زمنياً. أظهرت النتائج التجريبية المحققة ضمن إطار العمل المقترح تحسّناً ملحوظاً في جودة خطط الاستعلامات المولّدة، الأمر الذي انعكس مباشرة على تقليص أزمنة التنفيذ. وعلى نحو أكثر تحديداً، حققت المنهجية المقترحة انخفاضاً ملموساً في زمن تنفيذ الاستعلامات، إلى جانب تسريع ملحوظ وتحسين واضح في أداء عبء العمل وزيادة في عدد الخطط الأمثلية الناتجة، مما ينعكس إيجاباً على الاداء العام.

تتجسّد مساهمات هذه الدراسة في محورين رئيسين:

- ✓ توجيه خوارزمية UCT بمعلومات بنيوية عن الاستعلام: إذ أتاح دمج المعرفة ببنية الاستعلام تعزيز فهم الخوارزمية لسياق الاستعلام المعالج، والحدّ من مستويات عدم اليقين الملازمة للخوارزمية.
- ✓ اعتماد سياسة موجهة بطبيعة الاستعلام بدلاً من السياسات العشوائية: لأن استكشاف فضاء البحث الضخم بشكل عشوائي غير فعال يؤدي إلى ترتيبات ربط سيئة وبالتالي الحصول على خطط استعلام كارثية، بينما يضمن التوجيه المستند إلى خصائص الاستعلام مساراً أكثر فاعلية.

7. الأعمال المستقبلية:

استناداً إلى نتائج الدراسة الحالية وما تم تحقيقه من تقدم، يمكن توسيع نطاق العمل البحثي مستقبلاً عبر عدد من المسارات الواعدة، من أبرزها:

- ✓ دراسة إمكانية دمج طريقتنا المقترحة في مُحسِّن قواعد البيانات التقليدي، مما يُتيح لتقنيات التحسين التقليدية الاستفادة من التعلّم الآلي.
- ✓ تسريع عملية التعلم من خلال دمج المعرفة المكتسبة من الخبراء في مرحلة التدريب نظراً لما تتطلبه بعض النماذج من موارد حسابية كبيرة وزمن تدريب طويل عند بنائها من الصفر.
- ✓ دراسة إمكانية مكاملة المزيد من المعلومات الخاصة بالاستعلام بخوارزمية MCTS بما يسهم في رفع كفاءة الخوارزمية وتحسين قدرتها على اتخاذ القرارات المثلى مما ينعكس إيجاباً على أزمنة تنفيذ الاستعلامات.

تلك المقترحات يمكن أن تسهم في تحسين أداء خوارزميات تحسين ترتيب عمليات الربط وزيادة كفاءتها في التعامل مع أعباء عمل متنوعة وقواعد بيانات مختلفة.

8. المراجع:

[1] Vidhya, V., Jeyaram, G. and Ishwarya, K.R, 2016- Database Management Systems. Alpha Science International.

[2] Heitz J, Stockinger K. Join query optimization with deep reinforcement learning algorithms. arXiv preprint arXiv:1911.11689. 2019 Nov 26.

[3] Li G, Zhou X, Cao L. AI meets database: AI4DB and DB4AI. InProceedings of the 2021 international conference on management of data 2021 Jun 9 (pp. 2859-2866).

[4] Sutton R., Barto, A., 2014- Reinforcement Learning: An Introduction. IEEE Transactions on Neural Networks, Vol. 9, No. 5, 1054.

[5] Hettiarachchi N, Yapa P. Advancements in SQL Query Optimization: A Review of Join Order and Index Selection. In2025 5th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE) 2025 Jun 13 (pp. 31-39). IEEE.

[6] Elmasri R, Navathe SB. Fundamentals of database systems seventh edition.

[7] Mazyavkina N, Sviridov S, Ivanov S, Burnaev E. Reinforcement learning for combinatorial optimization: A survey. Computers & Operations Research. 2021 Oct 1;134:105400.

[8] Zhang J, Zhou K, Schelter S. AlphaJoin: Join Order Selection à la AlphaGo. PhD@ VLDB. 2020 Aug 31;2652.

[9] Marcus R, Papaemmanouil O. Deep reinforcement learning for join order enumeration. InProceedings of the First International Workshop on Exploiting Artificial Intelligence Techniques for Data Management 2018 Jun 10 (pp. 1-4).

[10] Yu X, Li G, Chai C, Tang N. Reinforcement learning with tree-lstm for join order selection. In2020 IEEE 36th international conference on data engineering (ICDE) 2020 Apr 20 (pp. 1297-1308). IEEE.

[11] Xu X, Zhao Z, Zhang T, Kang R, Sun L, Chen J. Coool: A learning-to-rank approach for sql hint recommendations. arXiv preprint arXiv:2304.04407. 2023 Apr 10.

[12] Chen T, Gao J, Chen H, Tu Y. Loger: A learned optimizer towards generating efficient and robust query execution plans. Proceedings of the VLDB Endowment. 2023 Mar 1;16(7):1777-89.

[13] Liu C, Kamali A, Kantere V, Zuzarte C, Corvinelli V. A Novel Framework Using Deep Reinforcement Learning for Join Order Selection. arXiv preprint arXiv:2412.10253. 2024 Dec 13.

[14] Kemmerling M, Lütticke D, Schmitt RH. Beyond games: a systematic review of neural Monte Carlo tree search applications. arXiv preprint arXiv:2303.08060. 2023 Mar 14.

[15] Leis V, Radke B, Gubichev A, Mirchev A, Boncz P, Kemper A, Neumann T. Query optimization through the looking glass, and what we found running the join order benchmark. The VLDB Journal. 2018 Oct;27(5):643-68.

[16] Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning. Cambridge: MIT press; 2016 Nov 18.

[17] https://www.postgresql.org/about/. Last Visited 29/1/2026

[18] https://jdbc.postgresql.org/. Last Visited 29/1/2026

[19] https://pg-hint-plan.readthedocs.io/en/latest/description.html. Last Visited 29/1/2026

[20] https://pytorch.org/. Last Visited 29/1/2026

[21] https://pypi.org/project/torch/. Last Visited 29/1/2026

[22] Marcus R, Papaemmanouil O. Towards a hands-free query optimizer through deep learning. arXiv preprint arXiv:1809.10212. 2018 Sep 26.

[23] Browne CB, Powley E, Whitehouse D, Lucas SM, Cowling PI, Rohlfshagen P, Tavener S, Perez D, Samothrakis S, Colton S. A survey of monte carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in games. 2012 Feb 3;4(1):1-43.

تحسين الصيانة التنبؤية لآلات التفريز باستخدام هندسة الميزات ونماذج تعلّم الآلة الهجينة

إعداد : صفاء مصطفى محسن إشراف: د. حسام الوفائي

الملخص:

يركّز هذا البحث على تطوير نموذج صيانة تنبؤية يعتمد على تقنيات تعلّم الآلة بهدف تحسين موثوقية وكفاءة آلات التفريز في البيئات الصناعية الحديثة، تعتمد المنهجية على تحليل بيانات تشغيلية معيارية من قاعدة AI4I2020، ومعالجة التحديات المرتبطة بالصيانة التقليدية مثل التوقفات المفاجئة وارتفاع تكاليف التشغيل. يتضمن النموذج المقترح مراحل تنظيف البيانات واستخراج ميزات مشتقة تعبّر عن السلوك التشغيلي للآلة، ثم تدريب نماذج تعلم الآلة مثل Random Forest وXGBoost وGradient Boosting، إضافة إلى تصميم شبكة عصبونية مخصّصة للتعامل مع الأعطال النادرة.

تُظهر النتائج أن استخدام الميزات المشتقة حسّن أداء النماذج بشكل ملحوظ، خصوصاً في ، مقارنة بالاعتماد على الميزات الخام فقط، كما حققت F1-Scoreو Recall مؤشرات النماذج الشجرية التجميعية أداءً قوياً في تصنيف الأعطال الشائعة، بينما تطلّب اكتشاف استخدام تقنيات معالجة عدم التوازن مما رفع قدرة النموذج على TWF الأعطال النادرة مثل التعرّف على هذه الحالات رغم محدودية عيناتها، وتؤكد الدراسة أن دمج هندسة الميزات مع نماذج هجينة مخصّصة للفئات النادرة يوفّر إطاراً فعّالاً لبناء نظام صيانة تنبؤية قابل للتطبيق الصناعي.

كلمات مفتاحية: الصيانة التنبؤية، تعلّم الآلة، آلات التفريز، ميزات مشتقة، AI4I 2020.

Enhancing Predictive Maintenance of Milling Machines Using Feature

Abstract:

This research focuses on developing a predictive maintenance model based on machine learning techniques to enhance the reliability and operational efficiency of milling machines in modern industrial environments. The methodology relies on analyzing standardized operational data from the AI4I 2020 dataset and addressing challenges associated with traditional maintenance approaches, such as unexpected failures and high operational costs. The proposed model includes data cleaning, the extraction of engineered features that capture the machine's operational behavior, and the training of machine learning algorithms such as Random Forest, XGBoost, and Gradient Boosting, in addition to designing a neural network specifically tailored to handle rare failure cases.

The results show that incorporating engineered features significantly improved model performance, particularly in Recall and F1-Score metrics, compared to relying solely on raw features. Ensemble tree-based models demonstrated strong performance in classifying common failure modes, while detecting rare failures such as TWF required specialized imbalance-handling techniques, including RandomOverSampler, Focal Loss, and pos_weight, which enhanced the model's ability to identify these infrequent cases despite their limited representation.

The study confirms that combining feature engineering with hybrid models tailored for rare classes provides an effective framework for building an industrially applicable predictive maintenance system. It further recommends developing continuous monitoring mechanisms for the model, improving data quality, and expanding datasets related to rare failures to strengthen system reliability and reduce unplanned downtime**.**

Keywords: Predictive Maintenance, Machine Learning, Milling Machines, Engineered Features, AI4I 2020.

مقدمة:

تُعدّ الصناعات التحويلية من أهم القطاعات الداعمة للنمو الاقتصادي العالمي، نظراً لاعتمادها على منظومات إنتاجية متقدمة قادرة على تحويل المواد الخام إلى منتجات عالية القيمة، وقد أكدت الدراسات الحديثة أن التطور في تقنيات التصنيع الدقيقة أصبح عنصراً محورياً في تعزيز القدرة التنافسية للدول الصناعية، خاصة مع تزايد الطلب على منتجات ذات جودة عالية ومعايير هندسية صارمة [1].

وتُعد عمليات التشغيل الميكانيكي، وعلى رأسها التفريز، من الركائز الأساسية في تصنيع المكوّنات الدقيقة المستخدمة في قطاعات الطيران والسيارات والروبوتات والأجهزة الطبية.

تحتل آلات التفريز مكانة متقدمة بين آلات التشغيل، لما تتمتع به من مرونة عالية وقدرة على تشكيل المواد الصلبة وإنتاج أشكال هندسية معقدة، وتشير الدراسات الحديثة إلى أن التطور في أنظمة التحكم، وتحسين ديناميكية الماكينات، واعتماد تقنيات المراقبة الذكية، جعل آلات التفريز أكثر موثوقية وكفاءة في بيئات التصنيع المتقدمة [1].

ومع تزايد تعقيد الأنظمة الإنتاجية، أصبحت الأعطال المفاجئة لآلات التشغيل من أبرز التحديات التي تواجه المصانع الحديثة، لما تسببه من خسائر مالية كبيرة وتوقفات غير مخططة، وقد بينت دراسات صناعية حديثة أن التوقفات غير المخططة تمثل ما بين 15% و25% من إجمالي خسائر الإنتاج، مما يجعل إدارة الأعطال وتحسين موثوقية المعدات ضرورة استراتيجية [2].

وفي ظل هذا الواقع، برزت الصيانة التنبؤية كأحد أهم الاتجاهات الحديثة في الصناعة، نظراً لقدرتها على الانتقال من مفهوم الصيانة التقليدية القائمة على الجداول الزمنية أو التدخل بعد حدوث العطل، إلى نموذج أكثر ذكاءً يعتمد على تحليل البيانات التشغيلية للآلة بهدف التنبؤ بالأعطال قبل وقوعها.

كما أثبتت الدراسات الحديثة أن دمج الصيانة التنبؤية مع تقنيات تعلّم الآلة وإنترنت الأشياء الصناعي يعزز دقة التنبؤ بالأعطال ويجعل أنظمة الإنتاج أكثر كفاءة واستجابة لمتطلبات الصناعة الحديثة.

أن الصيانة التنبؤية تسهم في تقليل التوقفات غير المخططة بنسبة قد تصل إلى 40%، وتحسين كفاءة التشغيل بشكل كبير [3].

أثبتت الدراسات الحديثة أن خوارزميات تعلم الآلة تُظهر أداءً عالياً في مهام تشخيص الأعطال، خاصة في بيئات التشغيل المعقدة مثل آلات التفريز، كما أن دمج تقنيات تعلّم الآلة مع بيانات الاستشعار اللحظية يساهم في بناء نماذج تنبؤية أكثر دقة وموثوقية، مما يعزز من فعالية أنظمة الصيانة التنبؤية في الصناعة الحديثة [4].

وانطلاقاً من هذه التطورات، يهدف هذا البحث إلى توظيف تقنيات تعلّم الآلة في تحليل بيانات التشغيل الخاصة بآلات التفريز، بهدف التنبؤ بالأعطال وتحسين موثوقية الأنظمة الإنتاجية، بما يتماشى مع توجهات الصناعة الحديثة نحو التحول الرقمي ورفع كفاءة التشغيل.

تنظيم البحث:

تم تنظيم البحث وفق الترتيب التالي: تم البدء بمقدمة عن العمل المنجز ثم تم ذكر الهدف من البحث وطرائقه، وبعدها تم إجراء دراسة مرجعية في الفقرة 1 تستعرض أحدث الأبحاث التي تم إنجازها عن الصيانة التنبؤية، ثم تم في الفقرة 2 شرحاً تفصيلياً لآلية بناء النموذج المقترح، بدءاً من معالجة البيانات وتنظيفها واستخراج الميزات المشتقة، وصولاً إلى تدريب نماذج تعلّم الآلة المختلفة مثل Random Forest وXGBoost وGradient Boosting، إضافة إلى تصميم شبكة عصبونية مخصّصة للتعامل مع الأعطال النادرة وأخيراً تم مناقشة النتائج التي تم الحصول عليها في الفقرة 3 وتحليلها مع مقارنة هذه النتائج مع الأبحاث السابقة في الفقرة 4.

هدف البحث:

يهدف هذا البحث إلى تطوير نموذج صيانة تنبؤية يعتمد على تقنيات تعلّم الآلة والشبكات العصبونية للتنبؤ بالحالات التشغيلية والأعطال المحتملة في آلات التفريز، وذلك من خلال تحليل بيانات قاعدة AI4I 2020 وتقييم أداء النماذج المقترحة وملاءمتها لمتطلبات البيئات الصناعية الحديثة.

مواد وطرق البحث:

يعتمد هذا البحث على تحليل بيانات تشغيلية معيارية من قاعدة AI4I 2020 باستخدام منهجيات تعلم الآلة، تمت معالجة البيانات واستخراج الميزات المشتقة وتدريب النماذج تمّ عبر Jupyter Notebook مع بيئة البرمجة Python ومكتباتها المتقدمة مثل Pandas و Scikit-Learnو PyTorch.

1 دراسة مرجعية

تناولت العديد من الدراسات الحديثة موضوع الصيانة التنبؤية باستخدام تقنيات تعلّم الآلة، وقد اختلفت هذه الدراسات في نوع البيانات المستخدمة، الخوارزميات، ومستوى النضج التطبيقي.

حيث قام (Akyaz, T., & Engin) في [5] بتقديم دراسة لنظام صيانة تنبؤية لآلات الغزل الصناعية باستخدام خوارزميات مثل Random Forest وXGBoost وSVM والشبكات العصبونية، وتمثلت قوة الدراسة في وجود إطار تطبيقي متكامل يربط بين جمع البيانات وتحليلها وتنفيذ النموذج في بيئة صناعية فعلية، إضافة إلى استخدام مجموعة واسعة من الميزات التشغيلية، إلا أن الدراسة لم تقدّم معالجة متقدمة لمشكلة عدم توازن البيانات، كما لم تتضمن هندسة ميزات مشتقة أو تحليلاً تفسيرياً عميقاً للمتغيرات المؤثرة في الأعطال.

بينما استخدم (Baradaran, A. H, 2025) في [6] مجموعة من نماذج التعلم الخاضع للإشراف مثل Naive Bayes وSVM والانحدار اللوجستي وRandom Forest لتشخيص حالة المحركات الكهربائية، وتميزت الدراسة بشمولية المقارنة بين عدة خوارزميات تقليدية وسهولة تفسير نتائجها، إلا أنها اعتمدت على نماذج تقليدية فقط دون استخدام شبكات عصبونية أو نماذج هجينة، كما لم تعالج مشكلة عدم توازن البيانات، مما قلل من قدرتها على اكتشاف الأعطال النادرة، إضافة إلى غياب إطار تطبيقي متكامل.

أما (Waghulde et al, 2025) في [7] فقد ركز على تقييم أداء مجموعة واسعة من نماذج تعلم الآلة والتعلم العميق باستخدام قاعدة بيانات AI4I 2020 ، بما في ذلك Random Forest وSVM وXGBoost والشبكات العصبونية، وتمثلت قوة الدراسة في شمولية المقارنة وتغطية عدد كبير من الخوارزميات، إضافة إلى استخدام مقاييس أداء متعددة، ومع ذلك، اقتصرت الدراسة على التقييم فقط دون تقديم إطار تطبيقي متكامل، كما لم تتناول هندسة ميزات مشتقة أو تقنيات متقدمة لمعالجة الأعطال النادرة، مما يجعل نتائجها أقل قابلية للتطبيق الصناعي المباشر.

بينما (Zhao et al, 2025) في [1] قام بتطوير نموذج صيانة تنبؤية لآلة التفريز باستخدام خوارزميات تعلم الآلة مدعومة بتحليل SHAP لتفسير قرارات النموذج، وتمثلت قوة الدراسة في قدرتها على تحديد المتغيرات الأكثر تأثيراً في حدوث الأعطال، مما يعزز قابلية تفسير النموذج، إلا أن الدراسة اعتمدت على الميزات الخام فقط دون هندسة ميزات مشتقة، كما لم تقدّم معالجة كافية لمشكلة عدم توازن البيانات، مما يحدّ من قدرتها على اكتشاف الأعطال النادرة بدقة.

وتُظهر هذه الدراسات مجتمعة أهمية توظيف تقنيات تعلم الآلة في بناء أنظمة صيانة تنبؤية فعّالة، إلا أنها تختلف في منهجياتها من حيث نوع البيانات، الخوارزميات المستخدمة، وعمق معالجة مشكلة عدم التوازن.

كما أن معظمها يعتمد على الميزات الخام دون هندسة ميزات مشتقة أو نماذج مخصصة للأعطال النادرة، وهي فجوات يسعى هذا البحث إلى معالجتها من خلال تطوير نموذج أكثر

تكاملاً يعتمد على هندسة الميزات وتقنيات معالجة عدم التوازن ضمن إطار تطبيقي مخصص لآلات التفريز.

2 بناء النموذج التنبؤي المقترح:

2.1 البيانات المستخدمة كأساس لبناء النموذج المقترح:

اعتمدت هذه الدراسة على قاعدة البيانات المرجعية AI4I 2020 بوصفها المصدر الرئيسي لبناء وتدريب النموذج التنبؤي المقترح، نظراً لكونها من أكثر قواعد البيانات استخداماً في أبحاث الصيانة التنبؤية.

تُعد من قواعد البيانات الصناعية المصممة لتمثيل السلوك التشغيلي لآلات التفريز تحت ظروف تشغيل مختلفة، بما في ذلك الحالات الطبيعية وحالات التدهور والأعطال [8].

وتتألف القاعدة من نحو 10,000 عينة، وهو حجم مناسب لتدريب نماذج تعلّم الآلة دون التعرض لمشكلة فرط التعميم، تتضمن قاعدة البيانات:

- UDI : رقم تعريف فريد لكل سجل
- Product ID : يتضمن رمز جودة المنتج
- Type : نوع المنتج
- Air temperature [K] : درجة حرارة الهواء
- Process temperature [K] : درجة حرارة العملية
- Rotational speed [rpm]: سرعة الدوران
- Torque [Nm]: العزم
- Tool wear [min] : تآكل الأداة
- TWF (Tool Wear Failure) : عطل ناتج عن تآكل الأداة

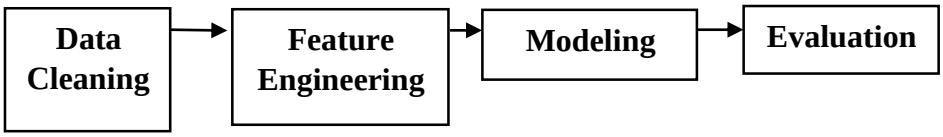
- HDF (Heat Dissipation Failure) : عطل في تبديد الحرارة
- PWF (Power Failure) : عطل في القدرة الكهربائية
- OSF (Overstrain Failure) : عطل بسبب الإجهاد الزائد
- RNF (Random Failure) : عطل عشوائي
- Machine failure: عطل الآلة

يُعد التوزيع غير المتوازن بين الفئات أحد أبرز التحديات في هذه القاعدة، حيث تشكل الحالات السليمة النسبة الأكبر من البيانات، بينما تظهر الأعطال النادرة وخاصة TWF بنسب منخفضة جداً مقارنة ببقية الفئات. يؤدي هذا التفاوت إلى انحياز النماذج نحو الفئات الشائعة، ويضعف قدرتها على اكتشاف الأعطال النادرة [9].

وبناءً على ما سبق، تشكل قاعدة AI4I 2020 أساساً قوياً لتطوير نموذج صيانة تنبؤية يعتمد على هندسة الميزات ومعالجة عدم التوازن، وهو ما سيتم تفصيله في القسم التالي ضمن منهجية بناء النموذج المقترح.

2.2 المنهجية العملية لتصميم نموذج الصيانة التنبؤية:

يوضّح شكل 1 مخطط سير العمل المعتمد في بناء النموذج التنبؤي المستخدم في هذه الدراسة، والذي يتكوّن من أربع مراحل رئيسية: تنظيف البيانات، واستخراج الميزات، وبناء النماذج، ثم التقييم.



شكل 1 المنهجية العملية لتصميم نموذج الصيانة التنبؤية

- معالجة البيانات وتنظيفها

تُعدّ معالجة البيانات خطوة أساسية في بناء أي نموذج يعتمد على تقنيات تعلّم الآلة، إذ تؤثر جودة البيانات وتنظيمها بشكل مباشر في دقة النموذج وقدرته على التعميم، وبناءً على ذلك، تم اتباع منهجية منهجية دقيقة لمعالجة البيانات التشغيلية المستمدة من قاعدة AI4I 2020 قبل البدء بمرحلة التدريب.

بدأت عملية المعالجة بإجراء تحليل أولي للبيانات بهدف فهم طبيعتها وتوزيعها واكتشاف أي قيم شاذة أو غير منطقية قد تؤثر في أداء النماذج.

بعد ذلك، تم تنفيذ ترميز للمتغيرات الفئوية مثل نوع المنتج باستخدام أسلوب الترميز العددي، وذلك لضمان إمكانية استخدامها ضمن نماذج تعلّم الآلة التي تتطلب مدخلات رقمية، كما تم تطبيق التطبيع بهدف تقليل الفروقات الكبيرة في نطاق القيم وتحسين استقرار عملية التدريب، خاصة في النماذج الحساسة لمقاييس البيانات مثل الشبكات العصبونية، وفي خطوة لاحقة، تم تقسيم البيانات إلى مجموعتي تدريب واختبار بنسبة مناسبة لضمان تقييم موضوعي لأداء النماذج [10].

أسهمت هذه الخطوات في تجهيز البيانات بصورة مثالية لبناء النموذج التنبؤي، وضمان أن تكون المدخلات نظيفة، متسقة، ومهيأة لعمليات هندسة الميزات ومعالجة عدم التوازن التي تشكل المرحلة التالية في منهجية بناء النموذج المقترح.

- هندسة الميزات

تُعدّ هندسة الميزات خطوة محورية في بناء النموذج التنبؤي، إذ تعتمد دقة النماذج وفعاليتها على جودة التمثيل الرقمي للسلوك التشغيلي للآلة، فالاعتماد على الميزات الخام وحدها غالباً لا يكفي لالتقاط الأنماط الدقيقة المرتبطة بالأعطال، خصوصاً الأعطال النادرة التي تتطلب توصيفاً

أعمق للسلوك الديناميكي للآلة، ومن هذا المنطلق، تم في هذه الدراسة تطوير مجموعة من الميزات المشتقة التي تعزز قدرة النماذج على التمييز بين الحالات السليمة وحالات العطل.

تمثل هذه الميزات المشتقة أحد العناصر الأساسية في أنظمة الصيانة التنبؤية، إذ تُستخلص من المتغيرات التشغيلية اعتماداً على العلاقات الفيزيائية والوظيفية المرتبطة بأنماط الأعطال المختلفة، وتتيح هذه الميزات توصيفاً كمياً أكثر دقة للحالة الصحية لآلة التفريز تحت ظروف تشغيل وتحميل متباينة، مما يسهم في تحسين قدرة النماذج على اكتشاف التدهور المبكر والتنبؤ بالأعطال قبل وقوعها [11].

- تدريب نماذج تعلّم الآلة

بعد تجهيز البيانات ومعالجتها واستخراج الميزات المشتقة ومعالجة مشكلة عدم التوازن، تم الانتقال إلى مرحلة تدريب نماذج تعلّم الآلة بهدف بناء نظام قادر على التنبؤ بالأعطال التشغيلية في آلات التفريز بدقة وموثوقية.

وقد تم اختيار مجموعة من الخوارزميات الشجرية التجميعية وخوارزميات التعزيز التدريجي نظراً لقدرتها العالية على التعامل مع البيانات الصناعية المعقدة، إضافة إلى أدائها المتميز في مهام التصنيف متعددة الفئات [12].

- تصميم الشبكة العصبونية المخصّصة للأعطال النادرة

نظراً للتحديات التي تواجه نماذج تعلّم الآلة التقليدية في اكتشاف الأعطال النادرة، تم في هذه الدراسة تصميم شبكة عصبونية اصطناعية مخصّصة لمعالجة هذه الفئة من الأعطال [13].

تتكون الشبكة لدينا من:

❖ طبقة الإدخال(Input Layer)

هي الطبقة التي تســتقبل جميع الميزات بعد تجهيز البيانات، ويكون عدد خلاياها مســاوياً لعدد الميزات.

❖ الطبقة المخفية الأولى (Hidden Layer 1 — 64 Neurons)

هي طبقة كثيفة تحتوي على 64 خلية عصبونية، وتعمل على تعلّم العلاقات الواسعة بين الميزات والتقاط الأنماط غير الخطية وبناء تمثيل أولي للبيانات.

❖ الطبقة المخفية الثانية (Hidden Layer 2 — 32 Neurons)

تتكون من 32 خلية عصبونية، وتهدف إلى ضغط المعلومات وتقليل الضجيج واستخراج الأنماط الأكثر أهمية [14].

❖ الطبقة المخفية الثالثة (Hidden Layer 3 — 16 Neurons)

تضم هذه الطبقة 16 خلية عصبونية، وتعمل على تعلّم الأنماط الدقيقة المرتبطة بالعطل والتقاط العلاقات الصــــغيرة التي قد لا تظهر في الطبقات الســـابقة، كما تهيئ البيانات لطبقة الإخراج باعتبارها مرحلة فلترة نهائية قبل اتخاذ القرار.

وتعتمد جميع الطبقات المخفية على دالة التنشـــيط ReLU، نظراً لفعاليتها العالية في تمثيل العلاقات غير الخطية وسرعة حسابها [15].

❖ طبقة الإخراج (Output Layer)

يختلف تصميمها بحسب نوع المهمة، ففي حالة التصنيف الثنائي (1/0) تُستخدم خلية واحدة مع دالة التنشيط Sigmoid التي تحوّل المخرجات إلى قيمة بين 0 و1، مما يجعلها مناسبة للنماذج الاحتمالية التي تهدف إلى تحديد احتمال حدوث العطل.

أما في حالة التصنيف متعدد الفئات (0–5) فتتكون طبقة الإخراج من ست خلايا، وتُستخدم دالة التنشيط Softmax التي تحوّل المخرجات إلى توزيع احتمالي يمثّل احتمال انتماء العينة إلى كل فئة، بحيث يكون مجموع الاحتمالات مساوياً لـ 1 [16].

وتتمثل الوظيفة الأساسية لطبقة الإخراج في إنتاج الاحتمال النهائي للعطل أو تحديد نوع العطل، وفقاً لطبيعة المهمة المطلوبة في نظام الصيانة التنبؤية.

يمثل هذا النموذج العصبي خطوة محورية في تطوير إطار صيانة تنبؤية متكامل، إذ يجمع بين قوة النماذج الشجرية في اكتشاف الأعطال الشائعة ومرونة الشبكات العصبونية في التعامل مع الأعطال النادرة، مما يعزز موثوقية النظام التنبؤي ويجعله مناسباً للتطبيق في البيئات الصناعية الحديثة.

3 النتائج

3.1 المرحلة الأولى: الكشف عن العطل

في المرحلة الأولى من النظام المقترح، تم التركيز على مهمة الكشف عن العطل في الآلة.

3.1.1 دراسة عطل الآلة باستخدام الميزات الخام

تم تدريب النماذج على متغير Machine Failure باعتباره المتغير الهدف وذلك بهدف التمييز بين حالتي "وجود عطل" و"عدم وجود عطل" في الآلة، تم تطبيق مجموعة من خوارزميات تعلّم الآلة على الميزات الخام لتقييم قدرتها على اكتشاف العطل بشكل أولي.

يوضح جدول 1 تقييم أداء هذه النماذج باستخدام مجموعة من المقاييس تشمل (Accuracy)، (Recall)، (Precision)، ومقياسF1 ، بالإضافة إلى تحليل مصفوفة الالتباس (Confusion Matrix).

وقد ساعد هذا التنوع في النماذج على فهم سلوك البيانات بشكل أعمق وتحديد النموذج الأنسب للانتقال إلى المرحلة الثانية، نتج مايلي:

جدول 1 تقييم أداء خوارزميات تعلم الآلة باستخدام الميزات الخام

Model	Accuracy	Recall	Precision	F1-Score
XGBoost Baseline	0.9875	0.6912	0.9216	0.7899
Gradient Boosting Baseline	0.9835	0.5882	0.8889	0.7080
Random Forest Baseline	0.9825	0.5441	0.9024	0.6789
Neural Network (MLP 64→32→16)	0.9785	0.5588	0.7451	0.6387
Decision Tree Baseline	0.9785	0.4706	0.8205	0.5981
SVM Baseline	0.9200	0.8529	0.2788	0.4203
Logistic Regression Baseline	0.8265	0.7794	0.1377	0.2340

على الرغم من أن جميع النماذج المحسوبة باستخدام الميزات الخام حققت مستويات مرتفعة من الدقة، إلا أنّ قيم كل من Recall و Precision كانت منخفضة نسبياً، مما يشير إلى ضعف قدرة النماذج على اكتشاف حالات العطل بشكل صحيح وتمييزها عن الحالات الطبيعية، وتبرز هذه النتائج الحاجة إلى إدخال ميزات مشتقة أكثر تعبيراً عن سلوك الآلة، بهدف تحسين قدرة النماذج على التقاط الأنماط المرتبطة بالعطل ورفع أداء مقاييس التصنيف الحساسة للفئات النادرة.

ولضمان بناء نماذج تنبؤية مستقرة وذات موثوقية عالية وقابلة للتطبيق في البيئات الصناعية تم التركيز على الأعطال الأربعة الأكثر تمثيلاً والأعلى ارتباطاً بالمتغيرات التشغيلية على الترتيب:

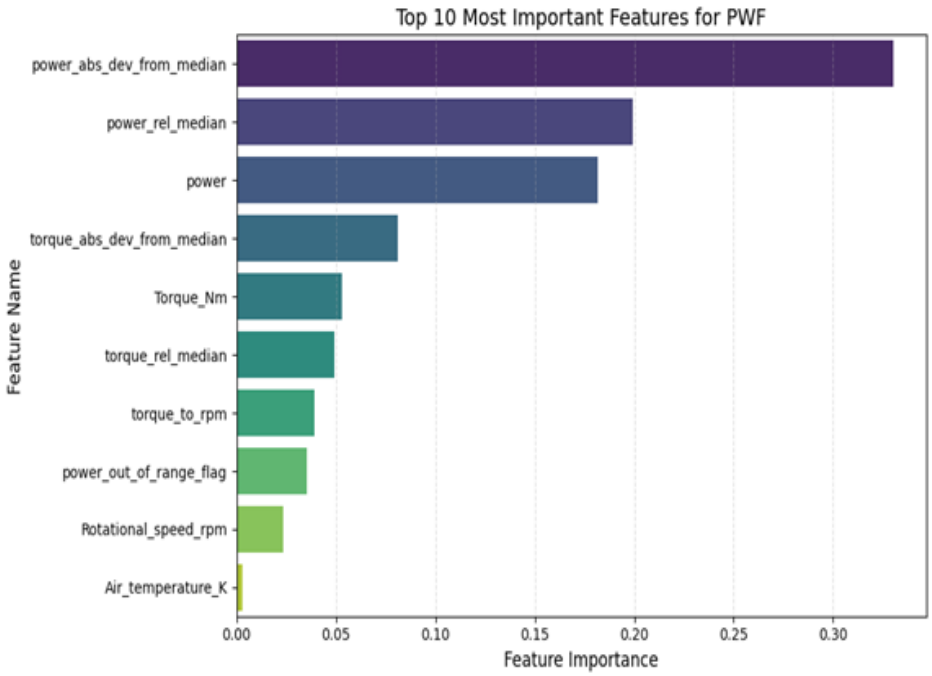
1. PWF
2. HDF
3. OSF
4. TWF

3.1.1.1 تحديد أهمية الميزات المؤثرة في كل عطل باستخدام Random Forest

تم الاعتماد على نموذج Random Forest نظراً لقدرته العالية على التقاط التفاعلات غير الخطية بين المتغيرات وتقدير أهمية كل ميزة ضمن سياق التصنيف، يتميز هذا النموذج بآلية داخلية لحساب أهمية الميزات، مما يتيح تحديد الميزات الأكثر ارتباطاً بكل نمط عطل، وبالتالي توجيه عملية هندسة الميزات نحو بناء تمثيل أكثر دقة لسلوك الآلة.

- **PWF**

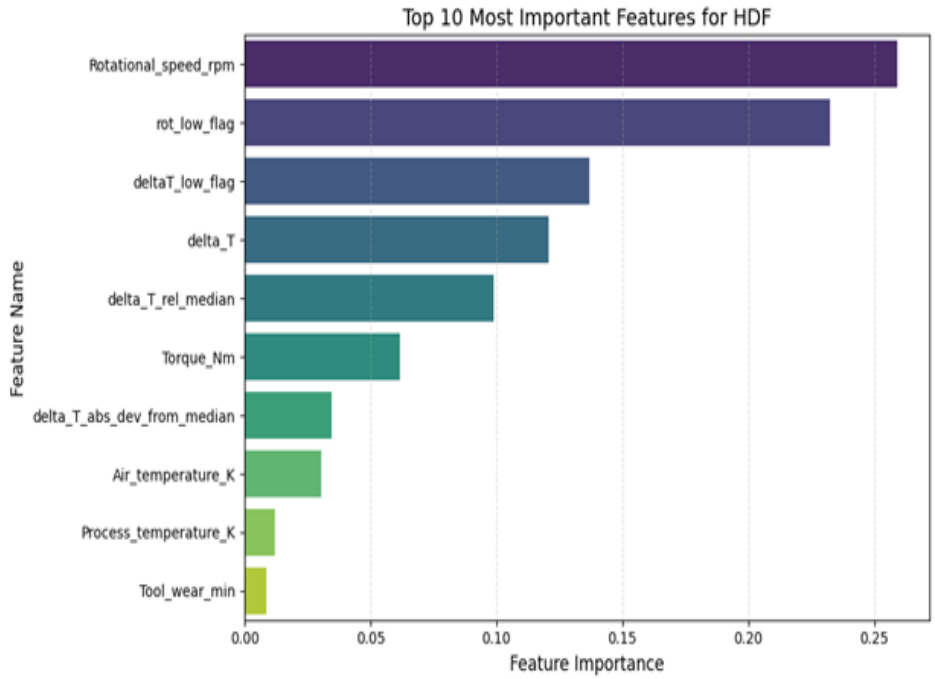
يبين شكل 2 ترتيب أهم الميزات لنمط العطل PWF نلاحظ أن الميزات المرتبطة بالقدرة Power مثل power_abs_dev_from_median و power_rel_median هي الأكثر مساهمة، ما يدل على أن انحرافات القدرة عن السلوك الطبيعي تلعب الدور الأكبر في كشف العطل،



شكل 2 ترتيب أهم الميزات لنمط العطل PWF

- **HDF**

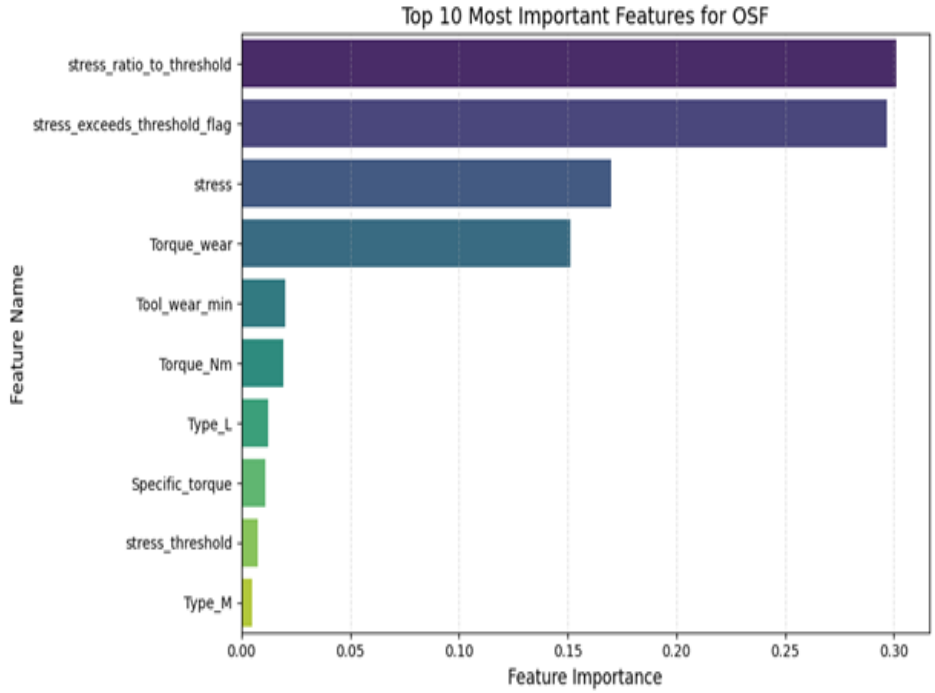
يوضـــح شـــكل 3 ترتيب أهم الميزات لنمط العطل HDF نلاحظ أن الميزات المرتبطة بانخفاض الســرعة الدورانية مثل Rotational_speed_rpm و rot_low_flag كانت الأكثر مساهمة، ما يدل على أن سلوك السرعة هو العامل الأبرز في كشف هذا النوع من الأعطال.



شكل 3 ترتيب أهم الميزات لنمط العطل HDF

- **OSF**

نلاحظ من الشكل 4 ترتيب أهم الميزات لنمط العطل OSF، أن الميزات المرتبطة بالإجهاد الميكانيكي مثل stress_ratio_to_threshold و stress_exceeds_threshold_flag هي الأكثر تأثيراً، ما يشير إلى أن تجاوز الإجهاد للحدود التصميمية هو العامل الأساسي في حدوث هذا العطل.



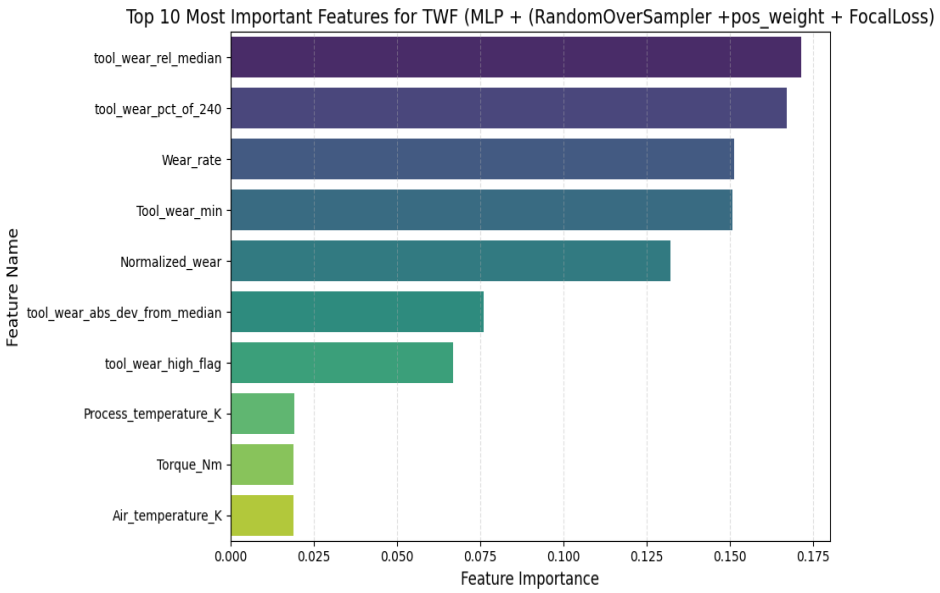
شكل 4 ترتيب أهم الميزات لنمط العطل OSF

- **TWF**

نظراً لندرة عطل TWF مقارنةً ببقية الأعطال، تم التعامل معه كنموذج فرعي مستقل ضمن بنية هجينة لضمان رفع حساسية الكشف عنه.

ولتحقيق ذلك، تم تدريب شبكة عصبونية متعددة الطبقات (MLP) مخصّصة لاكتشاف هذا العطل باستخدام مزيج من تقنيات معالجة اختلال التوزيع الطبقي.

يوضح شكل 5 أهم الميزات لنمط العطل TWF بعد معالجة عدم التوازن:



شكل 5 ترتيب أهم الميزات لنمط العطل TWF بعد معالجة عدم التوازن

3.1.2 دراسة عطل الآلة باستخدام الميزات المشتقة

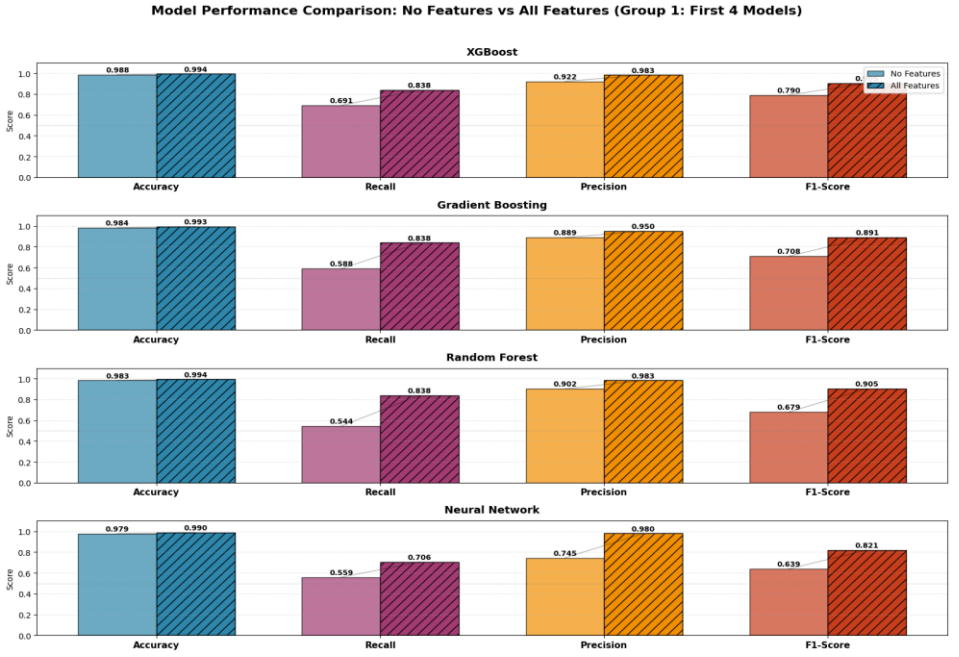
يبين جدول 2 نتائج تطبيق مجموعة من خوارزميات تعلّم الآلة على هذه الميزات المشـتقة بهدف تقييم قدرتها على اكتشـاف حالات العطل، وذلك من خلال تحليل أداء النماذج باسـتخدام مقاييس التقييم.

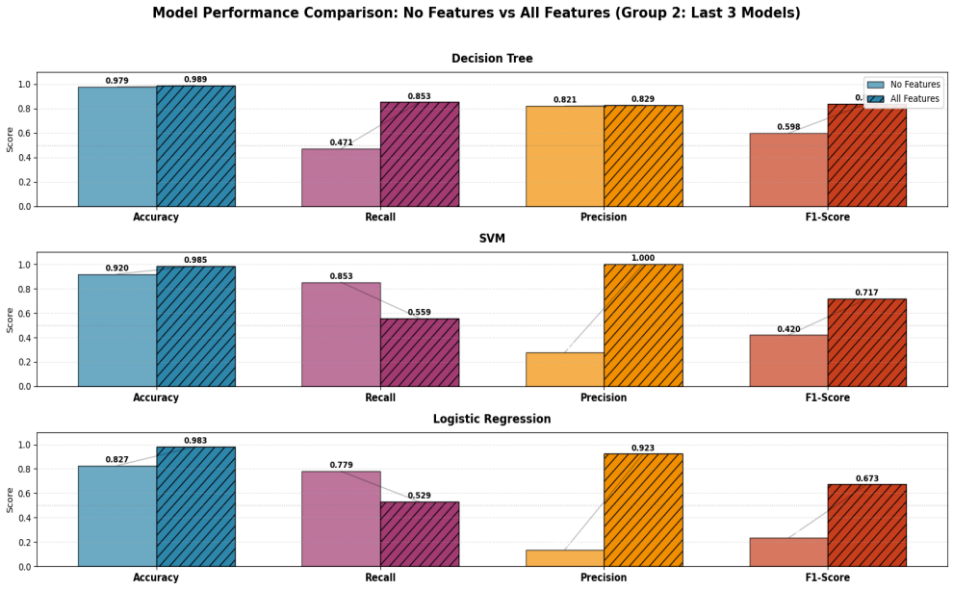
Model	Accuracy	Recall	Precision	F1-Score
XGBoost (Baseline)	0.9940	0.838235	0.982759	0.904762
Gradient Boosting (Baseline)	0.9930	0.838235	0.950000	0.890625
Random Forest (Baseline)	0.9940	0.838235	0.982759	0.904762
Neural Network (MLP 64→32→16)	0.9895	0.705882	0.979592	0.820513

Decision Tree (Baseline)	0.9890	0.852941	0.828571	0.840580
SVM (LinearSVC Baseline)	0.9850	0.558824	1.000000	0.716981
Logistic Regression (Baseline)	0.9825	0.529412	0.923077	0.672897

جدول 2 تقييم أداء خوارزميات تعلم الآلة باستخدام الميزات المشتقة

3.1.3 مقارنة أداء النماذج قبل وبعد استخدام الميزات المشتقة





شكل 6 مقارنة أداء النماذج قبل وبعد استخدام الميزات المشتقة

يوضح شكل 6 المقارنة بين أداء النماذج قبل وبعد استخدام الميزات المشتقة أثر هندسة الميزات في تحسين قدرة النماذج على اكتشاف حالات العطل، حيث أظهرت معظم النماذج تحسّناً ملحوظاً في مقاييس الاسترجاع وF1 Score بعد إدراج الخصائص المشتقة التي تعبّر بشكل أدق عن السلوك الميكانيكي للنظام.

ولكن على الرغم من أن استخدام الميزات المشتقة قد أدى إلى تحسين ملحوظ في دقة النماذج وارتفاع كبير في Precision وF1 Score، إلا أن قيمة Recall انخفضت في SVMوLogistic Regression ويُعزى هذا الانخفاض إلى أن الميزات المشتقة، رغم قدرتها على تعزيز الفصل بين الفئات، قد جعلت النموذج أكثر تحفظاً في تصنيف العينات على أنها حالات عطل، فالنماذج الخطية مثل SVM وLogistic Regression تعتمد على حدود قرار حادة، وعندما تصبح الميزات أكثر تمييزاً، فإن النموذج يميل إلى تضييق منطقة القرار الخاصة

بالفئة الإيجابية (العطل)، مما يقلل من عدد العينات المصنّفة كعطل ويؤدي بالتالي إلى انخفاض Recall.

3.2 المرحلة الثانية : تحديد نوع العطل

في هذه المرحلة، تم توحيد جميع أنواع الأعطال الجزئية في متغير واحد متعدد الفئات بهدف إنشاء هدف تصنيفي موحّد يسهل التعامل معه في نماذج تعلّم الآلة، ولتحقيق ذلك، جرى إنشاء عمود جديد باسم Machine failure كما هو موضح في جدول 3:

جدول 3 جدول ترميز فئات أعطال الآلة

Condition	Assigned Value
No Failure	0
TWF == 1	1
HDF == 1	2
PWF == 1	3
OSF == 1	4
RNF == 1	5

وبعد تنفيذ النماذج الأربعة الأولى على متغير Machine Failure المرمّز، كانت نتائج الأداء كما هي مبينة في جدول 4 :

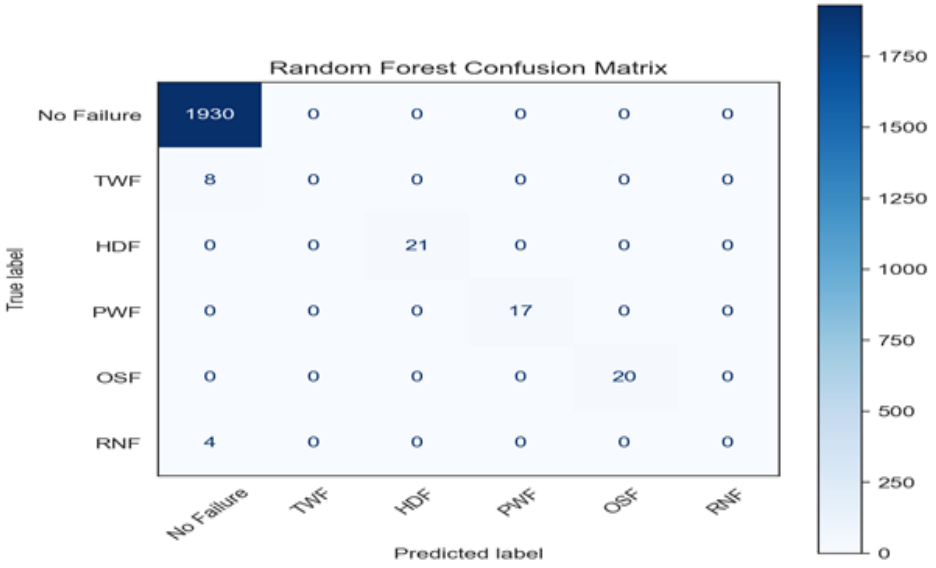
جدول 4 مقارنة أداء النماذج في مرحلة تحديد نوع العطل

Model	Accuracy	Recall	Precision	F1-Score
XGBoost Classifier	0.9940	0.9940	0.988037	0.991009
Gradient Boosting	0.9925	0.9925	0.987537	0.990005
Random Forest	0.9940	0.9940	0.988037	0.991009

Neural Network (MLP 64→32→16)	0.9880	0.9880	0.981427	0.983787

تُظهر نتائج النماذج أن أداء خوارزميات تعلّم الآلة ظلّ مرتفعاً عند التعامل مع متغير Machine failure متعدد الفئات وتؤكد هذه النتائج أن عملية الترميز لم تُضعف أداء النماذج، بل مكّنتها من التعرّف على أنماط الأعطال المختلفة بكفاءة عالية، مما يجعلها مناسبة للتنبؤ بنوع العطل بشكل أكثر دقة وموثوقية.

يعرض شكل 7 مصفوفة الالتباس الخاصة بنموذج Random Forest بعد تطبيقه على مهمة تصنيف نوع العطل باستخدام متغير Machine Failure المرمّز.



شكل 7 مصفوفة الالتباس لنموذج الغابات العشوائية (تصنيف الأعطال)

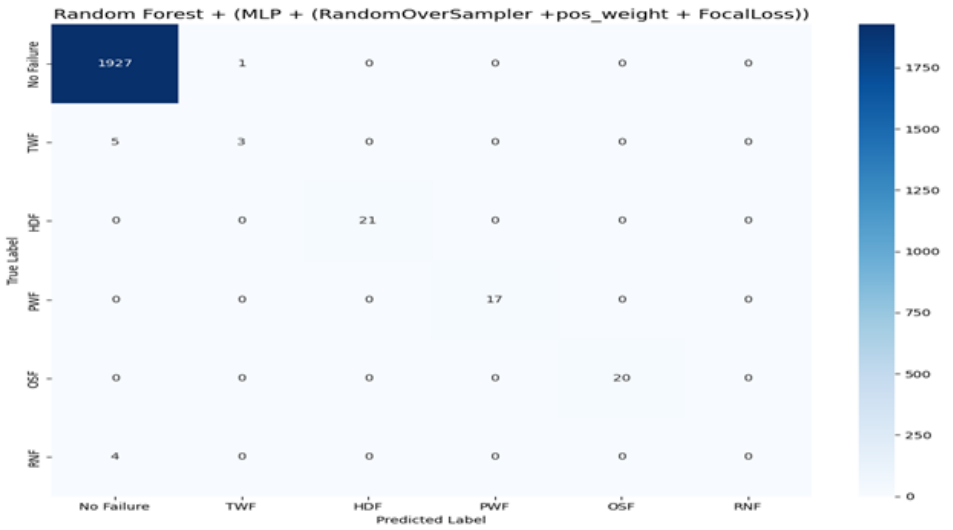
توضح المصفوفة أن النموذج نجح بشكل ممتاز في تصنيف الحالات السليمة، حيث تعرّف على جميعها دون أخطاء، كما أظهر دقة عالية في تصـنيف أعطال HDF و PWF وOSF، إذ تم التعرّف عليها بالكامل.

في المقابل، لم يتمكّن النموذج من التعرّف على أي حالة من حالات TWF ، حيث تم تصنيفها جميعاً ضمن الفئة السليمة، مما يشير إلى ضعف النموذج في التعامل مع الأعطال النادرة.

ولمعالجة مشكلة ضعف النماذج في التعرّف على عطل TWF تم استخدام نهج هجين كما في المرحلة الاولى وقد أسـهم هذا النهج الهجين في رفع حسـاسـية النظام تجاه عطلTWF دون التأثير على أداء النماذج في الفئات الأخرى، مما أدى إلى بنية أكثر توازناً وموثوقية في تحديد نوع العطل [17].

أما عطل RNF، فقد تم اسـتبعاده من النموذج المخصّـص للأعطال النادرة نظراً لغياب نمط تشـغيلي واضـح يمكن للنموذج تعلّمه، إضـافةً إلى ندرة العينات المرتبطة به، مما يجعل بناء نموذج مسـتقل له غير مسـتقر وغير قابل للتعميم، ويؤكد هذا الواقع محدودية القدرة على التنبؤ بهذا النوع من الأعطال، نظراً لعدم توفر خصـائص مميزة أو سـلوك ميكانيكي ثابت يمكن استغلاله نمذجياً.

يعرض شـكل 8 نتائج النموذج الهجين الذي يجمع بين Random Forest و MLP ، مع استخدام تقنيات معالجة عدم التوازن لتحسين التعرّف على الأعطال النادرة



شكل 8 مصفوفة الالتباس النهائية لنموذج التصنيف الهجين

تُظهر المصفوفة تحسناً ملحوظاً في اكتشاف عطل TWF مقارنة بالنموذج السابق، حيث تم تصنيف ثلاث حالات بشكل صحيح بعد أن كانت جميعها تُخطئ سابقاً، كما حافظ النموذج على دقة عالية في تصنيف الأعطال الشائعة HDF وPWF وOSF، دون أي تراجع في الأداء، وتُبرز هذه النتائج فعالية النهج الهجين في تعزيز حساسية النموذج تجاه الأعطال النادرة دون التأثير على دقته في الفئات الأخرى.

4 مقارنة نتائج النموذج المقترح مع الدراسات السابقة

بعد دراسة وتحليل مقاييس الأداء الخاصة بالنموذج المقترح في هذا البحث، تم إجراء مقارنة منهجية بين نتائج النموذج الحالي ونتائج مجموعة من الدراسات المرجعية الحديثة [5]،[6]،[7]،[1] في مجال الصيانة التنبؤية كما هو موضح في جدول 5.

جدول 5 مقارنة مع الأعمال ذات الصلة

الدراسة	نوع البيانات	الميزات المستخدمة	معالجة عدم التوازن	أفضل نموذج	تصنيف الأعطال (Multi-Class)
5	بيانات حساسات حقيقية من آلات غزل صناعية	ميزات خام	غير مستخدمة	Random Forest	✗
6	بيانات تشغيل كهربائية حقيقية لمحركات كهربائية	ميزات خام	Oversampling بسيط	XGBoost	✗
7	AI4I 2020	ميزات إحصائية بسيطة	غير مستخدمة	SVM	✗
1	AI4I 2020	ميزات خام	Class Weights	Gradient Boosting	✓
Our Work	AI4I 2020	ميزات خام و ميزات مشتقة	Oversampling + pos_weight + Focal Loss	MLP - XGBoost - Gradient Boosting -	✓

				Random Forest	

تُبرز المقارنة بين النموذج المقترح والدراســـات المرجعية الأربع تفوق المنهجية المعتمدة في هذا البحث، والتي تجمع بين هندســـة الميزات المتقدمة، ومعالجة عدم التوازن، واســـتخدام نماذج هجينة، فقد أظهرت الدراســـات الســـابقة اعتماداً كبيراً على الميزات الخام والنماذج التقليدية دون معالجة فعّالة لعدم التوازن، مما أدى إلى انخفاض واضـح في مؤشـرات Recall و F1-Score للفئات النادرة رغم ارتفاع الدقة العامة.

كما تفوقت الشبكة العصبونية المخصّصة على جميع النماذج المستخدمة في الدراسات السابقة من حيث قدرة اكتشـــاف الأعطال النادرة، مما يؤكد فعالية التصـــميم الهجين المقترح وملاءمته للتطبيقات الصناعية التي تتطلب دقة عالية في رصد حالات العطل.

5 خاتمة

تم في هذه الدراسـة تطوير نموذج صـيانة تنبؤية قادر على التعرّف على الأعطال التشـغيلية في آلات التفريز بدقة عالية، مع قدرة مميزة على اكتشـاف الأعطال النادرة مثلTWF . وقد سـاهم دمج هندسـة الميزات مع تقنيات معالجة عدم التوازن في تحسـين مؤشـرات الأداء بشـكل واضـح، بينما أظهر النموذج الهجين الذي يجمع بين النماذج الشـــجرية والشـــبكة العصـــبونية تفوقاً في تصنيف الأعطال الشائعة والنادرة على حد سواء، وتؤكد النتائج أن المنهجية المقترحة توفر إطاراً عملياً وفعالاً للصـــيانة التنبؤية، يدعم تقليل التوقفات غير المخططة ويعزز موثوقية المعدات في البيئات الصناعية.

المراجع

1. Zhao, W., Ding, J., Huang, X., & Zhang, Y. (2025). Research on milling machine predictive maintenance based on machine learning and SHAP analysis in intelligent manufacturing environment (arXiv:2512.01205v1). arXiv.

2. Kumar, R., et al. (2023). Impact of machine downtime on modern manufacturing systems. Procedia CIRP.

3. Carvalho, T., et al. (2023). Predictive maintenance in Industry 4.0: Recent developments and challenges. Sensors.

4. Zhao, W., et al. (2025). Machine learning based fault diagnosis in CNC milling machines. Applied Sciences.

5. Akyaz, T., & Engin, D. (2024). Machine learning-based predictive maintenance system for artificial yarn machines. IEEE Access, 12, 125446–125461.

6. Baradaran, A. H. (2025). Predictive maintenance of electric motors using supervised learning models: A comparative analysis (arXiv:2504.03670). arXiv.

7. Waghulde, R. R., Rai, R. K., Chadhar, R. M., Rane, M., & Yadav, V. (2025). Evaluating machine learning and deep learning models for predictive maintenance: A study using the AI4I 2020 dataset. Journal Name.

8. AI4I 2020 Predictive Maintenance Dataset- UCI Machine Learning Repository (2020)

9. Wang, Z., Liu, Q., & Zhao, R. (2022). Data imbalance challenges in predictive maintenance: A comprehensive review. IEEE Transactions on Industrial Informatics, 18(8), 5234–5248.

10. https://www.sanfoundry.com/data-preprocessing-ml-cleaning-normalization-encoding

11. Tremblay, R. (2025). Understanding equipment failure behavior: Types, patterns, and maintenance strategies. Spartakus.

12. Z. M. Çınar et al., "Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0," Sustainability, vol. 12, art. 8211, Oct. 2020.

13. Buabeng, A. Simons, N. K. Frempong, and Y. Y. Ziggah, "Hybrid Intelligent Predictive Maintenance Model for Multiclass Fault Classification," Soft Computing, vol. 28, pp. 8749–8770, Jul. 28, 2023.

14. Chugani, V. (2025). A concise guide to overfitting. Statology.

15. Crudu, A., & MoldStud Research Team. (2025). Mitigating vanishing gradients with advanced activation functions — A comprehensive guide.

16. Gu, J., Li, C., Liang, Y., Shi, Z., & Song, Z. (2024). Exploring the frontiers of Softmax: Provable optimization, applications in diffusion model, and beyond (arXiv preprint arXiv:2405.00000). arXiv.

17. J. Wiercioch, "Hybrid predictive maintenance model – study and implementation example," Production Engineering Archives, vol. 30, no. 3, 2024.