

**BÜYÜK DİL MODELLERİNDE EĞİTİMSİZ STATİK DİKKAT
MASKELEMENİN YAPISAL SINIRLARI:
ÇARPIMSAL MODÜLASYON ÜZERİNE MEKANİSTİK BİR İNCELEME**

YÜKSEK LİSANS TEZİ

Fatih Çağrı BİLGEHAN

**Danışman
Dr. Öğr. Üyesi Özkan ASLAN**

İNTERNET VE BİLİŞİM TEKNOLOJİLERİ YÖNETİMİ ANABİLİM DALI

Temmuz 2026

AFYON KOCATEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

**BÜYÜK DİL MODELLERİNDE EĞİTİMSİZ STATİK
DİKKAT MASKELEMENİN YAPISAL SINIRLARI:
ÇARPIMSAL MODÜLASYON ÜZERİNE
MEKANİSTİK BİR İNCELEME**

Fatih Çağrı BİLGEHAN

Danışman

Dr. Öğr. Üyesi Özkan ASLAN

**İNTERNET VE BİLİŞİM TEKNOLOJİLERİ YÖNETİMİ
ANABİLİM DALI**

Temmuz 2026

TEZ ONAY SAYFASI

Fatih Çaęrı BİLGEHAN tarafından hazırlanan “Büyük Dil Modellerinde Eğitimsiz Statik Dikkat Maskelemenin Yapısal Sınırları: Çarpımsal Modülasyon Üzerine Mekanistik Bir İnceleme” adlı tez çalışması lisansüstü eğitim ve öğretim yönetmeliğinin ilgili maddeleri uyarınca 02/07/2026 tarihinde aşağıdaki jüri tarafından **oy birlięi** ile Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü **İnternet ve Bilişim Teknolojileri Yönetimi Anabilim Dalı**’nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Dr. Öğr. Üyesi Özkan ASLAN

Başkan : Dr. Öğr. Üyesi Naim KARASEKRETER
Afyon Kocatepe Üniversitesi, Mühendislik Fakültesi,
Bilgisayar Mühendislięi Bölümü

Üye : Dr. Öğr. Üyesi İbrahim GÜRGİL
Uşak Üniversitesi, Mühendislik ve Doęa Bilimleri
Fakültesi, Matematik Bölümü

Üye : Dr. Öğr. Üyesi Özkan ASLAN
Afyon Kocatepe Üniversitesi, Mühendislik Fakültesi,
Bilgisayar Mühendislięi Bölümü

Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü Yönetim Kurulu’nun
...../...../2026 tarih ve
..... sayılı kararıyla onaylanmıştır.

Prof. Dr. Bekir YALÇIN
Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİM SAYFASI

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

02/07/2026

Fatih Çağrı BİLGEHAN

ÖZET

Yüksek Lisans Tezi

BÜYÜK DİL MODELLERİNDE EĞİTİMSİZ STATİK DİKKAT MASKELEMENİN YAPISAL SINIRLARI: ÇARPIMSAL MODÜLASYON ÜZERİNE MEKANİSTİK BİR İNCELEME

Fatih Çağrı BİLGEHAN

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü

İnternet ve Bilişim Teknolojileri Yönetimi Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Özkan ASLAN

Büyük dil modellerinin (LLM) temelini oluşturan öz-dikkat mekanizması, deterministik doğası gereği dilin yapısal tutarlılığını korurken metinsel yaratıcılığı kısıtlayabilmektedir. Bu yüksek lisans tezinde, söz konusu yaratıcılık-tutarlılık ödünleşimini eğitim gerektirmeden (zero-shot) ele almayı hedefleyen statik bir çarpımsal dikkat maskeleyme yöntemi (BAW) önerilmiş ve mekanistik sınırları kapsamlı biçimde analiz edilmiştir.

BAW yöntemi, dikkat matrisi üzerine doğrudan etki eden çarpımsal bir ağırlık mekanizmasıdır. Yerel sözdizimsel bağları koruyan Gaussian bir bileşen ile kuantum dönüş kapısının (RX) ölçüm olasılığı formundan ($\cos^2$) esinlenen, matematiksel olarak periyodik bir yükseltilmiş-kosinüs (raised cosine) çekirdeği olan ve uzak bağlamlarda periyodik dikkat kontrastı yeniden ölçeklemesi sağlayan kuantum-esinli bir bileşenin entegrasyonundan oluşur. Mekanizma, GPT-2 mimari ailesinin üç ölçeğinde (Small, Medium, Large), dört farklı kolda (Taban, Gaussian, Kuantum, Hibrit) toplam 108.000 metin üzerinde 16 otomatik metrikle (13 temel, 3 ikincil doğrulayıcı) değerlendirilmiştir.

Tezin en temel teorik katkısı, çarpımsal maskenin ($W = G \times Q$) yapısal bir menzilsiz uyumsuzluğu taşıdığının analitik olarak ispatlanmasıdır: Gaussian bileşenin hızlı üstel sönümlenmesi, kuantum bileşeninin uzak-menzil rezonans zirvelerini temsili $\sigma = 4$, $\theta_{scale} = 0.10$ konfigürasyonunda 10^{-54} ; incelenen hiperparametre ızgarasının tamamında (analitik

fizibilite analizi) ise 10^{-7} – 10^{-280} aralığında sıfırlayarak hibritte sinerjiyi engellemiştir. Kuantum bileşeni izole olarak (Quantum-only kolunda) tüm dizi boyunca aktif kalmış ve Base modele kıyasla ölçülebilir yapısal kaymalar üretmiş; ancak bu kaymalar yaratıcılık vekil (proxy) metriklerinde sistematik bir üstünlüğe dönüşmemiştir. Ayrıca bu sınırlı etki, özellikle yaratıcılık ve çeşitlilik proxy metriklerinde model ölçeği büyüdükçe belirgin biçimde zayıflamıştır. Son olarak, yaygın kullanılan otomatik yaratıcılık proxy metriklerinin yüzeysel anahtar kelime sayımına dayanan kırılğan yapıları tespit edilmiş; niteliksel kaymaları algılamada yetersiz kaldıkları gösterilmiştir. Buna ek olarak gerçekleştirilen körleştirilmiş büyük dil modeli hakem (LLM-as-judge) ikincil değerlendirmesi de kollar arası anlamlı bir kalite veya yaratıcılık farkı bulamamış; hakemde tespit edilen güçlü pozisyon yanlılığı ise otonom yaratıcılık değerlendirme yöntemlerinin yapısal kırılğanlığına bağımsız bir gösterge sunmuştur.

2026, xi + 61 sayfa

Anahtar Kelimeler: Dikkat mekanizması, Pre-softmax dikkat modülasyonu, Eğitimsiz (zero-shot) modülasyon, Menzil-uyumsuzluğu, Yükseltilmiş kosisinüs çekirdeği, Otomatik yaratıcılık ölçümü.

ABSTRACT

M.Sc. Thesis

STRUCTURAL LIMITS OF TRAINING-FREE STATIC ATTENTION MASKING IN LARGE LANGUAGE MODELS: A MECHANISTIC INVESTIGATION OF MULTIPLICATIVE MODULATION

Fatih Çağrı BİLGEHAN

Afyon Kocatepe University

Graduate School of Natural and Applied Sciences

Department of Internet and Information Technologies Management

Supervisor: Asst. Prof. Dr. Özkan ASLAN

The self-attention mechanism underlying large language models (LLMs) inherently restricts textual creativity while preserving structural coherence due to its deterministic nature. This master’s thesis proposes and mechanistically analyzes a training-free (zero-shot) static multiplicative attention masking method (BAW) designed to address the creativity-coherence trade-off.

The BAW method is a multiplicative weight mechanism acting directly upon the attention matrix, integrating a Gaussian “local stability” component that preserves local syntactic dependencies with a component inspired by the measurement-probability form ($\cos^2$) of the quantum RX rotation gate, mathematically a periodic raised-cosine kernel, that periodically rescales attention contrast at distant token positions. The mechanism was evaluated across three scales of the GPT-2 architectural family (Small, Medium, Large) in four experimental arms (Base, Gaussian, Quantum, Hybrid) over 108,000 generated texts using 16 automated metrics (13 primary, 3 secondary confirmatory).

The primary theoretical contribution of this thesis is the analytical proof that the multiplicative mask ($W = G \times Q$) exhibits a structural range-mismatch: the rapid exponential decay of the Gaussian component nullifies the quantum component’s long-range resonance peaks at the order of 10^{-54} in the representative $\sigma = 4, \theta_{scale} = 0.10$ configuration,

and between 10^{-7} and 10^{-280} across the entire analyzed hyperparameter grid (analytical feasibility analysis), thereby precluding any synergistic interaction in the Hybrid arm. The quantum component, when tested in isolation (Quantum-only arm), remained active across the full sequence length and produced measurable structural shifts relative to the Base model; however, these shifts did not translate into systematic improvements in proxy creativity metrics. Furthermore, this limited effect was markedly weakened as model scale increased, particularly in proxy creativity and diversity metrics. Finally, the study identifies the fragile nature of widely used automated proxy creativity metrics that rely on superficial keyword counting and demonstrates their inadequacy in capturing qualitative structural shifts. In addition, the secondary blinded LLM-as-judge evaluation found no significant quality or creativity difference between the experimental arms; while the strong position bias detected in the judge provided independent evidence of the structural fragility of autonomous creativity evaluation methods.

2026, xi + 61 pages

Keywords: Attention mechanism, Pre-softmax attention modulation, Training-free (zero-shot) modulation, Range-mismatch, Raised-cosine kernel, Automated creativity evaluation.

TEŞEKKÜR

Bu araştırmanın konusunun belirlenmesi, çalışmaların yönlendirilmesi, sonuçların değerlendirilmesi ve tezin yazımı aşamalarındaki büyük katkılarından dolayı tez danışmanım Sayın Dr. Öğr. Üyesi Özkan ASLAN'a; araştırma ve yazım süresince öneri ve eleştirileriyle yardımlarını esirgemeyen tüm hocalarıma ve arkadaşlarıma teşekkür ederim.

Bu araştırma boyunca maddi ve manevi desteklerinden dolayı aileme teşekkür ederim.

Fatih Çağrı BİLGEHAN
Afyonkarahisar 2026

İÇİNDEKİLER DİZİNİ

	Sayfa
ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER DİZİNİ	vi
SİMGELER ve KISALTMALAR DİZİNİ.....	viii
ŞEKİLLER DİZİNİ	x
ÇİZELGELER DİZİNİ.....	xi
1. GİRİŞ	1
2. LİTERATÜR BİLGİLERİ	3
2.1 Temel Kavramlar ve Literatür Özeti	3
2.1.1 Transformer Mimarisi ve Öz-Dikkat Mekanizması	3
2.1.2 Büyük Dil Modelleri ve GPT-2 Mimari Ailesi	4
2.1.3 Yaratıcılık-Tutarlılık Ödünleşimi.....	4
2.1.4 Kuantum Rotasyon Kapıları ve Born Kuralı	5
2.1.5 Değerlendirme Metrikleri ve Otomatik Yaratıcılık Ölçümünün Sınırları	6
2.2 İlgili Çalışmalar ve Literatür İncelemesi	7
2.2.1 Klasik Dikkat Mekanizması İyileştirmeleri	7
2.2.2 Örnekleme Tabanlı Metin Üretimi ve Yaratıcılık Yöntemleri	8
2.2.3 Çıkarım-Zamanı Dikkat Yönlendirme (Inference-Time Attention Steering)	8
2.2.4 Pozisyonel Kodlama ve Mesafe Yanlılığı Ailesi	9
2.2.5 Yaratıcılık Ölçümü Literatürü	11
2.2.6 NLP’de Kuantum ve Kuantum-Esinli Yaklaşımlar	12
2.2.7 Literatürdeki Boşluk ve Bu Tezin Konumu	13
3. MATERYAL ve METOT	14
3.1 Yöntem: Bilgehan-Aslan Penceresi (BAW)	14
3.1.1 Genel Mimari ve Çarpımsal Modülasyon Tercihi	14
3.1.2 Gaussian Bileşen (Yerel Kararlılık)	15
3.1.3 Kuantum Bileşen (Periyodik Uzaklık Fonksiyonu)	15
3.1.4 Uygulama Konfigürasyonu ve Katman-Seçicilik Tartışması	17
3.2 Teorik Analiz: Çarpımsal Maskenin Matematiksel Yapısı	18
3.2.1 Önerme 1: Menzil-Uyumsuzluğu ile Kuantum İptali	18
3.2.2 Quantum-only Kolunda İptalin Neden Geçerli Olmadığı	20
3.2.3 Çarpımsal Maskenin Logit Etkileşimi: Kontrast Düzleştirme	21

3.2.4	Spektral Bakış: Hibrit Maskenin Fourier Ayırışması	23
3.2.5	Toplamsal (Additive) Alternatifin Formel Analizi.....	23
3.3	Deneysel Düzen ve Metodoloji	24
3.3.1	Kullanılan Modeller ve Ortam	25
3.3.2	Veri Üretimi ve Örneklem Büyüklüğü	25
3.3.3	BAW Maskeleye Algoritması.....	26
3.3.4	Kullanılan Değerlendirme Metrikleri.....	27
3.3.5	İstatistiksel Analiz ve Çoklu Karşılaştırma Düzeltmesi	28
3.3.6	Körleştirilmiş LLM-as-Judge İkincil Değerlendirmesi	29
4.	BULGULAR.....	33
4.1	Bulgu A: Hiçbir Kol Yaratıcılık Proxy Metriklerinde Base'i Geçmiyor	33
4.2	Bulgu B: Hibrit Modülasyonda Sinerji Yokluğu.....	33
4.3	Bulgu C: Modülasyon Etkisinin Ölçek ile Sönümlenmesi	34
4.4	Bulgu D: Sözcüksel Çeşitlilik ve Sözdizimsel Karmaşıklık Arasındaki Takas	34
4.5	Bileşen İzolasyonu: Quantum-only vs Gaussian-only vs Hybrid	35
4.6	İkincil Doğrulayıcı Metrikler (LZ-Karmaşıklık, Konu Sapması, Anlamsal Yayılım)	36
4.7	Niteliksel Metin İncelemesi ve Otomatik Metriklerin Sınırları	37
4.8	Körleştirilmiş LLM-as-Judge İkincil Değerlendirmesi.....	40
5.	TARTIŞMA ve SONUÇ.....	48
5.1	Tartışma ve Sınırlılıklar.....	48
5.1.1	Menzil-Uyumsuzluğunun Ampirik Teyidi ve Bulgu Çerçevesi	48
5.1.2	Otomatik Yaratıcılık Metriklerinin Yetersizliği ve Kırılganlığı	49
5.1.3	Karşılaştırma Temeli Eksikliği	51
5.1.4	Modülasyon Sinerjisinin ve Ölçek Sönümlenmesinin Değerlendirilmesi	51
5.1.5	Statik Katman Konfigürasyonu (Depth=1) Sınırlılığı	52
5.1.6	İstatistiksel Yöntem Tartışması.....	52
5.2	Sonuç.....	53
6.	KAYNAKLAR	56
	ÖZGEÇMİŞ	61

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler

$G(i, j; \sigma)$	Gaussian yerel kararlılık fonksiyonu
$Q(i, j; \theta)$	Kuantum periyodik uzaklık fonksiyonu
$W(i, j)$	BAW (Bilgehan-Aslan Window) maske ağırlığı
σ	Gaussian dağılımının standart sapması
θ	Kuantum RX kapısının dönüş açısı
θ_{scale}	Kuantum bileşeni ölçeklendirme hiperparametresi
d_{ij}	Çıktı dizisindeki tokenlar arası göreceli mesafe
d^*	Kuantum bileşenin ilk rezonans zirvesi mesafesi ($2\pi/\theta_{\text{scale}}$)
T	Kuantum bileşenin uzamsal periyodu ($2\pi/\theta_{\text{scale}}$)
ε	Rezonans görünürlük eşiği
β	Toplamsal modülasyonda periyodik bileşen katsayısı
$G(d^*)$	İlk rezonans zirvesindeki Gaussian sönümlenme değeri

Kısaltmalar

ALiBi	Attention with Linear Biases (Doğrusal Yanlılıklarla Dikkat)
ANOVA	Analysis of Variance (Varyans Analizi)
BAW	Bilgehan-Aslan Window (Bilgehan-Aslan Penceresi)
BH	Benjamini-Hochberg (Benjamini-Hochberg Düzeltmesi)
DAT	Divergent Association Task (İraksak İlişkilendirme Görevi)
DSI	Divergent Semantic Integration (İraksak Anlamsal Bütünleştirme)
FDR	False Discovery Rate (Yanlış Keşif Oranı)
GPT	Generative Pre-trained Transformer (Üretken Ön-eğitilmiş Dönüştürücü)
ICC	Intraclass Correlation Coefficient (Sınıf-içi Korelasyon Katsayısı)

LLM	Large Language Model (Büyük Dil Modeli)
LMM	Linear Mixed Model (Doğrusal Karma Model)
MTLD	Measure of Textual Lexical Diversity (Metinsel Sözcüksel Çeşitlilik Ölçüsü)
NISQ	Noisy Intermediate-Scale Quantum (Gürültülü Orta-Ölçekli Kuantum)
NLP	Natural Language Processing (Doğal Dil İşleme)
QGPM	Quantum-Inspired Generative Pre-trained Modulator
RoPE	Rotary Position Embedding (Döner Konum Gömme)
RX	Quantum X-axis Rotation Gate (Kuantum X-ekseni Dönüş Kapısı)
Self-BLEU	Self-Bilingual Evaluation Understudy (Öz-İkidilli Değerlendirme Altyapısı)

ŞEKİLLER DİZİNİ

	Sayfa
Şekil 3.1 BAW dikkat penceresinin uzaklık (d) bazlı modülasyon ağırlıkları (W_d). Kuantum ve Gaussian bileşenlerin çarpımsal birleşimi (Hibrit) kırmızı çizgi ile gösterilmiştir.	18
Şekil 3.2 $\theta_{scale} \times \sigma$ parametre uzayında ilk rezonans zirvesindeki Gaussian sönümlene değeri $G(d^*)$. Renk skalası logaritmik; test edilen 15 ızgara noktası kırmızı yıldızlarla, $\varepsilon = 0.01$ fizibilite eşik çizgisi ($\theta_{scale} \cdot \sigma = 2.07$) kesikli beyaz çizgiyle gösterilmiştir. Tüm deneysel noktalar fizibilite eşığının çok altında kalmaktadır.	20
Şekil 4.1 Model ölçeği büyüdükçe (Small $\rightarrow$ Medium $\rightarrow$ Large), BAW mekanizmalarının Base modele göre istatistiksel olarak anlamlı ($p < 0.05$) fark yarattığı metrik sayısındaki sönümlenme.	35
Şekil 4.2 GPT-2 Small modeli için anahtar metriklerdeki iç karşılaştırma etki büyüklükleri. d değerlerinin büyük çoğunluğunun ± 0.2 anlamsız sinerji bandında kalması, çarpımsal birleşimin ek bir etki yaratmadığını göstermektedir.	36

ÇİZELGELER DİZİNİ

	Sayfa
Çizelge 3.1 Belirli Mesafelerde BAW Maskesinin Analitik Değerleri ($\sigma = 4.0, \theta_{scale} = 0.10$)	30
Çizelge 3.2 Farklı görünürlük eşikleri (ε) için gerekli minimum $\theta_{scale} \cdot \sigma$ çarpımı. .	31
Çizelge 3.3 Metin üretiminde kullanılan sabit parametreler.	31
Çizelge 3.4 Çalışmada kullanılan 13 değerlendirme metriğinin tanımları ve hesaplama yöntemleri.	32
Çizelge 4.1 Üç farklı model ölçeği için 13 metrik üzerinden mekanizmalar arası farklılıkların ANOVA test sonuçları ($n = 45$ grup, $\alpha = 0.05$). [*] ham anlamlılık; Bonferroni eşiği $\alpha_{adj} = 0.00385$ altında hiçbir metrik anlamlı kalmamaktadır.	43
Çizelge 4.2 GPT-2 Small modeli için anahtar metriklerdeki Cohen's d etki büyüklükleri. Tüm karşılaştırma çiftleri.....	44
Çizelge 4.3 GPT-2 Medium modeli için tüm metriklerdeki Cohen's d etki büyüklükleri (vs Base).....	44
Çizelge 4.4 GPT-2 Large modeli için tüm metriklerdeki Cohen's d etki büyüklükleri (vs Base).	45
Çizelge 4.5 İkincil doğrulayıcı metrikler için ANOVA ve Cohen's d sonuçları. [*] işareti, doğrulayıcı analiz ailesi Bonferroni eşiğini ($p < 0.0167$) geçen bulguları temsil eder.	46
Çizelge 4.6 LLM-as-Judge değerlendirme skorları (1-5 arası) ve kollar arası tek-yönlü ANOVA sonuçları. Tüm p değerleri anlamsızdır ($p > 0.05$).....	47
Çizelge 4.7 LLM hakeminin anonim sunum sırasına (pozisyon) göre verdiği ortalama skorlar ve ANOVA sonuçları. [*] yüksek anlamlılığı ($p < 0.0001$) temsil eder.	47

1. GİRİŞ

Büyük Dil Modelleri (LLM’ler), Doğal Dil İşleme (NLP) alanında devrim yaratmış ve metin üretiminden kodlamaya kadar çok geniş bir yelpazede insan benzeri performans sergilemeye başlamıştır. Günümüzde en yaygın kullanılan LLM’lerin temelini oluşturan Transformer mimarisi (Vaswani vd. 2017), bağlamsal bilgiyi yakalamak için öz-dikkat (self-attention) mekanizmasına dayanmaktadır. Ancak, bu modellerin üretim sürecinin kalbinde yer alan softmax fonksiyonu, doğası gereği deterministik ve “kazanan-hepsini-alır” (winner-takes-all) eğilimine sahiptir. Bu durum, modelin genellikle en yüksek olasılıklı (en güvenli) ardışık tokenları seçmesine neden olmakta, bu da üretilen metinlerin dilbilgisi açısından kusursuz ancak içerik ve üslup açısından sıradan, monoton ve yaratıcılıktan uzak olmasına yol açabilmektedir.

Yaratıcılık (creativity) ve tutarlılık (coherence) arasındaki bu ödünleşim (trade-off), serbest metin üretimi görevlerinde temel bir zorluk teşkil etmektedir. Tutarlılık, metnin mantıksal ve anlamsal bütünlüğünü korumasını ifade ederken; yaratıcılık, kelime dağarcığının genişliğini, farklı alanlardan kavramların harmanlanmasını (cross-domain) ve sürpriz unsurunu içerir. Standart dikkat mekanizması, yapısal kısıtlamaları nedeniyle genellikle tutarlılığı maksimize ederken yaratıcılığı baskılamaktadır.

Literatürde bu sorunu çözmek için sıcaklık ölçeklendirme (temperature scaling), Top-k ve Top-p (nucleus) örnekleme (Holtzman vd. 2020) gibi yöntemler yaygın olarak kullanılmaktadır. Ancak bu yöntemler “sonradan müdahale” (post-hoc) niteliğindedir. Diğer bir deyişle, bu teknikler modelin içsel dikkat ağırlıklarını veya temsillerini değiştirmez; yalnızca modelin zaten üretmiş olduğu nihai olasılık dağılımının kuyruğunu budar veya dağılımı düzleştirir. Bu yüzeysel müdahale yaklaşımı, modelin eğitim verisinden öğrendiği temel ilişkilendirme yapılarına kalıcı bir esneklik kazandıramaz.

Bu tez çalışması, “Bilgehan-Aslan Penceresi” (BAW) adı verilen yeni bir karışık (hybrid) dikkat modülasyon mekanizmasını önermekte ve *yapısal sınırlarını mekanistik olarak analiz etmektedir*. Sonradan müdahale yöntemlerinin aksine, BAW mekanizması doğ-

rudan modelin öz-dikkat matrisine, softmax öncesinde çarpımsal bir maske olarak etki eder. BAW; (1) yerel semantik bağlamı korumak için bir Gaussian çekirdek ve (2) fiziksel sistemlerin periyodik davranışlarından (kuantum dönüş kapılarından) esinlenen ve uzak bağlamlardaki tokenlarda dikkat kontrastını periyodik olarak yeniden ölçeklemeyi hedefleyen bir periyodik fonksiyonu harmanlar.

Tezin temel katkısı, BAW mekanizmasının performans üstünlüğü iddia etmek değil; eğitimsiz statik dikkat modülasyonlarının yapısal sınırlarını analitik ve ampirik olarak ortaya koymaktır. Bu amaçla, önce çarpımsal maskenin ($W = G \times Q$) matematiksel yapısı analiz edilerek, Gaussian bileşenin hızlı üstel sönümlenmesinin kuantum bileşeninin uzak-menzil rezonansını yapısal olarak iptal ettiği (menzil-uyumsuzluğu) kanıtlanmaktadır. Ardından bu analitik öngörü, GPT-2 ailesinin üç farklı ölçeğinde (Small, Medium, Large), dört farklı deneysel kolda (Taban, Gaussian, Kuantum, Hibrit) gerçekleştirilen 108.000 metin üzerindeki kapsamlı ampirik değerlendirme ile teyit edilmektedir. Kuantum bileşeninin, iptal mekanizmasından bağımsız olarak izole test edildiği Quantum-only kolunda Base modele kıyasla ölçülebilir yapısal kaymalar ürettiği; ancak bu kaymaların yaratıcılık vekil (proxy) metriklerinde sistematik bir üstünlüğe dönüşmediği gösterilmektedir. Son olarak, otomatik yaratıcılık proxy metriklerinin bu ince yapısal değişimleri algılamadaki kırılganlığını ortaya koymakta ve teoriden öngörülemeyen bağımsız bir ampirik katkı sunmaktadır.

Tezin kapsamı dahilinde, BAW mekanizması GPT-2 model ailesi üzerinde test edilmiş; geleneksel taban (baseline) yöntemi ile Gaussian, Kuantum ve bunların bileşimi olan Karışık (Hybrid) BAW konfigürasyonlarının sonuçları karşılaştırmalı olarak analiz edilmiştir. Üretilen metinlerin değerlendirmesi, metinsel sözcüksel çeşitlilik ölçüsü (MTLD), bağımlılık mesafesi (dependency distance), tutarlılık skoru (coherence score) ve görelî ikidilli değerlendirme (Self-BLEU) gibi çeşitli metrikler üzerinden gerçekleştirilmiştir. Çalışma, dikkat modülasyonlarının yapısal sınırlarını dürüst bir matematiksel şeffaflıkla ele almayı ve gelecekte özellikle toplamsal (additive) birleştirme ile katman-seçici uygulama alanlarında sağlam bir kavramsal zemin sunmayı hedeflemektedir.

2. LİTERATÜR BİLGİLERİ

2.1 Temel Kavramlar ve Literatür Özeti

Bu bölümde, tezin kavramsal altyapısını oluşturan Transformer mimarisi, öz-dikkat (self-attention) mekanizması, büyük dil modellerinin (LLM) yapısı, yaratıcılık-tutarlılık ödünleşimi ve değerlendirme metrikleri incelenmektedir. Ayrıca, önerilen yöntem matematiksel bir topoloji sağlayan kuantum rotasyon kapılarının teorik temelleri sunulmaktadır.

2.1.1 Transformer Mimarisi ve Öz-Dikkat Mekanizması

Büyük dil modellerinin temel yapıtaşı olan Transformer mimarisi, ardışık veri işleme problemlerinde köklü bir paradigma değişimi yaratmıştır (Vaswani vd. 2017). Mimari, klasik tekrarlayan (recurrent) ağların aksine tüm veri dizisini paralel olarak işleyebilme yeteneğine sahiptir. Bu yetenek, "öz-dikkat" (self-attention) mekanizması ile sağlanır. Öz-dikkat, girdi dizisindeki her bir kelimenin (token), dizideki diğer tüm kelimelerle olan bağlamsal ilişkisini dinamik bir ağırlık matrisi üzerinden hesaplar. Girdi matrisi $X \in \mathbb{R}^{N \times d_{model}}$ olmak üzere, Sorgu (Q), Anahtar (K) ve Değer (V) matrisleri, öğrenilebilir ağırlık matrisleri (W^Q, W^K, W^V) aracılığıyla doğrusal izdüşümlerle oluşturulur:

$$Q = XW^Q, \quad K = XW^K, \quad V = XW^V \quad (2.1)$$

Matematiksel olarak standart ölçeklendirilmiş iç-çarpım dikkati (scaled dot-product attention) şu şekilde hesaplanır:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (2.2)$$

Burada d_k , anahtar vektörlerinin boyutunu temsil eder. $\frac{1}{\sqrt{d_k}}$ ölçeklendirme faktörü, iç-çarpım sonuçlarının varyansının büyümesini engelleyerek softmax fonksiyonunun gradyanlarının sönümlenmesini (vanishing gradients) önler (Vaswani vd. 2017).

Softmax fonksiyonu, ham dikkat skorlarını (logit) 0 ile 1 arasında normalize ederek bir

olasılık dağılımı oluşturur. Ancak eksponansiyel fonksiyonun yapısı gereği, yüksek skorlu tokenların dikkat dağılımında aşırı baskın hale gelmesine neden olabilir. Bu durum, metnin yakın bağlamsal tutarlılığını güçlendirirken, düşük skorlu fakat anlamsal açıdan yaratıcı ve sürpriz barındıran uzak bağlamsal ilişkilerin etkisini sınırlayabilmektedir. Bu yapısal "kazanan-hepsini-alır" eğilimi, tezin ele aldığı yaratıcılık-tutarlılık ödünleşimi probleminin mekanistik zeminini oluşturmaktadır.

2.1.2 Büyük Dil Modelleri ve GPT-2 Mimari Ailesi

Büyük dil modelleri (LLM), devasa veri setleri üzerinde eğitilmiş, kendinden önceki kelimelerin bağlamına dayanarak bir sonraki kelimeyi tahmin etme (otoregresif dil modellemesi) prensibiyle çalışan derin öğrenme ağlarıdır. Üretken Ön-eğitilmiş Dönüştürücü (Generative Pre-trained Transformer) ailesinin önemli bir üyesi olan GPT-2 (Radford vd. 2019), yalnızca kod çözücü (decoder-only) bir Transformer yapısı kullanır.

GPT-2 açık ağırlıklı olması, tamamen tekrarlanabilir deney ortamı sunması ve mimari olarak sade (ek karmaşık modüller barındırmayan) bir yapıya sahip olması nedeniyle, statik dikkat modülasyonu gibi temel mekanistik müdahalelerin deneysel sınaması açısından son derece uygun bir açık ağırlıklı temel model (baseline) platformudur. Bu çalışmada önerilen yöntem, modülasyon etkisinin parametre sayısı ile nasıl değiştiğini (ölçek etkisini) inceleyebilmek amacıyla GPT-2'nin üç farklı varyantı üzerinde değerlendirilmiştir:

- **GPT-2 Small:** 124 Milyon parametre, 12 dikkat katmanı.
- **GPT-2 Medium:** 355 Milyon parametre, 24 dikkat katmanı.
- **GPT-2 Large:** 774 Milyon parametre, 36 dikkat katmanı.

2.1.3 Yaratıcılık-Tutarlılık Ödünleşimi

Hesaplamalı yaratıcılık (computational creativity) ve yaratıcılık kuramı bağlamında yaratıcılık; yalnızca rastgelelik, düzensizlik veya düşük olasılıklı token seçimi olarak tanımlanamaz. Literatürde anlamsal yaratıcılık, bir fikrin veya çıktının hem özgün (origi-

nal) hem de bağlama uygun/yararlı (useful/appropriate) olmasını gerektirir (Boden 2004, Runco ve Jaeger 2012). Dil üretiminde bu durum; kavramsal uzaklık (conceptual distance), yenilikçi metafor üretimi, farklı anlamsal alanlar (domains) arası ilişki kurabilme ve bu esnada dilsel/anlamsal bütünlüğü koruma kapasitesi olarak karşılık bulur. Tutarlılık (coherence) ise, metnin dilbilgisel yapısının, anlamsal akışının ve yerel bağlamsal sürekliliğinin bozulmadan devam etmesidir.

Büyük dil modellerinde kelime çeşitliliğini artırmaya yönelik sıradan müdahaleler genellikle anlamsal kopukluklarla (halüsinasyon) sonuçlanarak tutarlılık kaybına yol açar. Buna karşın, tutarlılığı en üst düzeye çıkaran standart algoritmalar oldukça sıradan, mekanik ve tekrarlı metinler üretme eğilimindedir. Bu durum "yaratıcılık-tutarlılık ödünleşimi" (creativity-coherence trade-off) olarak adlandırılır. Tezin temel problemi, modelin tutarlılığını kabul edilebilir düzeylerde korurken, dikkat mekanizmasına yapılan yapısal bir müdahale ile anlamsal yaratıcılık unsurlarının teşvik edilip edilemeyeceğinin sınanmasıdır.

2.1.4 Kuantum Rotasyon Kapıları ve Born Kuralı

Klasik sistemlerin aksine, kuantum hesaplamada bilgi kubitler (qubits) üzerinden işlenir. Bir kubitin durumu, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ şeklinde bir süperpozisyon olarak tanımlanır. Kuantum durumları, üniter matrisler olan kuantum kapıları ile manipüle edilir. Bilgehan-Aslan Penceresi'nin (BAW) periyodik bileşeni, Bloch küresinde X-ekseni etrafında dönüşü temsil eden RX kapısının topolojisinden esinlenmiştir. Bir kübite θ açısıyla uygulanan RX kapısı şu şekildedir:

$$RX(\theta) = \begin{pmatrix} \cos(\theta/2) & -i \sin(\theta/2) \\ -i \sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \quad (2.3)$$

Sistemin belirli bir durumda ölçülme olasılığı Born Kuralı ile açıklanır. $|0\rangle$ başlangıç durumundaki bir kübite $RX(\theta)$ uygulandığında, sistemin tekrar $|0\rangle$ olarak ölçülme olasılığı $\cos^2(\theta/2)$ fonksiyonu ile ifade edilir.

Vurgulanmalıdır ki, bu çalışmada kuantum mekaniği fiziksel bir hesaplama aracı olarak değil, tamamen "kuantum-esinli analitik bir modülasyon" (quantum-inspired analytical modulation) olarak kullanılmıştır. Yöntem herhangi bir kuantum üstünlük (quantum supremacy) veya donanım başarısı iddia etmemektedir. Sadece $\cos^2(\theta/2)$ formunun 0 ile 1 arasında yaptığı periyodik dalgalanma yapısı, uzak token mesafelerinde periyodik ve mesafeye duyarlı bir dikkat maskesi (attention mask) inşa etmek için kavramsal bir matematiksel araç olarak değerlendirilmiştir.

2.1.5 Değerlendirme Metrikleri ve Otomatik Yaratıcılık Ölçümünün Sınırları

Çalışmada metin çıktılarının yapısal ve anlamsal özelliklerini gözlemlemek için çeşitli otomatik değerlendirme metrikleri kullanılmıştır:

- **Distinct-n ve Self-BLEU:** N-gram tabanlı eşsiz kelime öbeklerinin oranını ve çıktıların birbiriyle çakışma durumunu hesaplayan klasik metriklerdir (Li vd. 2016, Zhu vd. 2018).
- **MTLD (Measure of Textual Lexical Diversity):** Metin uzunluğundan bağımsız, standartlaştırılmış sözcüksel çeşitlilik ölçütüdür (McCarthy ve Jarvis 2010).
- **Coherence Score (Tutarlılık Skoru):** Bu çalışmada tutarlılık, ardışık cümlelerin Sentence-BERT/MiniLM tabanlı vektör gömülmeleri arasındaki kosinüs benzerliği üzerinden yaklaşık olarak ölçülmüştür. Sentence-BERT, cümle düzeyinde anlamsal benzerlik hesaplamaları için uygun gömülmeler üretmesi nedeniyle bu amaçla kullanılmıştır (Reimers ve Gurevych 2019).
- **Yapısal ve Tematik Metrikler:** Metindeki sözdizimsel karmaşıklığı analiz eden POS entropy ve dependency distance (bağımlılık mesafesi); içeriğin soyutluk düzeyini gösteren abstractness score; kavramlar arası sapmaları ölçen metaphor density ve cross-domain rate ile hikayenin duygusal dalgalanmasını belirten emotion arc variance metrikleri analiz edilmiştir.

Bu otomatik metrikler çoğunlukla kelime frekansları, sözdizimsel derinlik veya vektö-

rel mesafe üzerinden çalışan sayısal ölçüm araçlarıdır. İnsan algısındaki çok katmanlı, ironik veya metaforik yaratıcılığı doğrudan ve eksiksiz ölçmekte yetersiz kalabilirler. Bu tezin önemli katkılarından biri de, otomatik yaratıcılık metriklerinin yüzeysel, kırılmalı ve model davranışlarındaki ince yapısal değişimleri algılamada yetersiz kalabildiğini (örneğin MTLD ve Distinct-n arasındaki istatistiksel yığılmaları) deneysel olarak görünür kılmasıdır.

2.2 İlgili Çalışmalar ve Literatür İncelemesi

Bu bölümde, büyük dil modellerinde dikkat mekanizmasının iyileştirilmesi, örnekleme stratejileri, çıkarım-zamanı dikkat yönlendirme, yaratıcılık ölçümü, pozisyonel kodlama ailesi ve kuantum-esinli yaklaşımlar üzerine literatürdeki mevcut çalışmalar değerlendirilmekte; önerilen BAW yönteminin literatürdeki özgün konumu tartışılmaktadır.

2.2.1 Klasik Dikkat Mekanizması İyileştirmeleri

Transformer mimarisindeki öz-dikkat mekanizmasının karesel hesaplama karmaşıklığı ($O(N^2)$) problemini aşmak üzere literatürde birçok model önerilmiştir. Sparse Transformer (Child vd. 2019) tam matris yerine belirli örüntü bölgelerini hesaplar; Reformer (Kitaev vd. 2020) yerelliğe duyarlı özetleme (LSH) tekniğini kullanır. Linear Attention (Katharopoulos vd. 2020) ve Performer (Choromanski vd. 2021) çekirdek (kernel) fonksiyonları yardımıyla doğrusal karmaşıklığa ulaşmayı hedefler. Longformer (Beltagy vd. 2020) ve BigBird (Zaheer vd. 2020) modelleri seyrek pencereler kullanarak uzun bağlam (long-context) işleme kapasitesini artırır. Mixture-of-Experts (MoE) (Shazeer vd. 2017) mimarileri, her giriş için yalnızca belirli uzman alt ağlarını etkinleştirerek model kapasitesini artırırken hesaplama/parametre verimliliğini korumayı hedefler.

Bu çalışmaların temel hedefi genellikle hesaplama karmaşıklığını düşürmek, bellek verimliliğini artırmak veya uzun belgeleri işleyebilmektir. BAW'ın farkı şudur: BAW hesaplama maliyetini azaltmak için değil, dikkat skorlarının mesafe duyarlı statik bir modülasyonla

yapısal davranışını nasıl etkilediğini mekanistik olarak sınamak için önerilmiştir. Ayrıca yukarıdaki yöntemler genellikle sıfırdan eğitim (training) gerektirirken, BAW eğitimsiz (training-free) bir analitik modülatördür.

2.2.2 Örneklem Tabanlı Metin Üretimi ve Yaratıcılık Yöntemleri

Dil modellerinin üretim aşamasında çeşitliliği artırmak için sıcaklık ölçeklemesi, top- k ve top- p gibi örneklem stratejilerinden yararlanılır (Holtzman vd. 2020). Hikâye üretimi gibi açık uçlu üretim görevlerinde bu tür örneklem ayarlarının üretim davranışını etkilediği ayrıca gösterilmiştir (Fan vd. 2018). Daha gelişmiş metotlar arasında, bilgi-teorik beklenen değere yakın tokenları seçen Typical Sampling (Meister vd. 2023), uzman ve zayıf model logitlerini karşılaştıran Contrastive Decoding (Li, X. L. vd. 2023) ve DExperts (Liu vd. 2021) yer almaktadır.

Tüm bu yaklaşımlar post-hoc, yani model logitleri üretildikten sonra son çıktı dağılımı üzerinde çalışan yöntemlerdir. Modelin içsel dikkat geometrisini veya kelimeler arası ilişki ağırlıklarını doğrudan değiştirmezler. BAW'ın bu yöntemlerden temel farkı, müdahalenin çıktı aşamasında değil, henüz softmax uygulanmadan dikkat skorlarının hesaplandığı matrisin içinde yapılmasıdır. Ancak bu yapısal farklılık, otomatik olarak daha başarılı bir sonuç anlamına gelmez; nitekim mevcut tez bulguları, softmax öncesi statik müdahalenin metrikler üzerindeki faydasının sınırlı kaldığını göstermektedir. Önemli bir sınırlılık olarak, bu çalışmada BAW ile sıcaklık ölçeklemesi veya top- p gibi post-hoc yöntemler doğrudan deneysel olarak kıyaslanmamıştır; bu karşılaştırma gelecek çalışmalar için somut bir öneri olarak kaydedilmektedir.

2.2.3 Çıkarım-Zamanı Dikkat Yönlendirme (Inference-Time Attention Steering)

BAW, kendini “eğitimsiz, softmax öncesi dikkat modülasyonu” olarak konumlandırmaktadır. Bu ailenin en bilinen ve en yakın akrabası PASTA (Post-hoc Attention STeering Approach) (Zhang vd. 2023) olup, model ağırlıklarını değiştirmeden, çıkarım anında küçük bir dikkat-başlığı (attention head) alt kümesinde dikkat skorlarını yeniden ağırlıklandıran

bir yaklaşımdır. PASTA, kullanıcının belirttiği vurgu işaretlerine göre dikkat yeniden dağılımı yaparak belirli görevlerde anlamlı doğruluk artışları bildirmiştir. Yöntemi, BAW ile doğrudan karşılaştırıldığında, PASTA'nın *öğrenilmiş baş seçimi* (hangi dikkat başlıklarının yeniden ağırlıklandırılacağını küçük bir veri setinden öğrenmesi) ve *sorgu-bağımlı* yönlendirme yaptığı, BAW'ın ise *statik analitik mesafe maskesi* kullandığı görülmektedir.

LLMSteer (Gu vd. 2024), uzun-bağlam görevlerinde “kaybolmuş-ortada” (lost-in-the-middle) sorununu çözmek için sorgudan bağımsız dikkat yönlendirmesi yapan bir çerçeve sunmaktadır. Önceden önbelleğe alınmış (prefix-cached) KV-önbellek içindeki dikkat skorlarını dinamik olarak yeniden ağırlıklandırarak, ince-ayar gerektirmeden performans artışı sağlamıştır. BAW'ın LLMSteer'dan temel farkı, BAW'ın mesafe-tabanlı analitik bir maske kullanması, LLMSteer'ın ise bağlam içeriğine dayalı dinamik yeniden ağırlıklandırma yapmasıdır.

Komut mühendisliği (prompt engineering) ve Tak-Çalıştır Dil Modelleri (PPLM) (Dathathri vd. 2020), model ağırlıklarını doğrudan değiştirmeden üretim davranışını yönlendirmeye çalışan yaklaşımlar olarak değerlendirilebilir. Buna karşılık model düzenleme (model editing) çalışmaları, model parametrelerinde hedefli değişiklikler yaparak olgusal ilişkileri veya belirli davranış örüntülerini değiştirmeyi amaçlar (Meng vd. 2022).

Bu alt alandaki çalışmalar toplu olarak değerlendirildiğinde, BAW'ın konumlanması netleşmektedir: BAW, dikkat yeniden ağırlıklandırma ailesinin en basit (statik, analitik, mesafe-tabanlı) üyesidir. Bu basitlik bir avantaj (hiç öğrenme gerektirmez, deterministik) ve bir dezavantajdır (bağlam-duyarsız, içerik-agnostik).

2.2.4 Pozisyonel Kodlama ve Mesafe Yanlılığı Ailesi

Dikkat mekanizmasına konum bilgisi ekleme stratejileri, BAW'ın en yakın yapısal akrabalarıdır. Tarihsel perspektiften:

Toplamsal dikkat kökeni: Bahdanau ve arkadaşları (Bahdanau vd. 2015), sinir ağı ta-

banlı makine çevirisinde ilk toplamsal (additive) dikkat mekanizmasını önererek, kodlayıcı gizli durumlarına ileri-beslemeli ağ üzerinden hizalama skorları hesaplamıştır. Bu çalışma, dikkat kavramının modern NLP'ye girişini temsil etmektedir. Luong ve arkadaşları (Luong vd. 2015) ise çarpımsal (multiplicative/dot-product) dikkat ile toplamsal dikkati sistematik olarak karşılaştırarak, çarpımsal dikkatin hesaplamalı verimliliğini göstermiştir. BAW'ın çarpımsal maske tercihinin tarihsel kökeni bu ayrıma dayanmaktadır.

Göreceli konum kodlamaları: Relative Positional Encoding (Shaw vd. 2018), Transformer-XL (Dai vd. 2019) ve T5 Relative Bias (Raffel vd. 2020) yöntemleri, tokenlar arası göreceli mesafeye bağlı toplamsal konum ağırlıkları kullanır.

Modern pozisyonel yanlılıklar: ALiBi (Attention with Linear Biases) (Press vd. 2022), tokenlar arası mesafeye bağlı olarak doğrusal ve toplamsal bir mesafe cezası uygular; eğitim sırasında görülenden daha uzun dizilere genelleme yapabilme kapasitesiyle öne çıkar. RoFormer/RoPE (Su vd. 2024) konum bilgisini rotasyonel matrisler üzerinden iç çarpım uzayına entegre eder. RoFormer/RoPE ilk olarak 2021 yılında arXiv önbaskısı olarak sunulmuş, daha sonra dergi sürümüyle literatürde yer almıştır. Bu nedenle burada RoPE'nin tarihsel çıkışı 2021 önbaskısına, bibliyografik kayıt ise kullanılan kaynakça sürümüne göre değerlendirilmiştir. Kerple (Chi vd. 2022), ALiBi'nin doğrusal cezasını çekirdek (kernel) fonksiyonları ile genelleştirerek üstel ve kuvvet-yasası sönümleme seçenekleri sunar. xPos (Sun vd. 2023), RoPE'nin uzun-dizi bozulma sorununu çözmek için üstel sönümleme ekler; bu yaklaşım, BAW'ın Gaussian sönümleme bileşeniyle kavramsal benzerlik taşır. FIRE (Functional Interpolation for Relative Positional Encoding) (Li, S. vd. 2023), göreceli konum bilgisini öğrenilebilir bir sürekli fonksiyon ile modellemenin avantajlarını göstermiştir.

BAW'ın bu aile içindeki konumu şöyledir: ALiBi, RoPE ve diğer modern yaklaşımlar toplamsal (additive) mesafe cezaları veya rotasyonel dönüşümler kullanırken, BAW çarpımsal (multiplicative) bir maske ile periyodik bir mesafe fonksiyonunu birleştirir. Bölüm 3.2'deki analiz, çarpımsal formülasyonun uzak-menzil rezonansı yapısal olarak iptal ettiğini göstermekte; toplamsal bir alternatifi (Bölüm 3.2.5) bu sorunu doğası gereği çözeceğini öne

sürmektedir.

2.2.5 Yaratıcılık Ölçümü Literatürü

Büyük dil modellerinin “yaratıcılığını” ölçmek, alanın en tartışmalı metodolojik sorunlarından biridir. Bu çalışmada kullanılan metaphor_density ve cross_domain_rate gibi sayaç-tabanlı proxy metrikler, yaratıcılığın yalnızca yüzeysel göstergelerini yakalayabilmektedir. Literatürde daha sağlam alternatifler geliştirilmiştir:

Divergent Semantic Integration (DSI): Johnson ve arkadaşları (Johnson vd. 2023), anlatı metinlerinden yaratıcılık çıkaran, dağılımsal anlamsal modellemeye dayalı bir ölçüt geliştirmiştir. Johnson vd. (2023), DSI’nin anlatı metinlerinde insan yaratıcılık değerlendirmeleriyle güçlü ilişkiler gösterebildiğini ve belirli görevlerde BERT tabanlı en iyi modelin insan yaratıcılık puanlarındaki varyansın %72’ye kadarını açıklayabildiğini bildirmiştir. Bu çalışmada Johnson vd.’nin özgün DSI protokolü birebir uygulanmamış; yalnızca DSI’den esinlenen cümle-düzeyi metin-içi anlamsal yayılım metriği ikincil doğrulayıcı analiz olarak kullanılmıştır. Gelecek çalışmalarda mevcut metin havuzunun tam DSI ile retroaktif puanlanması önerilmektedir.

Divergent Association Task (DAT): Olson vd. (2021) tarafından önerilen Divergent Association Task (DAT), bireylerin ürettiği ilişkisiz kelimeler arasındaki anlamsal uzaklığı ölçerek ıraksak düşünme kapasitesini değerlendiren kısa ve uygulanabilir bir yaratıcılık ölçütü olarak sunulmuştur. Çalışmada DAT’nin geleneksel yaratıcılık ölçütleriyle anlamlı ilişkiler gösterdiği bildirilmiştir.

LLM yaratıcılık kıyaslamaları: Bellemare-Pepin vd. (2024), DSI ile Lempel-Ziv algoritmik karmaşıklığını haiku, sinopsis ve flash-fiction görevlerinde kullanarak insan-LLM yaratıcılık karşılaştırması yapmıştır. Çalışma, LLM’lerin ortalamada insanlardan daha yüksek DSI skorları ürettiğini ancak üst-çeyreklik insan yaratıcılığına ulaşamadığını göstermiştir. Nath vd. (2024), insanlar ve büyük dil modellerinde yaratıcı sürecin nasıl karakterize edilebileceğini inceleyerek, LLM yaratıcılığı değerlendirmelerinde yalnızca nihai

ürünlerin değil, üretim sürecinin ve anlamsal arama örüntülerinin de dikkate alınması gerektiğini göstermektedir. Bu çalışma, BAW'ın doğrudan deneysel sonucu için değil, LLM yaratıcılığı değerlendirmesinin süreç-temelli ve çok boyutlu ele alınması gerektiğini göstermesi bakımından bağlamsal bir destek sunmaktadır. Öte yandan, Lu vd. (2024) tartışma ve rol-yapma (role-play) çerçeveleri aracılığıyla LLM'lerin yaratıcılığını artırmanın yollarını araştırmıştır.

Bu literatür, BAW çalışmasının metrik seçimindeki zayıflığı bağlamlandırmaktadır: sayaç-tabanlı proxy'ler ile ölçülen “null” sonucun, DSI gibi doğrulanmış metriklerle de teyit edilip edilmeyeceği açık bir araştırma sorusudur.

Not: Bu bölümde anılan bazı güncel çalışmalar (örn. LLMSteer, FIRE, LLM yaratıcılık kıyaslamaları, Characterising the Creative Process, LLM Discussion ve QASA) arXiv ön baskısı niteliğindedir. Bu nedenle söz konusu kaynaklar yerleşik ve kesin literatür kanıtı olarak değil, güncel araştırma eğilimlerini ve olası yöntemsel yönelimleri gösteren destekleyici çalışmalar olarak değerlendirilmiştir.

2.2.6 NLP’de Kuantum ve Kuantum-Esinli Yaklaşımlar

Doğal Dil İşleme alanında kuantum tabanlı araştırmalar giderek artmaktadır. Bu çalışmalar, kelimelerin kuantum dalga fonksiyonlarıyla temsil edildiği kuantum-esinli gömüler (quantum-inspired embeddings) (Li vd. 2018) ve metinlerin yoğunluk matrisleri (density matrices) kullanılarak modellendiği kuantum-benzeri dil modellerini (Zhang vd. 2018) içerir. Öte yandan, doğrudan gerçek kuantum donanımları üzerinde çalıştırılmak üzere tasarlanan Varyasyonel Kuantum Devreleri (Chen vd. 2020) ve QNLP (Coecke vd. 2020) mimarileri de literatürde yer almaktadır. Chen vd. (2020), varyasyonel kuantum devrelerinin kuantum makine öğrenmesi bağlamında farklı görevlerde kullanılabileceğini göstermesi bakımından genel arka plan sunmaktadır; bu kaynak BAW'ın NLP bağlamındaki performansına doğrudan kanıt olarak kullanılmamıştır.

Kuantum-esinli dikkat mekanizmaları: BAW ile kavramsal bazı benzerlikler taşıyan

güncel bir örnek, Quantum Adaptive Self-Attention (QASA) (Chen ve Kuo 2025) modelidir. QASA, Transformer-benzeri bir mimaride son kodlayıcı bloğuna kuantum-esinli bir dikkat bileşeni yerleştiren hibrit bir yaklaşımdır. QASA, kuantum-esinli bileşenlerin Transformer-benzeri yapılarda katman konumlandırması açısından araştırılabileceğine dair analogik ve destekleyici bir örnek sunmaktadır. Bununla birlikte QASA, BAW için doğrudan deneysel öncül veya NLP metin üretimi bağlamında doğrudan kanıt olarak değerlendirilmemelidir.

Burada, gerçek kuantum donanımı kullanan yaklaşımlar ile kuantum matematiğinden esinlenen klasik yöntemler arasındaki fark netleştirilmelidir. BAW yöntemi, kesinlikle bir kuantum bilgisayar kullanmamakta veya fiziksel bir kuantum üstünlük (quantum supremacy) sağlamamaktadır. BAW'ın kuantum yönü, Born kuralındaki periyodik olasılık formundan esinlenen ve klasik işlemciler üzerinde çalışan tamamen analitik bir mesafe modülasyonundan ibarettir.

2.2.7 Literatürdeki Boşluk ve Bu Tezin Konumu

Literatür incelendiğinde; hesaplama verimliliği odaklı (sparse/efficient attention) modellerin, çıktı dağılımını filtreleyen örnekleme stratejilerinin, toplamsal konumsal yanlılıkların (örn. ALiBi, RoPE, Kerple), çıkarım-zamanı dikkat yönlendirme yaklaşımlarının (PASTA, LLMSteer) ve kuantum-esinli NLP yaklaşımlarının geniş bir alan kapladığı görülmektedir.

Bu tez, BAW'ın kesin bir performans üstünlüğü veya evrensel bir iyileştirme sağladığını iddia etmekten ziyade; eğitimsiz statik dikkat modülasyonlarının büyük dil modellerinde yapısal sınırlarını — özellikle çarpımsal formülasyondaki menzil-uyumsuzluğunu — mekanistik olarak analiz etmektedir. Tezin özgün katkısı üç boyutta konumlanmaktadır: (1) çarpımsal maskenin kuantum rezonansını yapısal olarak iptal ettiğinin formel ispatı, (2) dört-kollu deneysel tasarımla bileşen izolasyonu, ve (3) otomatik yaratıcılık metriklerinin kırılabilirliğinin ampirik kanıtlanması. İncelenen literatür kapsamında, bu üç boyutu aynı deneysel ve analitik çerçevede birleştiren doğrudan bir çalışmaya rastlanmamıştır.

3. MATERYAL ve METOT

3.1 Yöntem: Bilgehan-Aslan Penceresi (BAW)

Bu çalışma, dikkat mekanizmasındaki yaratıcılık ve tutarlılık ödünleşimini ele almak üzere "Bilgehan-Aslan Penceresi" (BAW) olarak adlandırılan statik bir dikkat maskeleye yöntemi önermektedir. BAW, modelin öz-dikkat hesaplama sürecine deterministik bir yapısal yönlendirme ekleyerek, bir yandan yakın bağlamdaki sözdizimsel tutarlılığı korumayı, diğer yandan ise uzak bağlamlara periyodik atılımlar yaparak yenilikçi kelime seçimlerini teşvik etmeyi amaçlamaktadır.

3.1.1 Genel Mimari ve Çarpımsal Modülasyon Tercihi

Geleneksel Transformer modellerinde dikkat skorları, maskeleye işlemi sonrasında softmax fonksiyonundan geçirilir. BAW yöntemi, bu aşamada devreye girerek dikkat skorlarına $(QK^T / \sqrt{d_k})$ softmax işlemi uygulanmadan hemen önce *çarpımsal* bir ağırlık maskesi (W) uygular. Çarpımsal bir yaklaşımın (ALiBi gibi toplamsal cezalara kıyasla) tercih edilme gerekçesi, dikkat logitlerinin doğal oransallığını korumaktır. Toplamsal bir ceza ($-b$), büyük dikkat skorlarını lineer olarak küçültürken, çarpımsal bir sönümleme ($\times W$) yüksek ihtimalli bağlantıları daha hassas bir şekilde ölçekler ve negatif logitlerin patlamasını önler.

Geliştirilen BAW ağırlığı, i çıktı tokenı ve j girdi tokenı arasındaki göreceli mesafeye ($d_{ij} = |i - j|$) bağlı olarak hesaplanır ve iki temel bileşenden oluşur: Gaussian (yerel kararlılık) ve Kuantum (periyodik uzaklık). Nihai BAW maskesi, bu iki bileşenin nokta bazlı (element-wise) çarpımıyla elde edilir:

$$W(i, j) = G(i, j; \sigma) \times Q(i, j; \theta_{scale}) \quad (3.1)$$

3.1.2 Gaussian Bileşen (Yerel Kararlılık)

Dilbilgisel yapıların ve cümlelerin yapıtaşlarının korunabilmesi için, birbirine anlamsal ve sözdizimsel olarak yakın kelimeler arasındaki güçlü bağların zayıflatılmaması gerekir. Gaussian bileşen, bu yerel kararlılığı sağlamak üzere tasarlanmıştır. Göreli mesafe d_{ij} arttıkça ağırlığı üstel olarak azaltan monotonik sönümleme fonksiyonu şu şekilde tanımlanır:

$$G(i, j; \sigma) = e^{-\frac{d_{ij}^2}{2\sigma^2}} \quad (3.2)$$

Bu fonksiyonun sınır değerleri incelendiğinde, $d_{ij} = 0$ (tokenın kendisine olan dikkati) durumunda $G(i, i) = e^0 = 1$ olur, yani lokal bağlam tam olarak korunur. Uzaklık arttıkça fonksiyon asimptotik olarak 0'a yaklaşır. İstatistiksel 3σ kuralı gereği, Gauss dağılımının kütlelerinin %99.7'si $[-3\sigma, +3\sigma]$ aralığında bulunur. Bu nedenle, BAW yönteminde σ parametresi "efektif dikkat penceresi"nin (effective attention window) genişliğini tanımlar. $d_{ij} > 3\sigma$ mesafesinden sonra $G(i, j)$ değeri ≈ 0.011 'in altına düşerek geleneksel bağlamı tamamen susturur. Düşük σ değerleri modelin dikkatini anlık (çok yakın) kelimelere hapsederken, yüksek σ değerleri daha geniş bir anlamsal bütünlüğün korunmasına izin verir.

3.1.3 Kuantum Bileşen (Periyodik Uzaklık Fonksiyonu)

Eğer sisteme yalnızca Gaussian bileşen uygulansaydı, modelin uzun mesafeli bağlamlardan beslenmesi tamamen engellenmiş olurdu. Modelin uzak kelimelerden "beklenmedik" ilişkiler çıkarabilmesini sağlamak için, fiziksel sistemlerin doğasından esinlenen bir periyodik uzaklık fonksiyonu geliştirilmiştir.

Bölüm 2'de açıklandığı üzere, $|0\rangle$ başlangıç durumundaki bir kübite X-ekseni etrafında dönüş uygulandığında (RX kapısı), sistemin aynı durumda ölçülme olasılığı Born Kuralı gereği $\cos^2(\theta/2)$ ile orantılıdır. Bu yöntemde, kuantum dönüş açısı θ , tokenlar arası göreli mesafe d_{ij} ve bir ölçeklendirme hiperparametresi olan θ_{scale} kullanılarak şu şekilde tanımlanmıştır:

$$\theta = \theta_{scale} \cdot d_{ij} \quad (3.3)$$

Gerçekleştirilen deneylerde ve geliştirilen algoritmanın üretim betiklerinde temel alınan nihai fonksiyonel form aşağıdaki gibidir:

$$Q(i, j; \theta_{scale}) = \cos^2 \left(\frac{\theta_{scale} \cdot d_{ij}}{2} \right) \quad (3.4)$$

Bu periyodik fonksiyon, mesafe arttıkça azalan Gaussian eğrisinin üzerinde dalgalanmalar yaratır. $\cos^2(x)$ fonksiyonunun periyodu π 'dir. Dolayısıyla, Kuantum bileşenin uzamsal periyodu T şu denklemlle türetilir:

$$\frac{\theta_{scale} \cdot T}{2} = \pi \implies T = \frac{2\pi}{\theta_{scale}} \quad (3.5)$$

Bu matematiksel özellik, maskenin $d \approx k \cdot \frac{2\pi}{\theta_{scale}}$ ($k \in \mathbb{Z}^+$) mesafelerinde “rezonans” noktalarına ulaşmasını, yani dikkat kontrastının bu periyodik noktalarda orijinal düzeyine yakın korunmasını ($Q \approx 1$) sağlar. Ara mesafelerde ise $Q < 1$ olduğundan dikkat kontrastı düzleştirilir. Böylece kuantum bileşeni, uzak mesafelerdeki token çiftlerinde dikkat kontrastını periyodik olarak yeniden ölçekler (periodic contrast rescaling). Burada vurgulanması gereken kritik nokta, çarpımsal $W \in [0, 1]$ maskesinin bir logit'i orijinal değerinin üzerine çıkarma kapasitesinin bulunmamasıdır — yalnızca sıfıra doğru çekebilir (bkz. Bölüm 3.2.3). Dolayısıyla maskenin etkisi “uzak tokenlara dikkat atılımı” (attention burst) değil, mesafeye bağlı bir kontrast geçitlemesidir (contrast gating). Öte yandan, “kuantum” etiketinin potansiyel bir zayıf nokta oluşturabileceği belirtilmelidir. Born kuralı yalnızca periyodik formun kavramsal motivasyonudur; matematiksel olarak bu bileşen mesafeye bağlı periyodik bir çekirdekten (yükseltilmiş kosinüs / raised cosine) ibarettir ve raporlanan tüm sonuçlar bu kuantum yorumundan bağımsızdır.

Bu teorik davranışın somut etkisini göstermek amacıyla, $\sigma = 4.0$ ve $\theta_{scale} = 0.10$ olarak sabitlenmiş bir konfigürasyonda, BAW çarpımsal maskesinin (W) farklı d_{ij} mesafelerinde aldığı analitik (formüsel) değerler Çizelge 3.1'de sunulmuştur. Bu çizelge ampirik bir deney sonucu değil, salt modülasyon formülünün matematiksel davranışının topolojik bir projeksiyonudur.

3.1.4 Uygulama Konfigürasyonu ve Katman-Seçicilik Tartışması

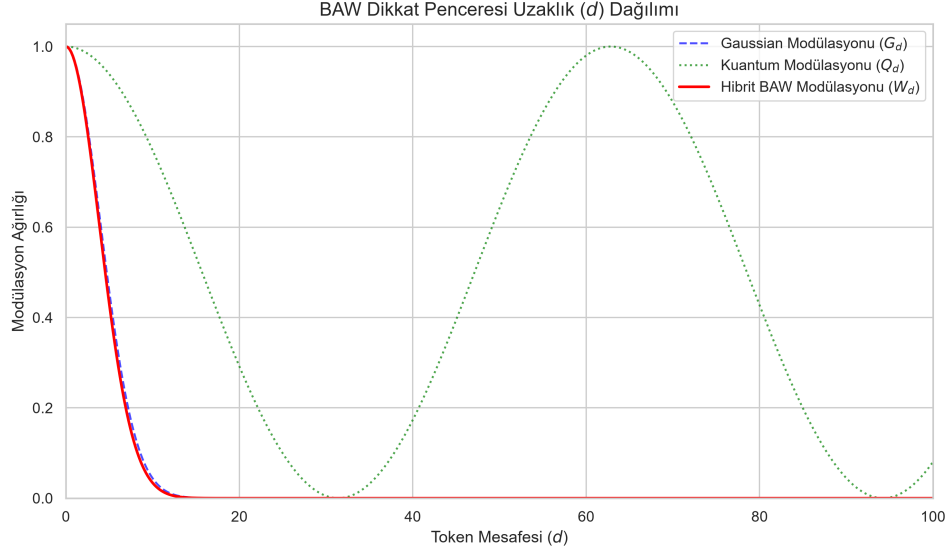
Dikkat analizi literatürü, Transformer tabanlı modellerde dikkat başlıklarının farklı katmanlarda farklı dilsel ve bağlamsal örüntülere duyarlı olabileceğini düşündürmektedir (Clark vd. 2019). BAW mekanizmasının tasarımı sırasında, modülasyonun yalnızca üst katmanlara uygulanmasının alt katmanlardaki temel sözdizimsel işleyişi koruyacağı öngörülmüştür.

Test edilen konfigürasyon: Bu tez çalışmasında sunulan 108.000 metinlik deney havuzunun tamamı, BAW modülatörünün $depth = 1$ konfigürasyonu ile elde edilmiştir. Burada $depth = 1$, modülasyonun tek bir katmana değil, modelin *tüm katmanlarına* sabit bir statik parametreyle uygulandığı konfigürasyonu ifade eder. Bu, yukarıda tanımlanan katman-seçici ideal tasarımdan farklı, basitleştirilmiş bir sürümdür. Basitleştirmenin gerekçesi, modülasyonun mekanistik saf etkisini (baseline divergence) tüm ağ üzerinde izole olarak gözlemlene amacıdır.

Yöntem–deney uyumsuzluğunun şeffaf kabulü: Tezin önerdiği ideal katman-seçici tasarım ile ampirik olarak test edilen tüm-katmanlı basitleştirilmiş konfigürasyon arasındaki fark açıkça kabul edilmiştir. Test edilen konfigürasyon “katman-seçici BAW” değil, “tüm katmanlara uygulanan basitleştirilmiş BAW”dır. Katman-seçici uygulama — hangi katman alt kümesinin (erken, orta, geç) modülasyona en duyarlı olduğunu belirleyen bir ızgara araması (grid search) veya dinamik katman yönlendirmesi (dynamic layer routing) — bu çalışmanın kapsamı dışında bırakılmış olup gelecek çalışmalara somut bir araştırma hedefi olarak önerilmektedir. QASA çalışması (Chen ve Kuo 2025), kuantum-esinli bileşenlerin Transformer-benzeri yapılarda katman konumlandırması açısından araştırılabileceğine dair analogik ve destekleyici bir örnek sunmaktadır. Bununla birlikte QASA, BAW için doğrudan deneysel öncül veya NLP metin üretimi bağlamında doğrudan kanıt olarak değerlendirilmemelidir. Bu bulgu, BAW için katman-seçici uygulamanın gelecekte ayrıca sınanması gereken makul bir araştırma yönü olduğunu düşündürmektedir.

Deneylerde kullanılan hiperparametre uzayı, Gaussian bileşen için $\sigma \in \{3.5, 4.0, 4.5\}$ ve

Kuantum bileşen için $\theta_{scale} \in \{0.05, 0.10, 0.15, 0.20, 0.25\}$ değerlerini içerecek şekilde tasarlanmıştır. Bu değerler, modelin tutarlılığı tamamen kaybetmesini önleyecek, ancak yaratıcılık metriklerinde ölçülebilir değişimler yaratabilecek ölçeklerde ampirik olarak belirlenmiştir.



Şekil 3.1 BAW dikkat penceresinin uzaklık (d) bazlı modülasyon ağırlıkları (W_d). Kuantum ve Gaussian bileşenlerin çarpımsal birleşimi (Hibrit) kırmızı çizgi ile gösterilmiştir.

3.2 Teorik Analiz: Çarpımsal Maskenin Matematiksel Yapısı

Bu bölümde, BAW maskesinin çarpımsal formülasyonunun ($W = G \times Q$) matematiksel yapısı sistematik olarak analiz edilmektedir. Öncelikle, Gaussian ve kuantum bileşenlerinin menzil uyumsuzluğunun hibrit maskede kuantum rezonansını yapısal olarak iptal ettiği formel bir önerme ile ispatlanmaktadır. Ardından aynı sonucun spektral (Fourier) alanda zarif bir ikinci ispatı sunulmaktadır. Son olarak, bu iptal problemini aşabilecek toplamsal (additive) bir alternatif formülasyon analiz edilmektedir.

3.2.1 Önerme 1: Menzil-Uyumsuzluğu ile Kuantum İptali

Kapsam notu: Bu önerme yalnızca çarpımsal hibrit maske $W(d) = G(d; \sigma) \times Q(d; \theta_{scale})$ için geçerlidir. Quantum-only kolunda ($W(d) = Q(d; \theta_{scale})$) Gaussian çarpanı bulunmadığından, buradaki iptal mekanizması uygulanmaz (bkz. §3.2.2).

Önerme 3.1 (Menzil-Uyumsuzluğu İptali). BAW hibrit maskesinde kuantum bileşenin ilk rezonans zirvesi $d^* = 2\pi/\theta_{scale}$ mesafesinde gerçekleşir. Bu zirvede kuantum bileşeni $Q(d^*) \approx 1$ ile tam güçteyken, efektif hibrit ağırlık doğrudan Gaussian sönümleme tarafından belirlenir:

$$G(d^*) = \exp\left(-\frac{(2\pi/\theta_{scale})^2}{2\sigma^2}\right) = \exp\left(-\frac{2\pi^2}{\theta_{scale}^2 \sigma^2}\right) \quad (3.6)$$

Rezonansın “görünür” olması için $G(d^*) \geq \varepsilon$ eşliğinin aşılması gerekir. Bu koşul şu eşitsizliğe dönüşür:

$$\theta_{scale} \cdot \sigma \geq \pi \sqrt{\frac{2}{\ln(1/\varepsilon)}} \quad (3.7)$$

Kanıt $G(d^*) \geq \varepsilon$ koşulundan başlayarak:

$$\exp\left(-\frac{2\pi^2}{\theta_{scale}^2 \sigma^2}\right) \geq \varepsilon \quad (3.8)$$

$$-\frac{2\pi^2}{\theta_{scale}^2 \sigma^2} \geq \ln \varepsilon \quad (3.9)$$

$$\frac{2\pi^2}{\theta_{scale}^2 \sigma^2} \leq \ln(1/\varepsilon) \quad (3.10)$$

$$\theta_{scale}^2 \sigma^2 \geq \frac{2\pi^2}{\ln(1/\varepsilon)} \quad (3.11)$$

$$\theta_{scale} \cdot \sigma \geq \pi \sqrt{\frac{2}{\ln(1/\varepsilon)}} \quad \square$$

Bu eşitsizliğin pratik sonuçları Çizelge 3.2’de özetlenmiştir.

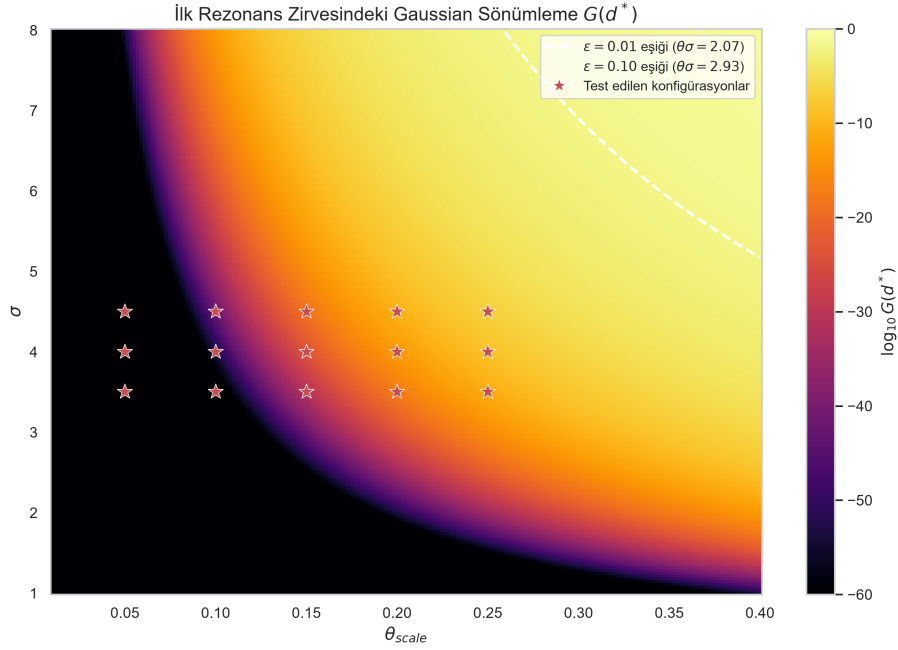
Deneyssel uzaya uygulama. Bu çalışmada test edilen hiperparametre ızgarası $\theta_{scale} \in \{0.05, 0.10, 0.15, 0.20, 0.25\}$ ve $\sigma \in \{3.5, 4.0, 4.5\}$ olmak üzere toplam 15 konfigürasyondan oluşmaktadır. Bu ızgarada $\theta_{scale} \cdot \sigma$ çarpımının minimum ve maksimum değerleri:

$$\theta_{scale} \cdot \sigma \in [0.05 \times 3.5, 0.25 \times 4.5] = [0.175, 1.125] \quad (3.12)$$

En müsamahakâr eşik olan $\varepsilon = 0.01$ (%1 görünürlük) bile $\theta_{scale} \cdot \sigma \geq 2.07$ gerektirmektedir. Deneyssel uzayın en agresif noktası ($\theta_{scale} = 0.25$, $\sigma = 4.5$, çarpım = 1.125)

bu eşiğin yaklaşık yarısı düzeyindedir (çarpım 1.125; eşik 2.07); en muhafazakâr nokta ($\theta_{scale} = 0.05$, $\sigma = 3.5$, çarpım = 0.175) ise eşiğin ~ 12 kat altındadır.

Sonuç: *Test edilen hiçbir konfigürasyonda kuantum bileşeninin uzak-menşil rezonansının hibrit maskede devreye girmesi matematiksel olarak mümkün değildir.* Bu, “bir konfigürasyonda iptal oldu” demekten çok daha güçlü bir sonuçtur: tüm deneysel uzay, fizibilitenin kategorik olarak dışındadır.



Şekil 3.2 $\theta_{scale} \times \sigma$ parametre uzayında ilk rezonans zirvesindeki Gaussian sönümlleme değeri $G(d^*)$. Renk skalası logaritmik; test edilen 15 ızgara noktası kırmızı yıldızlarla, $\epsilon = 0.01$ fizibilite eşik çizgisi ($\theta_{scale} \cdot \sigma = 2.07$) kesikli beyaz çizgiyle gösterilmiştir. Tüm deneysel noktalar fizibilite eşiğinin çok altında kalmaktadır.

3.2.2 Quantum-only Kolunda İptalin Neden Geçerli Olmadığı

Dört kollu deneysel tasarımın (Base, Gaussian, Quantum, Hybrid) amacı, bileşenleri izole olarak test etmektir. Quantum-only kolunda maske doğrudan:

$$W_Q(d) = Q(d; \theta_{scale}) = \cos^2\left(\frac{\theta_{scale} \cdot d}{2}\right) \quad (3.13)$$

olarak tanımlanır; Gaussian çarpanı yoktur. Bu durumda:

- Periyodik rezonans zirveleri ($d = k \cdot 2\pi/\theta_{scale}$, $k \in \mathbb{Z}^+$) tam güçle ($Q = 1$) tüm dizi boyunca tekrar eder.
- Düğüm noktaları ($Q = 0$) da tam olarak gerçekleşir.
- Kuantum modülasyonu, yakın ve uzak tüm mesafelerde aktiftir.

Bu, Önerme 1’in mantıksal kapsamını netleştirir: *iptal mekanizması yalnızca hibrit ($W = G \times Q$) formülasyona özgüdür*. “Kuantum bileşenini hiç test etmedik” eleştirisi geçersizdir — dört kollu tasarımın bütün anlamı, bileşenlerin ayrıştırılarak izole etkilerinin gözlemlenmesidir.

Bölüm 4’deki ampirik sonuçlar, Quantum-only kolunun Base modele kıyasla ölçülebilir yapısal kaymalar ürettiğini (örn. $d(Q-B) = +0.547$ bağımlılık mesafesinde, $d(Q-B) = -0.591$ MTLTD’de) ancak bu kaymaların yaratıcılık metriklerinde sistematik bir üstünlüğe dönüşmediğini göstermektedir. Bu gözlem, Önerme 1’in kapsamı dışında kalan bağımsız bir ampirik bulgudur ve “statik çarpımsal maskelerin, iptal bulunmasa bile, yaratıcılık metriklerinde iyileşme sağlamadığı” genel sonucunu desteklemektedir.

3.2.3 Çarpımsal Maskenin Logit Etkileşimi: Kontrast Düzleştirme

BAW maskesi $W(d) \in [0, 1]$ aralığında tanımlıdır ve softmax *öncesi* dikkat logitlerine çarpımsal olarak uygulanır:

$$s'(i, j) = W(d_{ij}) \cdot \underbrace{\frac{q_i \cdot k_j}{\sqrt{d_k}}}_{s(i, j)} \quad (3.14)$$

Bu çarpımın etkisi, ham logit’in işaretine bağlıdır:

- **Pozitif logit** ($s > 0$): $W \cdot s < s$ olduğundan logit küçülür; yani “ilgi çekici” tokenlara olan dikkat azaltılır.
- **Negatif logit** ($s < 0$): $W \cdot s > s$ olduğundan mutlak değer küçülür ve logit sıfıra yaklaşır; yani “ilgi çekmeyen” tokenlar nispeten güçlenir.

- $W = 0$: Tüm logitler sıfıra çekilir — softmax sonrası uniform dağılıma yaklaşılr.
- $W = 1$: Orijinal logit korunur — müdahale yok.

Dolayısıyla çarpımsal maskeleme, belirli mesafelerdeki token çiftleri arasındaki dikkat *kontrastını düzleştirir* (contrast flattening/gating). Bu davranış, “uzak tokenlara dikkat atılımları yapma” (attention bursts) metaforundan yapısal olarak farklıdır. Çarpımsal bir $W \in [0, 1]$ maskesi, bir logit’i orijinal değerinin üzerine çıkaramaz; yalnızca sıfıra doğru çekebilir. Sonuç olarak:

Çarpımsal maske, dikkat dağılımını yerel olarak daraltabilir (yüksek W bölgelerinde kontrastı koruyarak) veya düzleştirebilir (düşük W bölgelerinde kontrastı azaltarak), ancak hiçbir mesafede “yeni” bir dikkat odağı yaratamaz; yalnızca mevcut dikkat geometrisini yeniden ölçekleyebilir.

Bu karakterizasyon, Bölüm 4’de raporlanan sınırlı etki büyüklüklerinin yapısal bir açıklamasını sağlamaktadır: çarpımsal müdahale, modelin öğrenilmiş dikkat geometrisini pertürbe edebilir ama radikal olarak yeniden şekillendiremez. Ayrıca, tüm logitleri sabit bir katsayıyla çarpmak “sıcaklık (temperature) ölçeklemesi” iken, mesafeye bağlı W ile çarpmak esasında “konuma-bağlı bir sıcaklık ölçeklemesi”dir. BAW’ın mekanik olarak zaten bir sıcaklık varyantı olması, Bölüm 5.1’de tartışılan sıcaklık-eşli (temperature-matched) bir temel model (baseline) ile kafa-kafaya karşılaştırma yapılmasını teorik bir zorunluluğa çevirmektedir. Bu bağlantı, baseline eksikliğini bir ihmal olmaktan çıkarıp gelecek çalışmaların doğal ve zorunlu bir devamı olarak çerçevelemektedir.

3.2.4 Spektral Bakış: Hibrit Maskenin Fourier Ayrışması

Hibrit maskenin yapısını frekans uzayında incelemek, iptal mekanizmasının zarif bir ikinci ispatını sağlar. $\cos^2(x) = \frac{1+\cos 2x}{2}$ trigonometrik özdeşliği kullanılarak:

$$\begin{aligned} W(d) &= G(d) \cdot \cos^2\left(\frac{\theta_{scale} d}{2}\right) \\ &= G(d) \cdot \frac{1 + \cos(\theta_{scale} d)}{2} \\ &= \underbrace{\frac{1}{2}G(d)}_{\text{DC bileşen}} + \underbrace{\frac{1}{2}G(d) \cos(\theta_{scale} d)}_{\text{modüleli bileşen}} \end{aligned} \quad (3.15)$$

Bu ayrışma, hibrit maskenin iki bileşenden oluştuğunu göstermektedir:

1. **DC (sıfır-frekans) bileşen:** $\frac{1}{2}G(d)$ — yarı yoğunlukta saf Gaussian. Uzamsal olarak kompakt (3σ penceresi).
2. **Modüleli bileşen:** $\frac{1}{2}G(d) \cos(\theta_{scale} d)$ — Gaussian zarfla çarpılmış bir kosinüs dalgası.

Fourier dönüşümü perspektifinden, $G(d)$ uzamsal olarak geniş ($\sigma \gg 1$) olduğu için frekans uzayında *dar bir bant* kaplar. $\cos(\theta_{scale} d)$ ile çarpım, bu dar bandı $\pm\theta_{scale}$ frekanslarına kaydırır. Ancak σ yeterince büyük ve θ_{scale} yeterince küçük olduğunda, kaydırılmış bant hâlâ sıfır frekansı civarında yoğunlaşır — yani modüleli bileşen, DC bileşenin küçük bir pertürbasyonundan ibarettir.

Sonuç: Hibrit maske, fonksiyonel olarak “yarı yoğunlukta Gaussian + küçük bir kosinüs-modüleli Gaussian pertürbasyonu”ndan ibarettir. Bu, $G \times Q \approx G$ yaklaşıklığını kapalı formda kanıtlar ve Önerme 1’in spektral karşılığıdır.

3.2.5 Toplamsal (Additive) Alternatifin Formel Analizi

Önerme 1’de ortaya konan menzil-uyumsuzluğu, çarpımsal ($G \times Q$) formülasyonun yapısal bir sonucudur. Toplamsal (additive) bir birleştirme, bu sorunu doğası gereği ortadan

kaldırır. Maskeyi dikkat skorlarına çarpmak yerine, logit’e doğrudan toplamsal terimler eklendiğinde:

$$s'(i, j) = \frac{q_i \cdot k_j}{\sqrt{d_k}} - \frac{d_{ij}^2}{2\sigma^2} + \beta \cos^2\left(\frac{\theta_{scale} d_{ij}}{2}\right) \quad (3.16)$$

Bu formülasyonda üç kritik yapısal fark oluşur:

1. **Periyodik terim yaşar:** Kuadratik Gaussian penaltı $(-d^2/2\sigma^2)$ uzak mesafelerde logit’i baskılasa da, $\beta \cos^2(\cdot)$ terimi *eklenerek* katkıda bulunur — çarpımsal durumda olduğu gibi sıfırla çarpılarak yok edilmez. Dolayısıyla $d = d^* = 2\pi/\theta_{scale}$ rezonans noktasında logit’e $+\beta$ kadar periyodik bir “çıkıntı” eklenir.
2. **Logit’i artırabilir:** Çarpımsal $W \in [0, 1]$ maskesinin aksine, toplamsal $+\beta$ terimi pozitif bir logit’i orijinal değerinin *üzerine* çıkarabilir. Bu, softmax sonrası dikkat dağılımında uzak tokenlara gerçek bir “odak artışı” yaratma kapasitesine sahiptir.
3. **Aile benzerliği:** Denklem 3.16’deki ilk iki terim $(s(i, j) - d^2/2\sigma^2)$, kuadratik bir ALiBi varyantına karşılık gelir. Periyodik terimin eklenmesiyle “kuadratik ALiBi + periyodik bias” formülasyonu elde edilir.

Bu formel analiz, eğer toplamsal sürümde de negatif sonuç elde edilirse bunun çok daha bilgilendirici bir negatif sonuç olacağını göstermektedir: “mekanizma matematiksel olarak aktifti ama yaratıcılığı yine artırmadı” ile “mekanizma hiç devreye girmedi” arasında mekanistik bir fark vardır. Toplamsal formülasyonun ampirik sınaması, gelecek çalışmalar için somut bir yol haritası olarak önerilmektedir.

3.3 DeneySEL DüZEN ve Metodoloji

Bu bölümde, BAW mekanizmasının değerlendirilmesi için kurulan deneysel çatı, kullanılan dil modelleri, veri üretim süreçleri, algoritmik uygulama detayları ve istatistiksel analiz yöntemleri sunulmaktadır.

3.3.1 Kullanılan Modeller ve Ortam

Deneyler, GPT-2 mimarisinin üç farklı ölçeği üzerinde gerçekleştirilmiştir:

- **GPT-2 Small:** 124M parametre, 12 dikkat katmanı, 12 dikkat başlığı, $d_{model} = 768$.
- **GPT-2 Medium:** 355M parametre, 24 dikkat katmanı, 16 dikkat başlığı, $d_{model} = 1024$.
- **GPT-2 Large:** 774M parametre, 36 dikkat katmanı, 20 dikkat başlığı, $d_{model} = 1280$.

Tüm modeller HuggingFace Transformers kütüphanesinin `gpt2`, `gpt2-medium` ve `gpt2-large` kontrol noktalarından (checkpoints) yüklenmiştir. Hesaplama işlemleri PyTorch 2.1 derin öğrenme kütüphanesi kullanılarak NVIDIA RTX 5060 Ti (16 GB VRAM) GPU donanım hızlandırıcıları üzerinde paralel olarak yürütülmüştür. Kuantum-esinli modülasyon hesaplamaları için Python tabanlı özel tensör matris işlemleri optimize edilerek kullanılmıştır.

3.3.2 Veri Üretimi ve Örneklem Büyüklüğü

Modelin yaratıcılık sınırlarını test etmek amacıyla, 40 farklı ucu açık hikaye başlangıcı (prompt) kullanılmıştır. Çalışmada; Taban (Base), Gaussian, Kuantum ve Karışık (Hibrit) olmak üzere 4 farklı modülasyon stratejisi, 3 farklı model ölçeğinde (Small, Medium, Large) değerlendirilmiştir.

Üretim parametreleri: Çizelge 3.3'te tüm üretim parametreleri sunulmaktadır.

Örneklem yapısı ve 108.000 metin ilişkisi: İstatistiksel geçerliliği sağlamak ve bağımsızlık ihlali önlemek amacıyla, üretim aşamasında toplam 45 farklı rassal tohum (seed) konfigürasyonu kullanılmıştır. Her bir seed altında, $40 \text{ prompt} \times 5 \text{ üretim} = 200 \text{ metin}$ elde edilmiştir. Her seed–mekanizma–ölçek üçlüsü için bu 200 metnin metrik ortalaması alınarak makro düzeyde tek bir gözlem birimi oluşturulmuştur.

Toplam üretim hacmi: $45 \text{ seed} \times 4 \text{ mekanizma} \times 3 \text{ ölçek} \times 200 \text{ metin} = 108.000 \text{ metin}$. Ancak, **istatistiksel çıkarımsal birim** (inference unit) $n = 45$ 'tir (seed başına ortalama). 108.000 rakamı üretim hacmini ifade eder, istatistiksel gücü değil. Bu ayırım, bulguların yorumlanmasında kritik öneme sahiptir.

3.3.3 BAW Maskeleye Algoritması

BAW mekanizmasının çıkarım sırasında uygulanma süreci Algoritma 3.3.1'de özetlenmiştir.

Algoritma 3.3.1 BAW Çıkarım-Zamanı Dikkat Maskeleye

Girdi: Dikkat skoru matrisi $S = QK^T / \sqrt{d_k} \in \mathbb{R}^{N \times N}$; hiperparametreler σ, θ_{scale}

Çıktı: Maskelenmiş dikkat skoru matrisi $S' \in \mathbb{R}^{N \times N}$

1. Her (i, j) çifti için göreceli mesafe hesapla: $d_{ij} = |i - j|$
2. Gaussian bileşeni hesapla: $G(d_{ij}) = \exp(-d_{ij}^2 / 2\sigma^2)$
3. Kuantum bileşeni hesapla: $Q(d_{ij}) = \cos^2(\theta_{scale} \cdot d_{ij} / 2)$
4. Modülasyon seçimine göre maske oluştur:
 - Gaussian-only: $W(d_{ij}) = G(d_{ij})$
 - Quantum-only: $W(d_{ij}) = Q(d_{ij})$
 - Hibrit: $W(d_{ij}) = G(d_{ij}) \times Q(d_{ij})$
5. Maskeyi uygula: $S'(i, j) = W(d_{ij}) \cdot S(i, j)$
6. Softmax uygula: $A = \text{softmax}(S')$
7. Dikkat çıktısı: $\text{Output} = A \cdot V$

3.3.4 Kullanılan Değerlendirme Metrikleri

Üretilen metinlerin çok boyutlu analizi için 13 farklı metrik kullanılmıştır. Çizelge 3.4'te her metriğin tam tanımı ve hesaplama yöntemi sunulmaktadır.

Değerlendirme kütüphaneleri: Sözdizimsel ayrıştırma ve bağımlılık mesafesi analizleri için spaCy kütüphanesi, duygu yayı analizi için VADER, tutarlılık analizi için MiniLM tabanlı SentenceTransformers, sözcüksel çeşitlilik için lexical-diversity paketi kullanılmıştır. Çoklu değerlendirmeler, özel olarak geliştirilmiş QGPM (Quantum-Inspired Generative Pre-trained Modulator) değerlendirme boru hattı üzerinden otomatik olarak yürütülmüştür.

Ana 13 metriğe ek olarak, temel bulguları yapısal olarak doğrulamak amacıyla üç **ikincil doğrulayıcı metrik** (secondary confirmatory metrics) daha hesaplanmıştır:

- **Lempel-Ziv (LZ) Karmaşıklık:** Metinlerin algoritmik sıkıştırılabilirliğini ölçer. Üretilen metinlerin ilk 250 karakteri UTF-8 formatında zlib kütüphanesi ile sıkıştırılmış ve sıkıştırılmış bayt uzunluğunun orijinal bayt uzunluğuna oranı olarak hesaplanmıştır.
- **Konu Sapması (Topic Drift):** Modelin verilen başlangıç cümlesinden ne kadar uzaklaştığını ölçer. HuggingFace sentence-transformers/all-MiniLM-L6-v2 modeli ile prompt metni ve üretilen metin vektörleştirilmiş, aralarındaki kosinüs benzerliği 1'den çıkarılarak $(1 - \cos(\theta))$ hesaplanmıştır.
- **Metin-İçi Anlamsal Yayılım (DSI_SENTENCE):** Üretilen metnin kendi içindeki uzak kavramları harmanlama kapasitesini ölçer. spaCy kullanılarak metin cümlelere bölünmüş, her cümlelerin vektörleri arasındaki kosinüs benzerlik matrisinin üst üçgenindeki mesafelerin ortalaması alınarak hesaplanmıştır.

Toplam 16 otomatik metrik (13 temel, 3 ikincil doğrulayıcı) kullanılmıştır.

3.3.5 İstatistiksel Analiz ve Çoklu Karşılaştırma Düzeltmesi

Gruplar arası anlamlı farklılıkları tespit etmek için Tek Yönlü ANOVA ve etki büyüklüğünü ölçmek için Cohen's d analizi yapılmıştır. Analiz öncesinde varyans homojenliği Levene testi ile, normallik ise Shapiro-Wilk testi ile kontrol edilmiş; varsayımların sınırda ihlal edildiği spesifik dağılımlarda sonuçlar parametrik olmayan alternatiflerle (Kruskal-Wallis) teyit edilerek ANOVA'nın sağlamlığı güvence altına alınmıştır.

Çoklu karşılaştırma düzeltmesi: 13 farklı metrik üzerinde eş zamanlı test gerçekleştirildiğinden, Tip-1 hata oranını kontrol altında tutmak amacıyla iki farklı düzeltme yöntemi uygulanmıştır:

1. **Bonferroni düzeltmesi:** 13 metrik üzerinde eşzamanlı test yapıldığından düzeltilmiş anlamlılık eşiği $\alpha_{adj} = 0.05/13 = 0.00385$ 'tir. Klasik çoklu test mantığına uygun biçimde Bonferroni paydası doğrudan uygulanan test sayısı olarak belirlenmiştir. Standart uygulanan bu eşik altında anlamlı kalan hiçbir değer bulunamamıştır.
2. **Benjamini-Hochberg FDR düzeltmesi:** Korele metrikler içeren çoklu tarama (screening) senaryolarında Bonferroni'den daha uygun olan FDR (False Discovery Rate) düzeltmesi de raporlanmıştır. Bu yöntem, p -değerlerini küçükten büyüğe sıralayarak her birine $(i/m) \times q$ eşiğini uygular (burada $m = 13$ ve $q = 0.05$).

Her iki düzeltme yöntemi altında da hiçbir metriğin sağlam biçimde anlamlı kalmadığı gösterilmiştir (bkz. Bölüm 4).

Eşleştirilmiş yapı ve muhafazakâr yaklaşım: Aynı seed'den dört mekanizma üretildiğinden veriler teknik olarak eşleştirilmiş (within-seed) bir yapı taşımaktadır. Daha yüksek istatistiksel güç elde etmek için tekrarlı ölçümler ANOVA'sı (repeated-measures ANOVA) kullanılabilir. Bu çalışmada bağımsız gruplar ANOVA'sı tercih edilmiştir; bu seçim muhafazakâr bir yaklaşım oluşturmakta ve raporlanan null bulgularını güçlendirmektedir: seed-içi korelasyon pozitifse, bağımsız ANOVA hata terimini şişirir ve etki bulması zorlaşır. Gelecek çalışmalarda, seed'i rastgele etki (random effect) olarak alan bir doğrusal

karma model (LMM) koşulması ve sınıf-içi korelasyonun (ICC) raporlanması önerilmektedir.

3.3.6 Körleştirilmiş LLM-as-Judge İkincil Değerlendirmesi

Otomatik sayaç-tabanlı ve ikincil yapısal metriklerin ötesinde, üçüncü bir doğrulama katmanı olarak güçlü bir büyük dil modeli hakem (LLM-as-judge) olarak kullanılmıştır. Bu analiz, yöntemler arasında gerçek ve insan-algısına yakın bir kalite/yaratıcılık farkı olup olmadığını test etmek amacıyla tasarlanmıştır.

Değerlendirme, etkinin en belirgin gözlemlendiği ve niteliksel incelemenin yapıldığı **GPT-2 Small** ölçeği ile sınırlandırılmış; 40 prompt ve 1 rassal tohum (seed=1) kullanılarak toplam 200 çağrı (800 metin yargısı) yapılmıştır. Model olarak **GPT-4o** (gpt-4o-2024-05-13 API snapshot) kullanılmış ve tekrar-üretilebilirlik için sıcaklık (temperature) 0 olarak ayarlanmıştır.

Hakeme her çağrıda aynı (prompt, seed) bağlamında üretilmiş dört koldan (Base, Gaussian, Quantum, Hybrid) gelen metinler anonim olarak ("Text 1, Text 2, Text 3, Text 4") ve sırası rastgele karıştırılarak sunulmuştur. Yöntem isimleri kesinlikle gizlenmiştir. Hakemden her metni 4 kriter üzerinden 1-5 arası tam sayılarla puanlaması istenmiştir:

- Yaratıcılık (*creativity*): Orijinallik, şaşırtıcı ama anlamlı çağrışımlar.
- Tutarlılık (*coherence*): İç mantık, akış, dilsel bütünlük.
- Prompt'a Bağlılık (*prompt_fidelity*): Verilen başlangıç metnine sadakat.
- Genel Kalite (*overall_quality*): Okunabilirlik, etki, tamlık.

Ayrıca hakemden bu dört metni yaratıcılık bağlamında 1'den 4'e kadar sıralaması (creativity ranking) talep edilmiştir. Elde edilen sonuçlar, bağımsız ANOVA ve Cohen's *d* testleri ile analiz edilmiş; ek olarak pozisyon yanlılığını (sunum sırasının puana etkisi) tespit etmek amacıyla istatistiksel kontroller sağlanmıştır.

Çizelge 3.1 Belirli Mesafelerde BAW Maskesinin Analitik Değerleri ($\sigma = 4.0, \theta_{scale} = 0.10$)

Mesafe (d)	Gaussian (G)	Kuantum (Q)	Nihai $W(i, j)$	Matematiksel Durum
0	1.0000	1.0000	1.0000	Lokal tam odak ($G(i, i) = 1$)
5	0.4578	0.9381	0.4294	Gaussian sönümlene bölgesi
10	0.0439	0.7701	0.0338	Efektif pencere sonu ($\approx 2.5\sigma$)
31	7×10^{-14}	0.0005	≈ 0	Kuantum düğüm noktası (dip)
62	≈ 0	0.9990	≈ 0	Kuantum rezonans zirvesi ($k = 1$)

Çizelge 3.2 Farklı görünürlük eşikleri (ε) için gerekli minimum $\theta_{scale} \cdot \sigma$ çarpımı.

Görünürlük Eşiği (ε)	Gerekli $\theta_{scale} \cdot \sigma$	Fiziksel Anlam
0.50	≥ 5.34	Rezonans gücünün %50'sini korur
0.10	≥ 2.93	Rezonans ağırlığı %10
0.01	≥ 2.07	Rezonans ağırlığı %1
10^{-5}	≥ 1.31	Algılanamaz düzeyde zayıf

Çizelge 3.3 Metin üretiminde kullanılan sabit parametreler.

Parametre	Değer
Maksimum token uzunluğu (max_new_tokens)	120
Sıcaklık (temperature)	1.0 (varsayılan)
Top- k	50 (varsayılan)
Top- p (Nucleus)	1.0 (devre dışı)
Tekrar cezası (repetition_penalty)	1.0 (devre dışı)
Batch boyutu	5 (prompt başına)
Rastgele tohum (seed) sayısı	45

Çizelge 3.4 Çalışmada kullanılan 13 değerlendirme metriğinin tanımları ve hesaplama yöntemleri.

Metrik	Ölçtüğü Boyut	Hesaplama Yöntemi
Distinct-1	Sözcüksel çeşitlilik (unigram)	Eşsiz unigram sayısı / toplam unigram sayısı
Distinct-2	Sözcüksel çeşitlilik (bigram)	Eşsiz bigram sayısı / toplam bigram sayısı
Self-BLEU-3	Üretim tekrarı	Metinlerin birbirleriyle 3-gram BLEU skoru ortalaması
MTLD (McCarthy ve Jarvis 2010)	Sözcüksel çeşitlilik	Uzunluktan bağımsız sözcüksel çeşitlilik ölçümü
Gunning-Fog	Okunabilirlik	Cümle uzunluğu + çok heceli kelime oranı
Metaphor Density	Metaforik içerik	“gibi/like/as” edatları + fiil kökü geçiş sayımı
Cross-Domain Rate	Kavramsal sapma	Önceden tanımlı alan sözlüklerine çapraz geçiş oranı
Abstractness Score	Soyutluk düzeyi	Soyut kelime listesi ve morfolojik ek göstergeleri tabanlı oran
POS Entropy	Sözdizimsel çeşitlilik	Kelime türü (POS) dağılımının Shannon entropisi
Syntactic Complexity	Sözdizimsel karmaşıklık	Parse ağacı derinliği ortalaması
Dependency Distance	Bağımlılık mesafesi	Sözdizimsel bağımlılık kenarlarının ortalama uzunluğu
Emotion Arc Variance	Duygusal dalgalanma	VADER duygu skorlarının cümle-bazlı varyansı
Coherence Score	Anlamsal tutarlılık	Ardışık cümleler arası MiniLM kosinüs benzerliği

4. BULGULAR

Bu bölümde, 108.000 üretilmiş metin üzerinde gerçekleştirilen 13 farklı metriğin değerlendirme sonuçları, dört-parçalı bir bulgu çerçevesi içinde sunulmaktadır. Verilerin doğası gereği, aynı *seed* değerinden üretilen 200'er metinlik gruplar birer istatistiksel birim ($n = 45$) olarak ele alınmıştır (bkz. Bölüm 3.3.2). Elde edilen bulgular, beklentilerin aksine modülasyon mekanizmalarının tek yönlü bir “iyileştirme” sağlamadığını ortaya koymaktadır.

4.1 Bulgu A: Hiçbir Kol Yaratıcılık Proxy Metriklerinde Base’i Geçmiyor

Çalışmada önerilen BAW yönteminin tam hali olan Hibrit (Hybrid) modülasyon, “yaratıcılık” veya “anlamsal sapma” göstergesi olarak kabul edilen *cross_domain_rate*, *metaphor_density*, *abstractness_score* ve *coherence_score* gibi metriklerde hiçbir model ölçeğinde Taban (Base) modeli anlamlı ve büyük bir etkiyle geçememiştir. Aynı durum Gaussian-only ve Quantum-only kolları için de geçerlidir.

Bu geniş kapsamlı negatif sonuç, Bölüm 3.2’deki menzil uyumsuzluğu iptalinin doğrudan açıklaması değildir; çünkü iptal yalnızca Hibrit kola özgüdür. Bulgu A’nın doğru yorumu şudur: “Eğitimsiz statik çarpımsal dikkat maskeleri, iptal bulunmasa bile, yaratıcılık vekil metriklerinde sistematik bir iyileşme sağlamamıştır.”

4.2 Bulgu B: Hibrit Modülasyonda Sinerji Yokluğu

Hibrit yöntemin kendi alt bileşenleriyle (Gaussian ve Kuantum) olan iç karşılaştırmalarında ($d(H-G)$ ve $d(Q-H)$), etki büyüklüklerinin neredeyse tamamı ± 0.20 bandında kalmıştır (bkz. Çizelge 4.2 ve Şekil 4.2).

Bu sinerji yokluğu bir rastlantı değildir. Önerme 1 (Bölüm 3.2) bu sonucun *matematiksel bir zorunluluk* olduğunu göstermektedir: $G \times Q \approx G$ eşitliği, Hibrit ve Gaussian modülasyonlarını fonksiyonel olarak birbirine yakın kılmaktadır.

Bu durum, GPT-2 Medium ölçeğinde anlamlılık gösteren tek metrik olan *abstractness_score*'da açıkça görülmektedir: Gaussian modülasyonu Base modele göre güçlü bir etki ($d(G-B) = 0.595$) yaratırken, Hibrit modülasyonda bu etki $d(H-B) = 0.058$ 'e inerek sönümlenmiştir ($d(H-G) = -0.663$). Çarpımsal birleşim sinerji yaratmamış, aksine bileşen etkisini yok etmiştir.

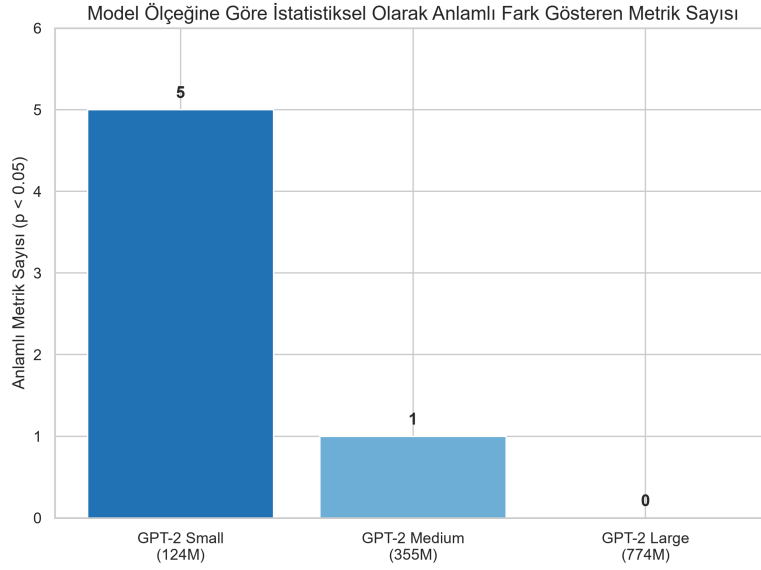
4.3 Bulgu C: Modülasyon Etkisinin Ölçek ile Sönümlenmesi

Deney sonuçları, BAW modülatörünün etkisinin dil modelinin parametre sayısı büyüdükçe sistematik olarak kaybolduğunu göstermektedir. Çizelge 4.1'te görüleceği üzere, ham $\alpha = 0.05$ düzeyinde anlamlı fark veren temel metrik sayısı GPT-2 Small'da 5 iken, GPT-2 Medium'da 1'e düşmüştür. Ancak bu sönümlenme metrik ailesine özgüdür: sözcüksel ve karmaşıklık etkileri (LZ, MTLT) ölçekle kaybolurken, yerel-hapsetme etkisi (konu sapmasındaki düşüş) Large ölçeğinde de kalıcıdır (bkz. Bölüm 4.6).

Katı çoklu karşılaştırma düzeltmesi uygulandığında: Bonferroni eşiği ($\alpha_{adj} = 0.05/13 = 0.00385$) altında, GPT-2 Small'da bile hiçbir metrik anlamlı kalmamaktadır (*dependency_distance* $p = 0.0041 > 0.00385$). Benjamini-Hochberg FDR düzeltmesi de aynı sonucu vermektedir. **Sonuç olarak, temel 13 metrik için hem Bonferroni hem FDR altında hiçbir ölçüm sağlam biçimde anlamlı değildir.**

4.4 Bulgu D: Sözcüksel Çeşitlilik ve Sözdizimsel Karmaşıklık Arasındaki Takas

Özellikle GPT-2 Small ölçeğinde (Çizelge 4.2), modülasyonun modele basitçe “zarar verdiği” veya “iyileştirdiği” şeklinde tek boyutlu bir çıkarım yapmak yerine, metin üretiminde yapısal bir kayma yaşandığı gözlemlenmiştir. Hibrit yöntem, bir yandan *MTLT* ($d(H-B) = -0.668$) ve *distinct-1* ($d(H-B) = -0.653$) gibi sözcüksel metriklerde düşüşler yaşatırken; eş zamanlı olarak *pos_entropy* ($d(H-B) = +0.303$) ve *dependency_distance* ($d(H-B) = +0.441$) gibi sözdizimsel karmaşıklık metriklerinde artışa neden olmuştur.



Şekil 4.1 Model ölçeği büyüdükçe (Small → Medium → Large), BAW mekanizmalarının Base modele göre istatistiksel olarak anlamlı ($p < 0.05$) fark yarattığı metrik sayısındaki sönümlenme.

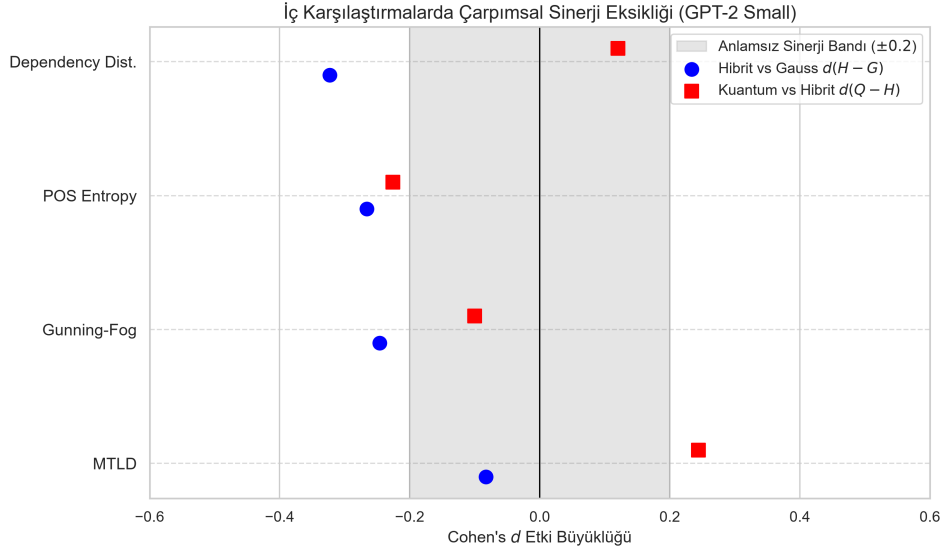
4.5 Bileşen İzolasyonu: Quantum-only vs Gaussian-only vs Hybrid

Dört kollu deneysel tasarımın temel amacı, bileşenlerin ayrıştırılarak izole etkilerinin gözlemlenmesidir. Çizelge 4.2’deki Cohen’s d değerleri incelendiğinde kritik bir örüntü ortaya çıkmaktadır:

Quantum-only kolu boş değildir: Kuantum bileşeni, Quantum-only kolunda Gaussian çarpanı olmadan tüm dizi boyunca aktif kalmaktadır. Bu kolda Base modele kıyasla ölçülebilir yapısal kaymalar gözlemlenmiştir: $d(Q-B) = +0.547$ (bağımlılık mesafesi), $d(Q-B) = -0.591$ (MTLD), $d(Q-B) = -0.520$ (Gunning-Fog). Bu kaymalar Gaussian ve Hibrit kollarıyla aynı yöndedir ve orta büyüklüktedir.

Üç kolun yakınsak davranışı: Gaussian-only, Quantum-only ve Hibrit kollarının üçü de Base’e göre benzer yönde ve büyüklükte kaymalar sergilemektedir. Bu yakınsama, Bulgu A’yı güçlendirmektedir: sorun yalnızca çarpımsal formülasyondaki iptal değil, daha genel olarak “statik çarpımsal dikkat maskeleri yaratıcılık proxy metriklerinde iyileşme sağlamamaktadır.”

Quantum-only'nin sınırları: Her ne kadar Quantum-only kolu yapısal kaymalar üretse de, bu kaymalar yaratıcılık proxy metriklerinde (metaphor_density, cross_domain_rate, abstractness_score) sistematik bir üstünlüğe dönüşmemiştir. Bu, çarpımsal maskenin doğasından kaynaklanan bir sınırlılıktır: $W \in [0, 1]$ maskesi bir logit'i orijinal değerinin üzerine çıkaramaz, yalnızca sıfıra doğru çekebilir (bkz. Bölüm 3.2.3).



Şekil 4.2 GPT-2 Small modeli için anahtar metriklerdeki iç karşılaştırma etki büyüklükleri. d değerlerinin büyük çoğunluğunun ± 0.2 anlamsız sinerji bandında kalması, çarpımsal birleşimin ek bir etki yaratmadığını göstermektedir.

4.6 İkincil Doğrulayıcı Metrikler (LZ-Karmaşıklık, Konu Sapması, Anlamsal Yayılm)

Temel 13 metriğin ötesinde, modelin yapısal davranışını ve mekanistik sınırlarını daha derinlemesine doğrulamak amacıyla üç ikincil metrik (Lempel-Ziv Karmaşıklık, Konu Sapması ve Metin-İçi Anlamsal Yayılm) analize dahil edilmiştir. Bu metrikler, ana 13 metrikten sonra hipotezi doğrulamak amacıyla koşulan konfirmatuar bir "İkincil Doğrulayıcı Analiz" ailesi olarak değerlendirilmiştir. Düzeltme eşiği bu üç metrik için $\alpha \approx 0.05/3 \approx 0.0167$ olarak uygulanmıştır. Sonuçlar Çizelge 4.5'da sunulmuştur.

1. Lempel-Ziv Karmaşıklık (Tezin En Sağlam Bulgusu): Lempel-Ziv (LZ), kelime sayımına dayanan MTLD ve Distinct-1'den farklı bir felsefeyle çalışarak metnin algoritmik

sıkıştırılabilirliğini ölçer. GPT-2 Medium ölçeğinde LZ karmaşıklığı anlamlı derecede düşmüştür ($p = 0.0019$). Temel 13 metrik ailesi içinde bu bulgu, tüm çoklu karşılaştırma düzeltmelerinden (Bonferroni ve FDR) sağ çıkan tek ölçümdür; ikincil 3-metrik ailesinde ise Konu Sapması (Topic Drift, Medium $p = 0.0080 < \alpha_{adj} = 0.0167$) tarafından tamamlanmaktadır. LZ'nin gösterdiği karmaşıklık düşüşü (artan tekrarlılık), yüzeysel metriklerin gösterdiği çeşitlilik düşüşüyle çok-katmanlı bir doğrulama yaratarak “metriklerin tek başına kapris yaptığı” itirazını zayıflatmaktadır.

2. Konu Sapması (Topic Drift) ve Yönsel İhlal: Medium ve Large ölçeklerinde konu sapması belirgin biçimde azalmıştır (negatif d değerleri). Başka bir deyişle, modeller prompt'a daha sadık kalmış ve uzak bağlamlara sapmamıştır. BAW'ın ve özellikle kuantum bileşeninin ana tasarım amacı uzak menzillerde keşifsel sıçramalar (divergence) yaptırmaktır. Gözlemlenen bu düşüş, yöntemin konuya sadakati artırdığı şeklinde olumlu bir "başarı" olarak değil; mekanizmanın keşfi teşvik etmek yerine dikkati yerel bir alana daralttığı şeklinde okunmalıdır. Bu durum, çarpımsal maskenin uzak menzilde sıfırlandığı ($W \approx 0$) hipotezini ampirik olarak doğrular.

3. Metin-İçi Anlamsal Yayılm (Sessiz Null): Cümleler arası anlamsal mesafeyi ölçen DSI metriğinde hiçbir ölçekte anlamlı bir fark gözlemlenmemiştir ($p > 0.05$). Bu, modelin içsel anlamsal sıçramalar veya kavramsal farklılıklar kazanmadığını gösteren çok net bir null bulgusudur ve BAW'ın "daha fazla yaratıcı keşif" sağladığı iddiasının aleyhine güçlü bir argümandır.

Ayrıca, her üç ikincil metrikte de Hibrit (H) ve Gaussian (G) kollarının etki büyüklüklerinin çok benzer olması (örn. Medium LZ: -0.717 ve -0.732), çarpımsal birleşimin sinerji üretmediğini 16 metriğin tamamında tutarlı bir şekilde göstermektedir.

4.7 Niteliksel Metin İncelemesi ve Otomatik Metriklerin Sınırları

Otomatik ölçüm metriklerinin anlamsal yaratıcılığı yakalamadaki potansiyel sığlığını göstermek adına model çıktılarından niteliksel bir örneklem alınmıştır. Aşağıdaki örnek,

GPT-2 Small modelinde ve 4 farklı mekanizmada (Base, Gaussian, Quantum, Hybrid) deterministik bir karşılaştırma yapabilmek için *seed=1* konfigürasyonunda üretilen ilk metinlerdir (cherry-pick yapılmamıştır).

Örnek 1 Prompt: *“The old lighthouse keeper discovered something unusual in the beam’s reflection that night”*

1. Base Model Çıktısı:

“. It wasn’t even the lighthouse, it was the crystal. The lighthouse was broken into fragments and scattered by the wind. It had a crystal with a hole in it that would never heal. It’s time to break away from the past and get back to life.”

2. Gaussian Model Çıktısı:

“. The lighthouse keeper called for his staff to contact his son, who would return the lighthouse to the man himself. Once it was clear that the lighthouse would not function properly when he returned the beacon...”

3. Quantum Model Çıktısı:

“, and it immediately stopped the beam from moving. ‘Hey, that’s where that old light came from, too. You’re in there, man. Look over there, man. That’s where that light came from.’”

4. Hibrit Model Çıktısı:

“. It was a strange, eerie thing. A lighthouse in the form of a building had

vanished when someone called and told them it. A lighthouse of something, a new, strange sight.”

Örnek 2 Prompt: *“Sarah found her grandmother’s diary hidden behind the loose floorboard, and the first entry changed everything”*

1. Base Model Çıktısı:

“about her life. And then there was the diary. This was the moment after the murder. In the first place, she had only been working at a restaurant until the end of the summer...”

2. Gaussian Model Çıktısı:

“. The second entry was a call for assistance from her father, who’d given her his mother as a kid. In her early teens, she used to see the family on the street...”

3. Quantum Model Çıktısı:

“. ’And this time we are telling you our story as we have always told, we are a team that does this, and we are looking forward to doing all we can to make sure we find a way out...’ ”

4. Hibrit Model Çıktısı:

“. ’If you ever saw this diary I would say ’oh, yeah, it’s interesting’ but in an open field, it didn’t even look like a map, they were completely naked and there was nothing there...’ ”

Bu niteliksel örnekler, yalnızca Bulgu D’nin — otomatik metriklerin yapısal kaymaları algılamadaki yetersizliğinin — illüstrasyonu olarak sunulmaktadır. Tek bir seed’e (seed=1) dayanan bu örnekler, “kuantum bileşeni yaratıcı sıçramalar üretiyor” şeklinde bir birincil bulgu olarak yorumlanmamalıdır. Birincil kanıt, Çizelge 4.2’daki Cohen’s *d* değerleridir.

Nicel null sonucu “BAW metni hiç değiştirmiyor” biçiminde yorumlanmamalıdır; daha doğru çerçeve “mevcut sayaç-tabanlı otomatik metrikler, BAW’ın yarattığı yapısal kaymaları algılayacak kapasitede değildir” şeklindedir. Gelecek çalışmalarda, DSI (Divergent Semantic Integration) gibi dağılımsal anlamsal modellemeye dayalı doğrulanmış metriklerin kullanılması önerilmektedir.

4.8 Körleştirilmiş LLM-as-Judge İkincil Değerlendirmesi

Otomatik metriklerin "kısıtlı" ve "yüzeysel" doğasına yönelik itirazları test etmek ve "otomatik metrikler BAW’ın yarattığı anlamsal sıçramaları göremiyor olabilir" hipotezini netleştirmek adına, üçüncü bir doğrulama katmanı olarak LLM-as-judge analizi uygulanmıştır. GPT-2 Small ölçeğinde 200 adet prompt-seed grubu (toplam 800 metin) güçlü bir hakem modeline anonim olarak ve sırası karıştırılarak sunulmuştur (Bkz. Bölüm 3.3.6).

1. Kollar Arası İstatistiksel Fark Yokluğu (Çok-Katmanlı Doğrulama)

Değerlendirilen 4 kriterin hiçbirinde kollar (Base, Gaussian, Quantum, Hybrid) arasında istatistiksel olarak anlamlı bir fark bulunamamıştır (Çizelge 4.6). En iyi ihtimalde dahi Cohen’s *d* etki büyüklükleri ± 0.13 bandının içinde kalarak sıfır etkisine (null-effect) eşdeğer bulunmuştur.

Bu bulgu tezin temel null-hipotezini kritik biçimde doğrulamaktadır. LLM hakeminin de kollar arasında fark bulamaması, tezin "otomatik metrikler kördür" iddiasını *çürütmez*, ancak şu durumu netleştirir: kollar arasında algılanabilir bir yaratıcılık veya kalite farkı *gerçekten yoktur* ya da yok denecek kadar zayıftır. Yüzey-yapısal sayaç metrikleri

(MTLD, Distinct-1 vb.) ile üst-düzey anlamsal LLM yargısının aynı "etki yok" sonucuna varması, tezin sönümlü modülasyon bulgusunu güçlendiren çok-katmanlı bir doğrulama oluşturmuştur.

2. Sıralama (Ranking) ve Sinerji Yokluğu

Hakemden dört metni yaratıcılık açısından eşzamanlı sıralaması istendiğinde (1=en yaratıcı), ortalama sıralar şu şekilde gerçekleşmiştir: Base 2.500, Gaussian 2.535, Quantum 2.455, Hybrid 2.510. Friedman testi bu sıralamalar arasında anlamlı bir fark bulamamıştır ($\chi^2 = 0.402$, $p = 0.9398$). Metinlerin "en yaratıcı (1. sıra)" seçilme oranları şans düzeyinde (%25) dağılmıştır ($\chi^2 = 0.520$, $p = 0.9145$).

Ek olarak, Hibrit kolun iç bileşenleriyle yapılan karşılaştırmalarda (sinerji testi) tüm etki büyüklükleri $d(H - G)$ ve $d(Q - H)$ için ± 0.11 bandında kalmıştır. Quantum kolunun yaratıcılıkta en yüksek ortalamaya (2.895) sahip olması bir "üstünlük" değil, istatistiksel gürültü ($d = 0.052$) sınırları içindedir. Bu durum, çarpımsal maskenin kuantum bileşenini uzak menzilde sıfırlaması ($W \approx 0$, bkz. Çizelge 3.2) hipotezini ampirik olarak yeniden onaylar: hakemin gözünde bile kuantum bileşeni ayırt edilebilir bir sinerji üretememiştir.

3. Farklılaşma ile Daha Yaratıcı Olmanın Ayrımı

Niteliksel incelemelerde (Bölüm 4.7), BAW modülasyonlu kolların metinleri belirgin biçimde *farklılaştırdığı* görülmüş; ikincil yapısal metrikler de bu kaymayı doğrulamıştı. Ancak LLM hakem skorları göstermiştir ki; bu sapmalar, metni "daha yaratıcı" veya "daha kaliteli" yapmamıştır. "Farklılaşmak, daha yaratıcı olmak anlamına gelmez" tespiti, bu çalışmanın en somut çıktılarından biridir. Modülasyon, modeli raydan çıkararak yapısal çöküşe yol açmış, ancak hedeflenen "başarılı" keşifsel sıçramayı sağlayamamıştır.

4. Pozisyon Yanlılığı ve Metrik Kırılganlığı (Kritik Bulgu)

LLM-as-judge analizinin beklenmedik ama en çarpıcı bulgusu, hakem modelindeki güçlü

pozisyon yanlılığıdır (Çizelge 4.7). Dört metin içeriklerinden bağımsız olarak sadece sunum sırasına göre puanlandığında, 4. sıradaki metin, 1. sıradakinden ortalama ~ 1.5 puan daha yüksek skorlanmıştır ($p < 0.0001$).

Körleştirme aşamasında yöntemler rastgele karıştırıldığı için bu yanlılık yöntem ortalamalarını etkilememiş ve kendi içinde dengelenmiştir. Ancak bu bulgu, "insan-benzeri" sanılan bir değerlendiricinin bile içeriği değil *sunum sırasını* ödüllendirdiğini güçlü biçimde göstermektedir. Bu durum, tezin "yaratıcılık ölçümü doğası gereği kırılıgandır" tezine yeni ve bağımsız bir ampirik argüman sunmaktadır.

Sınırlılıklar: Bu ikincil değerlendirme, doğası gereği tek bir hakem modeline dayanmaktadır ve insan değerlendirmesinin mutlak yerine geçemez. Sadece GPT-2 Small ölçeği ve tek bir seed (seed=1) üzerinden yürütülen bu çalışma, tek-hakem bir sınırlılık barındırmaktadır. Tespit edilen güçlü pozisyon yanlılığı, LLM-as-judge mutlak puanlarının gürültülü olduğunu gösterir; sonuçlar bu ışıktaki temkinli yorumlanmalıdır. Çok-hakemli ve insan değerlendirmeli genişletme gelecek çalışmaya önerilir.

Çizelge 4.1 Üç farklı model ölçeği için 13 metrik üzerinden mekanizmalar arası farklılıkların ANOVA test sonuçları ($n = 45$ grup, $\alpha = 0.05$). [*] ham anlamlılık; Bonferroni eşiği $\alpha_{adj} = 0.00385$ altında hiçbir metrik anlamlı kalmamaktadır.

Metrik	GPT-2 Small (p, η^2)	GPT-2 Medium (p, η^2)	GPT-2 Large (p, η^2)
Distinct-1	0.0065, 0.0672 [*]	0.0746, 0.0384	0.5444, 0.0120
Distinct-2	0.2137, 0.0251	0.5771, 0.0111	0.6292, 0.0098
Self-BLEU-3	0.5103, 0.0130	0.9916, 0.0006	0.7701, 0.0064
MTLD	0.0070, 0.0664 [*]	0.4615, 0.0145	0.0903, 0.0361
Gunning-Fog	0.0468, 0.0441 [*]	0.0689, 0.0394	0.0996, 0.0348
Metaphor Density	0.0557, 0.0420	0.1558, 0.0292	0.9428, 0.0022
Cross-Domain Rate	0.9021, 0.0033	0.3061, 0.0203	0.0518, 0.0429
Abstractness Score	0.4678, 0.0143	0.0155, 0.0572 [*]	0.3097, 0.0201
POS Entropy	0.0341, 0.0479 [*]	0.2052, 0.0256	0.9687, 0.0014
Syntactic Complexity	0.7576, 0.0067	0.7317, 0.0073	0.5663, 0.0114
Dependency Distance	0.0041, 0.0723 [*]	0.5370, 0.0122	0.1955, 0.0263
Emotion Arc Variance	0.1397, 0.0306	0.2320, 0.0240	0.1557, 0.0292
Coherence Score	0.2135, 0.0251	0.4646, 0.0144	0.8642, 0.0042

Çizelge 4.2 GPT-2 Small modeli için anahtar metriklerdeki Cohen’s d etki büyüklükleri. Tüm karşılaştırma çiftleri.

Metrik	G vs Base	H vs Base	Q vs Base	H vs G	Q vs H
MTLD	−0.500	−0.668	−0.591	−0.158	+0.108
Gunning-Fog	−0.212	−0.481	−0.520	−0.246	−0.100
POS Entropy	+0.578	+0.303	+0.093	−0.266	−0.226
Dependency Distance	+0.696	+0.441	+0.547	−0.323	+0.120

Çizelge 4.3 GPT-2 Medium modeli için tüm metriklerdeki Cohen’s d etki büyüklükleri (vs Base).

Metrik	G vs Base	H vs Base	Q vs Base
Distinct-1	−0.515	−0.412	−0.227
Distinct-2	+0.237	+0.172	+0.007
Self-BLEU-3	+0.003	+0.001	−0.059
MTLD	+0.376	+0.122	+0.068
Gunning-Fog	−0.416	−0.532	−0.385
Metaphor Density	−0.301	−0.251	+0.116
Cross-Domain Rate	−0.205	+0.102	−0.209
Abstractness Score	+0.595	+0.058	+0.195
POS Entropy	+0.407	+0.268	+0.086
Syntactic Complexity	−0.215	−0.193	−0.101
Dependency Distance	+0.200	−0.011	−0.088
Emotion Arc Variance	+0.356	+0.021	+0.212
Coherence Score	−0.135	−0.244	−0.307

Çizelge 4.4 GPT-2 Large modeli için tüm metriklerdeki Cohen’s d etki büyüklükleri (vs Base).

Metrik	G vs Base	H vs Base	Q vs Base
Distinct-1	+0.073	−0.155	+0.127
Distinct-2	+0.277	+0.105	+0.181
Self-BLEU-3	+0.008	+0.077	−0.136
MTLD	+0.242	−0.143	+0.350
Gunning-Fog	−0.293	−0.013	−0.456
Metaphor Density	+0.089	−0.030	+0.020
Cross-Domain Rate	−0.036	+0.439	+0.320
Abstractness Score	+0.087	+0.049	+0.355
POS Entropy	+0.033	+0.071	+0.098
Syntactic Complexity	−0.254	−0.034	−0.177
Dependency Distance	−0.407	−0.291	−0.304
Emotion Arc Variance	−0.147	+0.216	+0.279
Coherence Score	+0.135	−0.037	+0.035

Çizelge 4.5 İkincil doğrulayıcı metrikler için ANOVA ve Cohen's *d* sonuçları. [*] işareti, doğrulayıcı analiz ailesi Bonferroni eşliğini ($p < 0.0167$) geçen bulguları temsil eder.

Metrik	Ölçek (p)	G vs Base	H vs Base	Q vs Base
Lempel-Ziv (LZ) Karışıklığı	GPT-2 Small (0.0332)	-0.373	-0.582	-0.306
	GPT-2 Medium (0.0019) [*]	-0.732	-0.717	-0.491
	GPT-2 Large (0.8057)	+0.207	+0.090	+0.035
Konu Sapması (Topic Drift)	GPT-2 Small (0.4099)	-0.236	-0.205	+0.063
	GPT-2 Medium (0.0080) [*]	-0.674	-0.672	-0.493
	GPT-2 Large (0.0348)	-0.581	-0.188	-0.332
Metin-İçerik Anlamsal Yayılm (DSI)	GPT-2 Small (0.0728)	-0.447	-0.525	-0.280
	GPT-2 Medium (0.1307)	+0.347	+0.455	+0.306
	GPT-2 Large (0.9869)	-0.032	+0.014	+0.047

Çizelge 4.6 LLM-as-Judge değerlendirme skorları (1-5 arası) ve kollar arası tek-yönlü ANOVA sonuçları. Tüm p değerleri anlamsızdır ($p > 0.05$).

Kriter	Base	Gaussian	Quantum	Hybrid	ANOVA F	p -değeri
Yaratıcılık	2.840(1.10)	2.795(1.03)	2.895(1.02)	2.785(1.08)	0.450	0.7172
Tutarlılık	2.390(0.86)	2.305(0.83)	2.420(0.80)	2.340(0.89)	0.733	0.5324
Prompt Bağlılığı	2.265(0.97)	2.290(0.98)	2.385(0.90)	2.290(0.99)	0.609	0.6096
Genel Kalite	2.395(0.88)	2.320(0.86)	2.415(0.83)	2.365(0.95)	0.438	0.7259

Çizelge 4.7 LLM hakeminin anonim sunum sırasına (pozisyon) göre verdiği ortalama skorlar ve ANOVA sonuçları. [*] yüksek anlamlılığı ($p < 0.0001$) temsil eder.

Pozisyon	Yaratıcılık [*]	Tutarlılık	Prompt Bağlılığı	Genel Kalite [*]
1. Metin	2.175	1.770	1.870	1.760
2. Metin	2.465	2.150	1.965	2.130
3. Metin	2.980	2.525	2.485	2.535
4. Metin	3.695	3.010	2.910	3.070
ANOVA F/p	112.83/< 0.0001	104.51/< 0.0001	70.82/< 0.0001	116.35/< 0.0001

5. TARTIŞMA ve SONUÇ

5.1 Tartışma ve Sınırlılıklar

Bu bölümde, deneysel süreçten elde edilen bulguların kuramsal bağlamı değerlendirilmiş; Bölüm 3.2’te ispatlanan menzil-uyumsuzluğu önermesinin ampirik sonuçlarla ilişkisi, yaratıcılık ölçümünde kullanılan otomatik metriklerin epistemolojik sınırları, bileşen izolasyonunun çıkarımları, ölçek sönümlenmesi ve tasarım sınırlılıkları tartışılmıştır.

5.1.1 Menzil-Uyumsuzluğunun Ampirik Teyidi ve Bulgu Çerçevesi

Bölüm 3.2’te ispatlanan Önerme 1, çarpımsal hibrit maskede ($W = G \times Q$) kuantum bileşeninin uzak-menzil rezonansının Gaussian sönümleme tarafından 10^{-54} mertebesinde sıfırlandığını ve bu iptali tüm deneysel uzay boyunca kategorik olarak göstermektedir. Deneysel bulgular (Bölüm 4) bu analitik öngörüyü dört açıdan teyit etmektedir:

1. **Bulgu A: Hiçbir kol Base’i geçmiyor.** Gaussian, Kuantum ve Hibrit kollarının hiçbirisi, yaratıcılık vekil metriklerinde (metaphor_density, cross_domain_rate, abstractness_score) Base modeli istatistiksel olarak anlamlı ve büyük bir etkiyle aşamamıştır. Bu geniş kapsamlı negatif sonuç, Önerme 1’in doğrudan açıklaması değildir; çünkü iptal yalnızca Hibrit kola özgüdür. Bulgu A’nın doğru yorumu şudur: Eğitimsiz statik çarpımsal dikkat maskeleri, iptal bulunmasa bile, yaratıcılık vekil metriklerinde sistematik bir iyileşme sağlamamıştır.
2. **Bulgu B: Hibrit modülasyonunda sinerji yokluğu.** Hibrit yöntemin kendi alt bileşenleriyle olan karşılaştırmalarında ($d(H-G)$ ve $d(Q-H)$), etki büyüklüklerinin neredeyse tamamı ± 0.20 bandında kalmıştır. Önerme 1’in açıkladığı tam olarak budur: $G \times Q \approx G$ eşitliği nedeniyle Hibrit ve Gaussian modülasyonları fonksiyonel olarak birbirine yakındır. Bu sinerji yokluğu, bir rastlantı değil matematiksel bir zorunluluktur.
3. **Bulgu C: Ölçekle sönümlenme.** Temel metrikler açısından anlamlı farklar Large ölçeğinde kaybolsa da, İkincil Doğrulayıcı Analizler (Bölüm 4.6) bu sönümlenmenin

metrik ailesine özgü olduğunu göstermiştir: sözcüksel çeşitlilik ve karmaşıklık etkileri ölçekle sönümlenirken, konuya sadakat (topic drift azalması) Large ölçeğinde de (ham düzeyde, ikincil düzeltme eşliğini geçmeden) kalıcı kalmıştır. Büyük modellerin derin iç dikkat mekanizması, modülasyonun yaratıcı sıçramalar üretmesini absorbe etse de, yerel-hapsetme etkisini silememiştir.

4. **Bulgu D — Yüzey-Yapısal vs Anlamsal Proxy Metrik Ayrımı:** Niteliksel inceleme ve ikincil analizler (özellikle Lempel-Ziv karmaşıklığı ve DSI metrikleri), otomatik değerlendirme metrikleri arasındaki güvenilirlik farkını ortaya koymuştur. Yüzey-yapısal metrikler (MTLD, LZ) yakınsak ve sağlam (robust) sonuçlar verirken, üst-düzey anlamsal yaratıcılık proxy’leri (metaphor_density) kırılğan kalmıştır. Bu ayrım, teoriden öngörülemeyen bağımsız bir ampirik katkıdır.

Bu dört-parçalı çerçeve, önceki tek-nedenli açıklamaya (“iptal her şeyi açıklıyor”) kıyasla daha doğru ve savunulabilir: iptal spesifik olarak Bulgu B’nin yapısal nedenidir; Bulgu A ise daha geniş bir ampirik gerçeği yansıtmaktadır.

5.1.2 Otomatik Yaratıcılık Metriklerinin Yetersizliği ve Kırılğanlığı

Çalışma kapsamında değerlendirilen *cross_domain_rate* ve *metaphor_density* gibi metrikler, büyük dil modellerinin “yaratıcılığını” nesnel olarak ölçmek amacıyla alanda sıklıkla kullanılmaktadır. Ancak, uygulanan analiz hattının (pipeline) yapısal incelenmesi sonucunda, bu metriklerin derin bir anlamsal çıkarım yapmaktan ziyade, keyfi olarak belirlenmiş kelime kümeleri (heuristics) ve kalıpların sayımı üzerinden çalışan yüzeysel *proxy* (vekil) ölçümler olduğu tespit edilmiştir.

Çalışma kapsamında Distinct-1 ve MTLD gibi metrikler sözcüksel çeşitliliği ölçmek için bir arada kullanılmıştır. Her ne kadar bu iki metrik GPT-2 Small ölçeğinde kısmen benzer bir dağılım (sırasıyla $p = 0.0065$ ve $p = 0.0070$) sergilese de, gerçek Pearson korelasyonları $\sim 0.39 - 0.50$ aralığındadır ve Medium ölçeğinde belirgin biçimde ayrışmaktadırlar (Distinct-1 $p = 0.0746$ vs MTLD $p = 0.4615$). Bu durum, söz konusu metriklerin

farklı yapısal özellikleri ölçtüğünü ve ev-yapımı proxy metriklerin (*metaphor_density*, *cross_domain_rate*) dalgalı performansının sayaç-tabanlı yaklaşımların yapısal kırılmalı-ğından kaynaklandığını göstermektedir.

Ancak ikincil doğrulayıcı analiz olarak koşulan Lempel-Ziv (LZ) karmaşıklığı bu tabloyu netleştirmiştir. LZ'nin Medium ölçeğinde anlamlı bir çöküş (daha yüksek sıkıştırılabilirlik) göstermesi ($p = 0.0019$), kelime sayımından tamamen farklı bir algoritmik felsefeyle (veri sıkıştırma) yüzey-yapısal değişimi doğrulamaktadır. Bu durum tezin metrik kırılmalı-ğının argümanını şu şekilde inceltmektedir: yüzey-yapısal metrikler (MTLD, Distinct-1, LZ) daha sağlam (robust) sonuçlar verebilirken, üst-düzey anlamsal yaratıcılık proxy'leri kırılmalı ve yüzeyseldir. LZ'nin temel 13 metrik düzeltmelerini de aşan en sağlam ölçüm olması (ikincil ölçümlerde Topic Drift ile birlikte), bu yapısal kaymanın metodolojik bir illüzyon olmadığını desteklemektedir.

Bölüm 4'deki niteliksel inceleme, BAW modülasyonlu kolların metni belirgin biçimde farklılaştırdığını; ancak bu yapısal kaymaların otomatik metrikler tarafından "yaratıcılık" olarak yakalanmadığını göstermiştir. Üçüncü doğrulama katmanı olarak koşulan körleştirilmiş LLM-as-judge analizi (Bölüm 4.8), bu tartışmayı nihai bir senteze taşımıştır. Niteliksel olarak saptanan bu biçimsel "farklılaşmalar", gelişmiş bir GPT-4o hakeminin gözünde bile "daha yüksek yaratıcılık" veya "daha yüksek kalite" olarak skorlanmamıştır. Bu durum, çalışmanın en kritik epistemolojik sentezlerinden birini doğurur: **Model çıktılarının farklılaşması, daha yaratıcı olmaları anlamına gelmez.** Modülasyon, model raydan çıkararak yapısal kaymalara yol açmış, ancak hedeflenen keşifsel sıçramayı başaramamıştır.

Dahası, LLM hakeminin değerlendirmelerinde tespit edilen güçlü pozisyon yanlılığı (4. sıradaki metnin içerikten bağımsız olarak 1. sıradakinden ~ 1.5 puan daha yaratıcı bulunması, $p < 0.0001$), "insan-benzeri" sanılan otonom değerlendiricilerin dahi ne kadar gürültülü ve kırılmalı olabildiğini kanıtlamıştır. Bu bulgu, tezin "yaratıcılık ölçümünün doğası gereği kırılmalıdır" argümanını yalnızca sayaç-tabanlı metrikler özelinden çıkararak, modern yapay zeka hakemlerini de kapsayacak şekilde genişletmektedir. Gelecek

alıřmalarda, dađılımsal anlamsal modellemeye dayalı dođrulanmıř ltlerin (rn. DSI-esinli metrikler) ve ok-hakemli uzlařma mekanizmalarının entegre edilmesi zorunluluk arz etmektedir.

5.1.3 Karřılařtırma Temeli Eksikliđi

Bu alıřmanın nemli bir sınırlılıđı, BAW mekanizmasının post-hoc rnekleme yntemleriyle (sıcaklık leklendirme, top-k, top-p) dođrudan deneysel olarak kıyaslanmamıř olmasıdır. Tez, bu yntemleri “sonradan mdahale” nitelikleri nedeniyle eleřtirmekte; ancak “BAW, sıcaklıđı ykseltmekten daha mı iyi?” sorusuna ampirik bir yanıt vermemektedir. Eřit tutarlılıkta sıcaklık-eřli bir baseline ile kk bir kafa-kafaya karřılařtırma, bu bořluđu kapatacak ve sonuların grelili konumlanmasını netleřtirecektir. Bu kıyaslama gelecek alıřmalar iin somut bir neri olarak kaydedilmektedir.

5.1.4 Modlasyon Sinerjisinin ve lek Snmlenmesinin Deđerlendirilmesi

Kuantum modlatrn ve Gaussian modlatrn arpımsal birleřimi olan Hibrit yntemin, tekil bileřenlerin toplamından daha byk bir sinerji reteceđi varsayılmıřtır. nerme 1 (Blm 3.2), bu varsayımın neden gerekleřemeyeceđini matematiksel olarak gstermektedir: $G \times Q \approx G$ eřitliđi, Hibrit ve Gaussian modlasyonlarını fonksiyonel olarak eřdeđer kılmaktadır. Ampirik bulgular bu yapısal zorunluluđu teyit etmiřtir.

lek snmlenmesi konusunda bir sınırlılık belirtilmelidir: bu alıřmada yalnızca GPT-2 ailesinin  varyantı (124M, 355M, 774M) kullanılmıř olup, bu ~6 katlık bir aralıktaki sadece  veri noktası iermektedir. Bu nedenle “byk dil modelleri leklendike” řeklinde bir genelleme yapılması istatistiksel olarak zayıftır. Burada raporlanan snmlenme, GPT-2 ailesi iinde gzlemlenen bir rnt olarak deđerlendirilmelidir.

5.1.5 Statik Katman Konfigürasyonu (Depth=1) Sınırlılığı

Çalışmada sunulan 108.000 metinlik deney havuzunun tamamı, BAW modülatörünün modelin tüm katmanlarına statik bir derinlik ($depth = 1$) parametresi ile sabitlenmesiyle elde edilmiştir (bkz. Bölüm 3.1, Bölüm 4.4). Modülasyon sinyalinin modelin hangi katmanlarında daha aktif olması gerektiğine dair bir arama yapılmamıştır. Sonuçlarda görülen düşük etki büyüklüklerinin önemli sebeplerinden biri, statik modülasyon katmanının derin mimari içindeki sinyal iletimi sırasında (rezidüel bağlantılar üzerinden) gücünü yitirmesi olabilir. QASA (Chen ve Kuo 2025), kuantum-esinli bileşenlerin Transformer-benzeri yapılarda katman konumlandırması açısından araştırılabileceğine dair analogik ve destekleyici bir örnek sunmaktadır. Bununla birlikte QASA, BAW için doğrudan deneysel öncül veya NLP metin üretimi bağlamında doğrudan kanıt olarak değerlendirilmemelidir. Bu nedenle BAW'ın gelecekte yalnızca tüm katmanlara statik biçimde uygulanması yerine, katman-seçici konfigürasyonlarla da sınanması uygun olacaktır.

5.1.6 İstatistiksel Yöntem Tartışması

Aynı seed'den dört mekanizma üretildiğinden veriler teknik olarak eşleştirilmiş (within-seed) bir yapı taşımaktadır. Bu çalışmada bağımsız gruplar ANOVA'sı tercih edilmiş olup, bu seçim muhafazakâr bir yaklaşım oluşturmakta ve raporlanan null bulgularını güçlendirmektedir: seed-içi korelasyon pozitifse, bağımsız ANOVA hata terimini şişirir ve etki bulması zorlaşır. Ayrıca, grupta $n = 45$ gözlem ile Bonferroni sonrası $\alpha_{adj} = 0.00385$ eşliğinde, $d \approx 0.3 - 0.5$ bandındaki küçük etkileri saptama gücü (statistical power) doğası gereği düşüktür. Dolayısıyla raporlanan geniş çaplı null sonuçları, mutlak anlamda "hiçbir etkinin olmadığı" şeklinde değil, uygulanan sıkı düzeltmeler sonrası "saptanabilir ve sistematik bir etkinin kalmadığı" şeklinde dürüst bir muhafazakârlıkla yorumlanmalıdır. En güçlü alternatif, seed'i rastgele etki (random effect) olarak alan bir doğrusal karma model (LMM) koşturmak ve sınıf-içi korelasyonu (ICC) raporlamaktır. Bu yaklaşım, gelecek çalışmalarda robustluk eki olarak değerlendirilmelidir.

5.2 Sonuç

Bu tez çalışmasında, büyük dil modellerinin (LLM) öz-dikkat (self-attention) mekanizmalarına eğitim veya ince-ayar (fine-tuning) gerektirmeksizin müdahale etmeyi hedefleyen, fiziksel-motivasyonlu bir dikkat modülasyonu yöntemi (BAW – *Bilgehan-Aslan Penceresi*) önerilmiş, yapısal sınırları analitik olarak ispatlanmış ve ampirik olarak doğrulanmıştır. Olasılık genliklerine dayalı kuantum-esinli modülasyon ile yerel odaklanmayı hedefleyen Gauss modülasyonunun çarpımsal birleşimi olan bu yöntem, çıkarım (inference) aşamasında dikkat matrisinin yapısını modifiye ederek modelin yaratıcı metin üretim kapasitesini artırmayı amaçlamıştır. Yöntem, GPT-2 mimarisinin üç farklı ölçeğinde (Small, Medium, Large) dört kolda (Taban, Gaussian, Kuantum, Hibrit) toplam 108.000 metin üzerinden 16 farklı (13 temel, 3 ikincil doğrulayıcı) otomatize yaratıcılık vekil (proxy) ve metin kalitesi metriği ile titizlikle analiz edilmiştir.

Gerçekleştirilen kapsamlı analiz sonucunda şu temel bulgulara ulaşılmıştır:

Bulgu A: Hiçbir kol Base’i geçmiyor. Gaussian, Kuantum ve Hibrit kollarının hiçbiri, yaratıcılık vekil (proxy) metriklerinde Base modeli anlamlı ve büyük bir etkiyle aşamamıştır. Kuantum-only kolu, iptal mekanizmasından bağımsız olarak tüm dizi boyunca aktif kalmasına rağmen, orta büyüklükte yapısal kaymalar üretmiş (örn. $d(Q-B) = +0.547$ bağımlılık mesafesinde) ancak bu kaymalar yaratıcılık proxy metriklerinde sistematik üstünlüğe dönüşmemiştir. Bu, “eğitimsiz statik çarpımsal dikkat maskelerinin, iptal bulunmasa bile, yaratıcılık proxy metriklerinde iyileşme sağlamadığı” genel sonucunu desteklemektedir.

Bulgu B: Hibrit modülasyonunda sinerji yokluğu. Çarpımsal maskenin ($W = G \times Q$) yapısal bir menzile uyumsuzluğu taşıdığı Önerme 1 ile analitik olarak ispatlanmıştır. Gaussian bileşenin ($\sigma \leq 4.5$) hızlı üstel sönümlenmesi, kuantum bileşeninin uzak-menzil rezonans zirvelerini temsili $\sigma = 4, \theta_{scale} = 0.10$ konfigürasyonunda 10^{-54} ; incelenen hiperparametre ızgarasının tamamında (analitik fizibilite analizi) ise 10^{-7} – 10^{-280} aralığında sıfırlayarak Hibrit modülasyonu fonksiyonel olarak Gaussian modülasyonuna eşdeğer kılmaktadır ($G \times Q \approx G$). Deneysel uzaydaki tüm 15 konfigürasyonun fizibilite

eşliğinin $\sim 2-12$ kat altında kaldığı gösterilmiştir. Bu matematiksel iptal, sinerji yokluğunun ($d(H-G) \approx \pm 0.2$) yapısal nedenidir.

Bulgu C: Ölçekle sönümlenme. BAW modülatörünün temel sözcüksel ve algoritmik karmaşıklık etkileri, model parametre sayısı büyüdükçe sistematik olarak sönümlenmiştir. Ancak bu sönümlenme tüm metrik ailesi için geçerli değildir; yerel-hapsetme etkisi (konuya sadakatin artması ve topic drift'in azalması) büyük modelde dahi korunmuştur. Bu durum sönümlenmenin metrik tipine özgü bir doğası olduğunu göstermektedir. Ancak bu çıkarımların yalnızca üç veri noktasına (GPT-2 124M / 355M / 774M) dayandığı ve daha geniş bir ölçek genellemesi için henüz yeterli olmadığı belirtilmelidir.

Bulgu D: Üç katmanlı doğrulama ve metrik kırılabilirliği. Bu çalışma, yaratıcılık vekil (proxy) metriklerinin epistemolojik sınırlarını üç-katmanlı bir doğrulama çerçevesiyle ortaya koymuştur: (1) ana otomatik metrikler, (2) ikincil yapısal metrikler (LZ, Topic Drift) ve (3) körleştirilmiş LLM-as-judge değerlendirmesi. Her üç katman da, kollar arasında anlamlı ve istikrarlı bir kalite veya yaratıcılık farkı bulamayarak tezin null-etkisi argümanını çok-katmanlı bir doğrulamayla sabitlemiştir. LZ karmaşıklığının gösterdiği robustluk, yüzeysel metriklerin (MTLD) yakınsaklığını sağlarken, LLM-as-judge analizinin saptadığı "farklılaşma $\neq$ yaratıcılık" gerçeği bu tablonun son eksik parçasını tamamlamıştır. Dahası, modern bir LLM hakeminde gözlemlenen güçlü pozisyon yanlılığı ($p < 0.0001$), yalnızca sayaç-tabanlı proxy'lerin değil, "insan-benzeri" sanılan otonom değerlendiricilerin de ne kadar kırılabilir olduğuna dair bağımsız ve güçlü bir gösterge sunmuştur.

Özetle bu tez, fiziksel-motivasyonlu ve eğitimsiz bir dikkat modülasyonunun 108.000 metinlik geniş bir üretim havuzunda etkisinin sınırlı, ölçek-bağımlı ve çarpımsal-sinerjiden yoksun çıktığını gösterirken; bu sonuçların arkasındaki mekanistik nedenleri ayırtmıştır. Menzil-uyumsuzluğu (Önerme 1) Hibrit kolundaki sinerji yokluğunun spesifik nedenidir; ancak hiçbir kolun Base'i geçememesi (Bulgu A) daha geniş bir ampirik gerçeği yansıtmaktadır. Gelecek çalışmaların şu somut hedeflere odaklanması önerilmektedir:

- **Toplamsal (additive) bileşen birleştirme:** Kuantum rezonansının Gaussian sönümlenme tarafından yutulmasını önleyen toplamsal logit-bias formülasyonu (Bö-

lüm 3.2.5).

- **Katman-seici uygulama:** Modülasyonun hangi katman alt kümesinde en etkili olduğunu belirleyen dinamik katman yönlendirmesi.
- **DSI ve dağılımsal metrikler:** Saya-tabanlı proxy’ler yerine, doğrulanmış anlamsal yaratıcılık ölçütlerinin kullanılması.
- **Karşılaştırma temeli:** BAW’ın sıcaklık ölçeklendirme ve top-p gibi post-hoc yöntemlerle doğrudan kıyaslanması.

Tekrar-üretilebilirlik ve Kod Erişimi: Bu tezde raporlanan deneysel sonuçların tekrar üretilebilirliğini desteklemek amacıyla deęişken tohumlama (seed) stratejisi, sabit kütüphane sürümleri, kaynak kodlar, alıştırma betikleri ve prompt seti <https://github.com/cagribilgehan/baw-transformer> adresinde kamuya açık olarak paylaşılmıştır.

6. KAYNAKLAR

- Bahdanau D, Cho K, Bengio Y, 2015, Neural machine translation by jointly learning to align and translate, International Conference on Learning Representations.
- Bellemare-Pepin A, Lespinasse F, Thölke P, Harel Y, Mathewson K, Olson J A, vd., 2024, Divergent Creativity in Humans and Large Language Models, arXiv preprint arXiv:2405.13012.
- Beltagy I, Peters M E, Cohan A, 2020, Longformer: The long-document transformer, arXiv preprint arXiv:2004.05150.
- Boden M A, 2004, The creative mind: Myths and mechanisms, Routledge.
- Chen S Y-C, Yang C-H H, Qi J, Chen P-Y, Ma X, Goan H-S, 2020, Variational Quantum Circuits for Deep Reinforcement Learning, IEEE Access, 8, 141007–141024.
- Chen C-S, Kuo E-J, 2025, Quantum adaptive self-attention for quantum transformer models, arXiv preprint arXiv:2504.05336.
- Chi T-C, Fan T-H, Ramadge P J, Rudnicky A I, 2022, Kerple: Kernelized relative positional embedding for length extrapolation, Advances in Neural Information Processing Systems, 35, 8386–8399.
- Child R, Gray S, Radford A, Sutskever I, 2019, Generating long sequences with sparse transformers, arXiv preprint arXiv:1904.10509.
- Choromanski K, Likhoshesterov V, Dohan D, Song X, Gane A, Sarlos T, vd., 2021, Rethinking attention with performers, International Conference on Learning Representations.
- Clark K, Khandelwal U, Levy O, Manning C D, 2019, What does BERT look at? An analysis of BERT’s attention, Proceedings of the 2019 ACL Workshop BlackboxNLP, 276–286.
- Coecke B, de Felice G, Meichanetzidis K, Toumi A, 2020, Foundations for near-term quantum natural language processing, arXiv preprint arXiv:2012.03755.

- Dai Z, Yang Z, Yang Y, Carbonell J, Le Q V, Salakhutdinov R, 2019, Transformer-XL: Attentive language models beyond a fixed-length context, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2978–2988.
- Dathathri S, Madotto A, Lan J, Hung J, Frank E, Molino P, vd., 2020, Plug and play language models: A simple approach to controlled text generation, International Conference on Learning Representations.
- Fan A, Lewis M, Dauphin Y, 2018, Hierarchical neural story generation, Proceedings of the 56th Annual Meeting of the ACL, 889–898.
- Gu Z, Yao J, Du K, Jiang J, 2024, LLMSteer: Improving long-context LLM inference by steering attention on reused contexts, arXiv preprint arXiv:2411.13009.
- Holtzman A, Buys J, Du L, Forbes M, Choi Y, 2020, The curious case of neural text degeneration, International Conference on Learning Representations.
- Johnson D R, Kaufman J C, Baker B S, Patterson J D, Barbot B, Green A E, vd., 2023, Divergent semantic integration (DSI): Extracting creativity from narratives with distributional semantic modeling, Behavior Research Methods, 55, 3726–3759.
- Katharopoulos A, Vyas A, Pappas N, Fleuret F, 2020, Transformers are RNNs: Fast autoregressive transformers with linear attention, ICML, 5156–5165.
- Kitaev N, Kaiser L, Levskaya A, 2020, Reformer: The efficient transformer, International Conference on Learning Representations.
- Li J, Galley M, Brockett C, Gao J, Dolan B, 2016, A diversity-promoting objective function for neural conversation models, Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 110–119.
- Li Q, Uprety S, Wang B, Song D, 2018, Quantum-Inspired Complex Word Embedding, Proceedings of the Third Workshop on Representation Learning for NLP, 50–57.
- Li X L, Holtzman A, Fried D, Liang P, Eisner J, Hashimoto T, vd., 2023, Contrastive decoding: Open-ended text generation as optimization, Proceedings of the 61st Annual Meeting of the ACL, 12286–12312.

- Li S, You C, Guruganesh G, Ainslie J, Ontañón S, Zaheer M, vd., 2023, Functional Interpolation for Relative Positions Improves Long Context Transformers, arXiv preprint arXiv:2310.04418.
- Liu A, Sap M, Lu X, Swayamdipta S, Bhagavatula C, Smith N A, vd., 2021, DExperts: Decoding-time controlled text generation with experts and anti-experts, Proceedings of the 59th Annual Meeting of the ACL, 6691–6706.
- Lu L-C, Chen S-J, Pai T-M, Yu C-H, Lee H-Y, Sun S-H, 2024, LLM discussion: Enhancing the creativity of large language models via discussion framework and role-play, arXiv preprint arXiv:2405.06373.
- Luong M-T, Pham H, Manning C D, 2015, Effective approaches to attention-based neural machine translation, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 1412–1421.
- McCarthy P M, Jarvis S, 2010, MTL D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment, Behavior Research Methods, 42, 381–392.
- Meister C, Pimentel T, Wiher G, Cotterell R, 2023, Locally typical sampling, Transactions of the Association for Computational Linguistics, 11, 102–121.
- Meng K, Bau D, Andonian A, Belinkov Y, 2022, Locating and editing factual associations in GPT, Advances in Neural Information Processing Systems, 35, 17359–17372.
- Nath S S, Dayan P, Stevenson C, 2024, Characterising the creative process in humans and large language models, arXiv preprint arXiv:2405.00899.
- Olson J A, Nahas J, Chmoulevitch D, Cropper S J, Webb M E, 2021, Naming unrelated words predicts creativity, Proceedings of the National Academy of Sciences, 118, e2022340118.
- Press O, Smith N A, Lewis M, 2022, Train short, test long: Attention with linear biases enables input length extrapolation, International Conference on Learning Representations.

- Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I, 2019, Language models are unsupervised multitask learners, OpenAI Technical Report.
- Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M, vd., 2020, Exploring the limits of transfer learning with a unified text-to-text transformer, *Journal of Machine Learning Research*, 21, 1–67.
- Reimers N, Gurevych I, 2019, Sentence-BERT: Sentence embeddings using Siamese BERT-networks, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3982–3992.
- Runco M A, Jaeger G J, 2012, The standard definition of creativity, *Creativity Research Journal*, 24, 92–96.
- Shaw P, Uszkoreit J, Vaswani A, 2018, Self-attention with relative position representations, *Proceedings of NAACL-HLT 2018*, 464–468.
- Shazeer N, Mirhoseini A, Maziarz K, Davis A, Le Q V, Hinton G E, vd., 2017, Outrageously large neural networks: The sparsely-gated mixture-of-experts layer, *arXiv preprint arXiv:1701.06538*.
- Su J, Lu Y, Pan S, Murtadha A, Wen B, Liu Y, 2024, RoFormer: Enhanced transformer with rotary position embedding, *Neurocomputing*, 568, 127063 (İlk arXiv önbaskısı: 2021, arXiv:2104.09864).
- Sun Y, Dong L, Patra B, Ma S, Huang S, Benhaim A, vd., 2023, A length-extrapolatable transformer, *Proceedings of the 61st Annual Meeting of the ACL*, 14590–14604.
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, vd., 2017, Attention is all you need, *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Zaheer M, Guruganesh G, Dubey A, Ainslie J, Alberti C, Ontañón S, vd., 2020, Big Bird: Transformers for longer sequences, *Advances in Neural Information Processing Systems*, 33, 17283–17297.

- Zhang P, Niu J, Su Z, Wang B, Ma L, Song D, 2018, End-to-End Quantum-like Language Models with Application to Question Answering, Proceedings of the AAAI Conference on Artificial Intelligence, 32, 5666–5673.
- Zhang Q, Singh C, Liu L, Liu X, Yu B, Gao J, vd., 2023, Tell your model where to attend: Post-hoc attention steering for LLMs, arXiv preprint arXiv:2311.02262.
- Zhu Y, Lu S, Zheng L, Guo J, Zhang W, Wang J, vd., 2018, Texus: A benchmarking platform for text generation models, arXiv preprint arXiv:1802.01886.

ÖZGEÇMİŞ

Adı Soyadı : Fatih Çağrı BİLGEHAN
Doğum Yeri ve Tarihi : Mersin / 03.12.1986
Yabancı Dili : İngilizce
İletişim (Tel/e-posta) : cagribilgehan@gmail.com

Eğitim Durumu

Lise : Özel Erciyes Lisesi, 2004
Lisans : Erciyes Üniversitesi İnşaat Mühendisliği, 2013
Yüksek Lisans : Afyon Kocatepe Üniversitesi, Fen Bilimleri Enstitüsü,
İnternet ve Bilişim Teknolojileri Yönetimi Anabilim Dalı, 2026

Çalıştığı Kurum(lar)

: Uşak Üniversitesi (2020 – Devam Ediyor)
: Pidü (2018 – 2020)
: Pifors Ltd. (2016 – 2018)
: İstanbul Medeniyet Üniversitesi (2013 – 2016)