

Ultra-High Extinction Ratio Optical Modulation via Interference-Engineered Mach–Zehnder Modulator

Corresponding Author: Ms Zhengying Lou

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Review Comments

In this paper, the authors demonstrate a dual-branch ER-boost Mach–Zehnder modulator (MZM) that achieves an extinction ratio (ER) of 38 dB. The proposed configuration preserves the simplicity and broadband nature of a standard MZM. The study is carefully designed, well executed, and the results are clearly presented. I recommend a minor revision and publication after the authors address the following comments:

1. In the Introduction, the authors argue that high ER is essential for a wide range of applications, including coherent communication, photonic computing, quantum key distribution (QKD), and ultra-low-power interconnects. However, the ER requirements for these systems differ substantially. For example, in classical high-speed optical communication, an ER of this magnitude exceeds practical requirements; instead, electro-optic efficiency (e.g., $V_{\pi L}$) at high speed is typically more critical. In contrast, continuous-variable QKD indeed benefits from extremely high ER. It would strengthen the manuscript if the authors clarified which specific application spaces this ER-boost MZM is primarily intended to target.
2. In Fig. 1d, the reported ER appears robust against variations in the power-splitting ratios of the individual branches (γ_m or γ_e). However, the global power-split ratio of the initial Y-splitter (γ) is not explicitly discussed. What is the overall splitting ratio at the first splitter, and how sensitive is the ER to deviations in this parameter? A discussion of the tolerance of γ would be helpful.
3. Regarding Fig. 3c, is the plot showing the first ~70 seconds of bias-voltage evolution during optimization? It would improve clarity if the authors explained why the M-arm voltage exhibits oscillatory behavior during the first 15 seconds before reaching convergence.
4. My understanding is that the ER optimization is performed under DC bias conditions in the ER period. If so, when the device is operated with high-speed RF signals, does RF roll-off influence the achievable ER or hinder the optimization procedure? The presented experiments are limited to 400 MHz, where DC-bias-based optimization likely remains effective. At multi-GHz operation (1–10 GHz and above), will the same optimization method still be valid? If not, the authors should discuss potential approaches for optimizing the ER-boost MZM under high-speed modulation.
5. Linearity is an important performance metric for MZMs. Because the ER-boost design uses an additional interference branch to suppress residual optical leakage, how does this configuration affect the modulator's linearity relative to a standard MZM without the enhancement arm? Are there any trade-offs between ER enhancement and linearity that should be considered?

Reviewer #2

(Remarks to the Author)

In this work, the authors demonstrated a high-extinction-ratio Mach-Zehnder modulator by adding an auxiliary path. They further developed an FPGA- and microcontroller-based control board that enables bias control and real-time stabilization of

extinction ratio. Overall, this work has practical value and aligns well with the scope of Communications Engineering. Therefore, I recommend its publication after a major revision. My detailed comments are listed below.

1. I suggest not using the abbreviation "ER" in the title, since it is not a generally known term.
2. The drawbacks of this structure are not discussed. What is the insertion loss of this modulator? Is it higher than that of the single-branch structure? What is the power splitting ratio between the M-branch and the E-branch at the device input/output? Given that the E-branch is designed to generate an auxiliary field to cancel the weak residual leakage of the M-branch in the OFF state, the optical power propagating through this auxiliary arm must be significantly attenuated to match the low leakage level. Consequently, if a substantial fraction of the input power is distributed to the E-branch, a large portion of the optical energy would inevitably be dissipated, leading to a significant increase in the total insertion loss.
3. While the reported extinction ratio of 38 dB is impressive, the improvement enabled by this dual-branch structure is not clear. Have the authors also fabricated a standard (single-branch) modulator and compared the results? In addition, what factors are limiting further improvement of the extinction ratio? Is the electronic noise (e.g. PD and TIA) limiting the measured extinction ratio?
4. Optical power meters, which typically have large dynamic ranges, can be used to accurately characterize the extinction ratio under static conditions. Have the authors characterized the extinction ratio using an optical power meter?
5. More design details of the modulator should be provided, such as the waveguide width, layer thickness, and device length/width.
6. Fig. 3c: Is the unit of x-axis truly "second"? I also suggest indicating the three steps in this figure for clarity.
7. Fig. 5c: Are feedback controls applied to lock the extinction ratio to these values? In earlier sections, an ER above 38 dB is obtained. What is the reason for using a maximum ER of 32.1 dB here?
8. Supplementary, page 2: Please check whether "I-branch" and "Q-branch" are the proper terms here.

Reviewer #3

(Remarks to the Author)

The authors describe a novel and efficient implementation to improve extinction ratios of electro-optic modulators, limited by fabrication tolerances and drift. They practically demonstrate the feasibility and usefulness of the approach.

They show in detail the measured absolute performance of the device. It would help the reader if they would provide the relative performance improvement by

- A) Stating the native ER of a typical MZM fabricated in the particular platform (absolute ER improvement), and
- B) Comparing the ER drift over time (Fig. 4) with the real-time tracking on/vs off (ER stability improvement).

Below I have listed minor recommendations and requests for clarification.

1. In the abstract, the authors claim that ultra-high ERs are crucial to energy efficiency in classical optical communication systems. This claim is not further supported in the main text and in my opinion incorrect. I recommend to either remove the part "...both classical and" or to clarify under which particular circumstances ultra-high ERs are benefiting energy efficiency in classical communication systems.
2. In the introduction, the authors cite ref 18 and claim that they show that "...high ER is essential to suppress optical leakage and preserve quantum-state fidelity." I see in ref 18 an implementation of how to suppress optical leakage, but not an explanation why it is essential. I kindly ask the authors to either provide a reference that shows the essentiality, or show where ref 18 demonstrates essentiality, or rephrase the statement.
3. In the introduction, the authors implicitly claim to overcome the limitations of thermal sensitivity. At the end of the section they claim stability under "environmental perturbations". They don't show nor discuss thermal stability of their solution. If practical, it would be beneficial to the manuscript to show a time evolution of ER like in Fig. 4, but show with vs. without real-time tracking, and turn on a thermal perturbation in certain time intervals.
4. In the results section, there is a single closing bracket in line two. Presumably a typo.
5. In the "system architecture" section, the modulator is explained. It would help the reader to explain, that push-pull modulation is used and that the data signal is unipolar and differential, i.e. balanced.
6. In Fig. 1a, the configuration of the modulator electrodes is explained, but should be improved:
 - a. The V_{RF} is a differential signal, probably at least after the amplifier, this is not clear from the picture.
 - b. The V_{RF} appears to affect the upper arm twice as strongly as the lower one. My guess is, this is not intentional.
 - c. One arm of each of the bias electrodes seem to be grounded. It is not clear, whether these signals are differential.
 - d. V_{RF} is the only signal that is visually connected to the corresponding modulator.
 - e. The line labelled "electrical signal" is not the only electrical signal. "Data signal" would be a more adequate label to distinguish it.
 - f. The label "QPSK Drive Link" makes no sense in an intensity modulation scheme
7. The caption of Fig.1, subsection 2.2 and the supplement make use of the "power splitting ratios", labelled with the lowercase Greek letter gamma. It is not clear to me, what this should designate. The authors should define the figure. Similarly, the term "imbalance factor" is used without definition.
8. Subsection 2.2 explains the operation point of the MZM. It should be clarified, that intensity modulation is used in combination with unipolar differential drive signals, and that $V_{bias,M}$ is used to tune the static π phase shift between the arms of the M-branch.
9. In subsection 2.3, the optimization of the M-branch is described. The description refers to a saddle-shape in Fig.2c, which cannot be identified. I am sceptical of the plot in Fig.2c, as I would expect that at $V_{pp}=V_{pi}$ the optical average power is independent of $V_{bias,M}$. This can be derived from eq. 10 in the supplement by averaging $P_{out}(V_{in}=0V)$ and $P_{out}(V_{in}=V_{pi})$, where the result is constant and independent of $V_{bias,M}$. The authors should check the figure and indicate the saddle-point. They should also check the caption, the descriptions of b and c are incomprehensible, probably due to typos.

10. Fig. 3 and the caption of Fig. 3 show dynamic tuning of $V_{\text{bias},M}$ in the ER, while the list in subsection 2.3 mentions explicitly, that only $V_{\text{bias},E}$ and $V_{\text{bias},P}$ are adjusted in the ER period. Then in the caption of Fig. 4 it explained that the M-path is not tracked in real-time. The authors should resolve the inconsistencies. Ideally, Fig. 4 should compare the performance under recalibration of all paths to the recalibration without M-path and the performance without closed-loop control and apply a comparable thermal disturbance after a set amount of time.
11. Subsection 2.4 should mention the dynamic range of the SNSPD-based method, the maximum detectable ER, or a similar figure to compare the measured ER with the sensitivity of the setup.
12. In Figure 5c. and d. the colour scheme for the ER and the order in the legend is inconsistent. This would help to identify the correct measurements when comparing.
13. Section 3 should mention the relevant capabilities of the particular LNOI technology, i.e., the analog bandwidth of the MZM, the native ER of a typical MZM, the V_{π} .
14. Section 2 or 3 should mention whether the bias signals use the Pockels effect or whether thermo-optic phase shifters are used (or any other technology).
15. In the supplement, section 3 mentions the dropping of the subscript n and then the same two symbols are copied twice. - Probably a typo.

Reviewer #4

(Remarks to the Author)

The authors present a dual-parallel MZM architecture, where they use one branch to cancel out residual power of the other branch in the “off”-state. By doing so, the authors achieve on-off keying with 35 dB extinction ratio (ER) over 25 minutes. The device is use photon-counting based optical communication and achieve a bit-error-rate (BER) below 10⁻³.

Although the results are technically sound, in its current form, the manuscript lacks a comparison to current state-of-the-art and overall a better discussion on why the ER-boost MZM presented is better suited than other approaches. Therefore I would reject the manuscript in its current form and the authors should address the following points before resubmission:

1. Ultra-high ER modulation has been explored on various platforms. A more traditional approach uses two tunable beam-splitters instead of fixed Y-branches for the power splitting into the MZM. With such an architecture, Jin et al. (Opt. Lett. 44, 1265 (2019)) have achieved ~50 dB extinction with a broad optical bandwidth of 100 nm on the thin film lithium niobate (TFLN) platform. Similarly, cascaded MZI devices on silicon have achieved up to 60 dB ER (Wilkes et al. Opt. Lett. 41, 5318 (2016)). Although on silicon, the thermo-optic effect is used, the overall principle should be translatable to lithium niobate (LN). The authors should compare their approach to these cascaded architectures in terms of performance, fabrication tolerance, etc.
2. Along similar lines, in the introduction, the authors claim: “Unlike conventional static bias optimization or thermal tuning, the ER-Boost structure enables dynamic extinction enhancement through an electronically reconfigurable phase control, allowing the system to continuously track and compensate for phase drift.” Also in the discussion, they write: “The key advantage of the proposed dual-branch configuration lies in its active compensation of structural imbalance and dynamic phase drift through a reconfigurable interference mechanism.” To me it is not clear in how far the cascaded approaches mentioned in my previous point would not allow for a similar reconfigurable tuning mechanism.
3. Also in the introduction, the authors write “Recent developments in silicon photonics and hybrid material platforms have enabled on-chip integration and ultra-high ERs exceeding 40 dB [19,20]” in the introduction. First of all, their reference [20], besides only showing simulation results, is a meta-surface which is a free space application and not on-chip integrated. Additionally, there are other demonstrations of ultra-high ER modulation which should be cited:
 - Jin et al. Opt. Lett. 44, 1265 (2019)
 - Pohl et al. IEEE Photon. Technol. Lett. 33, 85 (2021)
4. The performance metrics reported in this contribution are subpar compared to state-of-the-art literature:
 - The reported ER of up to 38 dB is lower than the one reported by Jin et al. (Opt. Lett. 44, 1265 (2019)), which achieve up to 53 dB in a cascaded TFLN modulator, or Nelan et al. (IEEE Access 10, 100300 (2022)) which achieve 45 dB in a single TFLN modulator.
 - The experiments are performed at 400 MHz which is orders of magnitude lower than what TFLN is capable of (e.g. Wang et al. Nature 562, 101 (2018) and Pohl et al. IEEE Photon. Technol. Lett. 33, 85 (2021), both demonstrating bandwidths of 60 GHz and more)
 - The optical bandwidth is claimed to be high, however is not measured. Other high ER schemes reach bandwidth of 100 nm (Jin et al. Opt. Lett. 44, 1265 (2019))
 - CMOS compatible V_{π} have been demonstrated for example in Wang et al. Nature 562, 101 (2018). The current manuscript does not report V_{π} .

Ideally, the authors should measure the optical and electro-optic bandwidth and the V_{π} of their modulator and comment on the limiting factors.
5. The use of SNSPDs in the optical communication experiments makes the usage of their device in real-world optical communication links questionable. The authors should provide more context on the targeted applications. Can they also comment on if their enhanced ER would influence the link quality (e.g. the BER) if standard photodiodes would be used and whether their stepwise control strategy would produce the same results?

Overall, the manuscript mostly lacks the proper context of the current state-of-the-art of TFLN modulators and high-ER switching for optical communications, which is crucial for understanding and assessing the impact and quality of the demonstrated device architecture. In addition to these major points, I would also like to raise the following more minor points:

1. Can the authors provide more detail on the TFLN modulator? Was it fabricated in-house or by an external service? What is the film thickness, crystal cut, waveguide and electrode dimension?
2. More details on the communication experiments should be provided. Specifically the data rate and the link length should be mentioned.
3. For the bit-error-rates (BER) reported in fig. 5D, can the authors indicate the hard- and soft-decision forward error correction limits and whether they were able to reach those limits in their experiment?
4. The authors should provide the values of γ_E and γ_M used in their design and consider indicating those in fig. 1d. This would help the reader understand the better tolerance to fabrication imperfections.
5. In section 2.4, the “ultralow dark count rate” is mentioned. What was the dark count rate and the average rates during the signal and guard slots?

Reviewer #5

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #6

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have addressed all of my previous comments with detailed explanations. The manuscript has been significantly strengthened as a result, and I recommend it for publication in Communication Engineering.

Reviewer #2

(Remarks to the Author)

The authors have addressed my concerns. I now recommend the publication of this manuscript.

Reviewer #3

(Remarks to the Author)

I thank the authors for careful consideration of my comments. Most suggestions were implemented, questions answered, and inconsistencies resolved. Only a few comments remain.

1) My comment (reviewer 3, comment 2) regarding the phrase “...both classical and...” in the abstract was addressed in the “rebuttal”, but the corresponding phrase was not removed. See Abstract, end of line 2.

2) My comment (comment 7a) regarding the differential nature of the drive signal was addressed in the “rebuttal”, but the corresponding line in the figure is not explicitly drawn as a differential signal. The authors claim in the “rebuttal” that “the differential nature of the RF drive after amplification is now explicitly indicated”. – I would expect separate p and n (or S and S*) channels between the amplifier and the modulator.

3) My comment (comment 10) regarding Fig. 2c was addressed but not fully resolved.

a. The reference to a saddle-shaped surface in Fig. 2c (main text and caption) was not explained. I cannot see a saddle-point in Fig. 2c. I kindly ask the authors to indicate the saddle-point in the figure.

b. The typo remains, I think the authors meant to write “showing a sharp extinction boundary indicated with “E channel cutoff point”. c, Measured”

4) My comment (comment 11) regarding Fig. 3 was addressed but not fully resolved. The caption still claims “The bias controller dynamically tunes Vbias,M, Vbias,E, and Vbias,P during the ER period”, while the “rebuttal” states that “During Stage III (ER optimization), only Vbias,E and Vbias,P are adjusted, while the M-path bias is held constant.”

5) My comment (comment 13) regarding the legends of Fig. 5c-d was addressed but not fully resolved. The authors claim in the “rebuttal”, that “the order of the legend entries has been unified to follow the same ascending ER sequence for clarity”, but the legend entries are not in ascending order. The color scheme was unified, but the markers are still inconsistent between the two plots.

6) The caption of Fig. 5 uses brackets to identify subfigures, i.e. “(a)”, while the other captions use a comma, i.e. “a,”. This should be unified.

7) Table 2 is highly appreciated, but there is probably a typo in row 4, it should presumably read “DC ports”

8) Typo: In section 2.2 in the second paragraph, the first sentence ends with a double period.

Reviewer #4

(Remarks to the Author)

The authors have substantially improved the manuscript in response to my and the other referees' comments. The revisions not only enhance the clarity and flow of the manuscript but also place the work better in the current state-of-the-art in the field. Most of the concerns raised in the previous review round have been addressed and clarified. As a result, the manuscript is now more coherent and compelling, representing a clear advance over the original submission. In particular, the clarification that a ultra-high dynamic extinction ratio (ER) is the target here distinguishes their work from other cited literature and highlights the novelty of their work better. Therefore I would recommend publication, the authors may choose to address some of the comments below if they find it improves the manuscript:

- 1) Although in the manuscript they clearly distinguish between static and dynamic ER, in the abstract the authors may want to include this by including the word "dynamic".
- 2) In fig. 1c) the legend overlaps more than half the actual graph. It would be better to place the legend elsewhere.

Reviewer #5

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #6

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

I thank the authors for carefully addressing the remaining comments and requests. The manuscript has been significantly strengthened. I recommend it for publication in Communication Engineering.

Reviewer #4

(Remarks to the Author)

The authors have addressed all my remaining comments and I can recommend the manuscript for publication.

Reviewer #5

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Dear Editor Wenjie Wang,

We are writing to resubmit our manuscript entitled “Ultra-high extinction ratio modulation without perfect splitters using the ER-Boost Mach–Zehnder modulator” for consideration for publication in Communications Engineering with **COMMS-25-0880-T**. We would like to express our sincere gratitude to you and the reviewers for the comprehensive and insightful feedback provided for our work.

We are encouraged by the reviewers’ positive comments regarding the significance of achieving ultra-high extinction ratios (ER) for both classical and quantum optical systems. In this revised manuscript, we have carefully addressed all the technical and presentational points raised by the reviewers. Specifically, we have provided further clarification on the structural tolerance of our ER-Boost architecture and included more detailed discussions on the long term stability of our adaptive bias control system.

All major changes, including updated experimental analysis and expanded theoretical derivations, are highlighted in blue in the revised manuscript and summarized within the response boxes in this letter. Please note that the citations in this letter refer to the references listed at the end of this document, rather than those in the manuscript itself, unless stated otherwise. Similarly, figure indexing in this letter is independent of the manuscript.

We believe that the revisions have significantly strengthened the paper and addressed all the concerns raised. Thank you for your time and for the opportunity to improve our work.

Yours Sincerely,

Ziyuan Shi, Zhengying Lou, Jiajie Xu, and Mohamed-Slim Alouini

I. REVIEWER 1'S COMMENTS

We sincerely thank the reviewer for the thorough evaluation of our manuscript and for the constructive and insightful comments. The suggestions have helped us clarify the application scope, strengthen the technical rigor, and improve the presentation of the proposed ER-Boost architecture. In the revised version, we have:

(i) clarified the primary target application scenarios of the ER-Boost MZM, emphasizing photon-counting and quantum-inspired communication systems where ultra-high extinction ratio directly impacts system-level performance;

(ii) provided additional analysis on the tolerance of the global power-splitting ratio and discussed the robustness of the architecture against structural imbalance;

(iii) clarified the hierarchical bias optimization procedure and explained the transient oscillatory behavior observed during the initial convergence stage;

(iv) expanded the discussion on high-speed operation, including the influence of RF roll-off and potential strategies for broadband extinction optimization;

(v) added a dedicated discussion on linearity considerations and clarified that the ER-Boost architecture does not intrinsically degrade the modulation linearity of the underlying MZM.

We believe these revisions have significantly strengthened the technical clarity and contextual positioning of the work. We are grateful for the reviewer's careful reading and valuable suggestions.

Comment 1 *In the Introduction, the authors argue that high ER is essential for a wide range of applications, including coherent communication, photonic computing, quantum key distribution (QKD), and ultra-low-power interconnects. However, the ER requirements for these systems differ substantially. For example, in classical high-speed optical communication, an ER of this magnitude exceeds practical requirements; instead, electro-optic efficiency (e.g., $V_{\pi L}$) at high speed is typically more critical. In contrast, continuous-variable QKD indeed benefits from extremely high ER. It would strengthen the manuscript if the authors clarified which specific application spaces this ER-boost MZM is primarily intended to target.*

Response:

We thank the reviewer for this insightful and constructive comment. We fully agree that the extinction ratio (ER) requirements differ significantly across application scenarios and that,

in many classical high-speed optical communication systems, parameters such as electro-optic efficiency (e.g., V_π), bandwidth, and linearity are often more critical than achieving an ultra-high ER.

Our intention in the Introduction was to provide a broad context for the relevance of ER in diverse photonic systems. However, we acknowledge that the manuscript would benefit from a clearer specification of the primary target application spaces of the proposed ER-Boost MZM.

We would like to clarify that the ER-Boost architecture is particularly motivated by weak-signal and photon-sensitive scenarios where residual optical leakage directly limits system performance, including:

- photon-counting and pulse-position modulation (PPM) links, in which leakage in guard slots reduces modulation contrast and degrades error performance;
- quantum communication systems, especially continuous-variable QKD and related coherent protocols, where background leakage affects quantum-state fidelity and security margins;
- ultra-low-power optical interconnects and optical computing schemes, where leakage reduces the available dynamic range and energy efficiency.

In contrast, for conventional high-power intensity-modulation/direct-detection (IM-DD) or coherent links, an ER in the 10–15 dB range is often sufficient, and design trade-offs are typically dominated by bandwidth, and linearity rather than by pushing the ER into the 30–40 dB regime.

To address the reviewer’s suggestion, we have revised the manuscript to more clearly define the intended application scope of the proposed ER-Boost MZM. In particular, we have added explicit clarifications in both the Introduction and the Discussion sections to emphasize that our work is primarily motivated by photon-counting and quantum-inspired communication systems, where ultra-high ER directly translates into measurable system-level performance gains. The corresponding additions are provided below in the “Changes in the manuscript” section.

Changes in the manuscript:

(1) Added paragraph at the end of the Introduction section:

While the ER-Boost concept is, in principle, applicable to a broad class of interferometric modulators, the experiments and analysis in this work are specifically tailored toward photon-counting and quantum-inspired communication scenarios, where the achievable ER directly impacts the required average photon number per slot and thus the overall link budget. The

SCPPM-based photon-counting transmission experiments presented in this paper serve as a representative example of such systems, illustrating how the proposed architecture enables more than two orders of magnitude improvement in BER at a fixed received photon level by boosting the ER.

(2) Added clarification in the Discussion section:

Consequently, the ER-Boost MZM is particularly suited for photon-counting receivers, PPM modulation formats, and CV-QKD-like protocols, where leakage suppression is directly tied to sensitivity, security margins, and dynamic range. In such systems, improving the ER beyond conventional levels yields immediate system-level performance gains, in contrast to high-power classical links where electro-optic efficiency and bandwidth often dominate the design trade-offs.

Comment 2 *In Fig. 1(d), the extinction ratio (ER) is shown to be robust against variations in the power-splitting ratios of the individual branches (γ_m and γ_e). However, the global power-splitting ratio of the initial Y-splitter (γ) is not explicitly addressed. Please clarify the nominal splitting ratio of the first-stage splitter and discuss the sensitivity of the ER to deviations in γ . An analysis of the tolerance to variations in γ would further strengthen the manuscript.*

Response:

We thank the reviewer for this important observation. The nominal splitting ratio of the first-stage Y-splitter in our fabricated device is designed to be approximately 50:50. However, the ER-Boost architecture does not critically rely on an exact 50:50 global splitting ratio.

The key requirement for effective extinction enhancement is that the auxiliary (E) branch must provide sufficient optical field amplitude to cancel the residual leakage from the modulation (M) branch in the OFF state. Let the global splitting ratio be γ , such that the M-branch receives γP_0 and the E-branch receives $(1 - \gamma)P_0$, where P_0 is the input optical power.

If the intrinsic ER of the M-branch is limited (for example, 10 dB), the residual leakage power in the OFF state is approximately $0.1\gamma P_0$. Successful cancellation requires that the E-branch can

supply an optical field of comparable amplitude, leading to the condition

$$(1 - \gamma)P_0 \gtrsim 0.1\gamma P_0.$$

This yields $\gamma \lesssim 0.91$, indicating that as long as the global splitter is not extremely biased toward the M-branch, sufficient cancellation amplitude is available.

Therefore, moderate deviations of the global splitting ratio from 50:50 (e.g., within a wide range around balanced splitting) do not fundamentally limit the achievable ER. Only in extreme imbalance cases, where the E-branch receives insufficient optical power to compensate for the M-branch leakage, would the maximum achievable ER be reduced.

Changes in the manuscript:

To clarify the tolerance of the ER-Boost architecture to variations in the global splitting ratio of the first-stage Y-junction, we have added a brief discussion in Section 2.2 (Structural Design and Imbalance Guidelines). The added text is reproduced below:

In addition to the internal splitting ratios of the M- and E-branches, the global power-splitting ratio of the first-stage Y-junction determines the relative optical power available for modulation and extinction enhancement. The ER-Boost mechanism requires only that the E-branch receives sufficient optical power to cancel the residual leakage from the M-branch in the OFF state. Therefore, as long as the global splitter is not extremely unbalanced toward the modulation branch, moderate deviations from the nominal 50:50 ratio do not fundamentally limit the achievable extinction ratio. Only in highly asymmetric cases, where the E-branch receives insufficient power for effective cancellation, would the maximum ER become constrained.

Comment 3 *Regarding Fig. 3c, is the plot showing the first 70 seconds of bias-voltage evolution during optimization? It would improve clarity if the authors explained why the M-arm voltage exhibits oscillatory behavior during the first 15 seconds before reaching convergence.*

Response:

We thank the reviewer for this careful observation. Yes, Fig. 3c shows approximately the first 70 seconds of the initial optimization phase after system startup.

The automatic bias control follows a hierarchical optimization procedure. Since the current modulator integrates a monitoring photodetector (PD) only at the output port, all optimization

steps must rely on the total output optical power as the feedback signal. The control sequence proceeds as follows:

(1) *E-branch initialization*: The primary goal of this stage is to drive the E-branch to its effective OFF state by minimizing the output optical power. Because only the output-port PD is available, this requires iterative adjustments of the E-branch and phase-control voltages while observing the total power response.

(2) *M-branch operating-point search*: After the E-branch is initialized, the algorithm searches for the saddle-point operating condition of the M-branch to ensure balanced modulation and minimal sensitivity to bias and RF amplitude variations.

(3) *Global ER refinement*: Finally, the E- and P-branch biases are jointly fine-tuned during the dedicated ER-Boost periods to minimize the residual leakage and maximize the extinction ratio.

The oscillatory behavior observed in the first 15 seconds of Fig. 3c corresponds to the E-branch initialization stage. During this phase, the controller performs iterative bias adjustments to locate the minimum-output-power condition using only the output PD feedback. The observed voltage variations therefore reflect the transient convergence process of the optimization routine rather than instability of the system. Once the E-branch reaches its OFF state, the bias voltages stabilize and the subsequent refinement proceeds smoothly.

In future implementations, we plan to integrate additional monitoring PDs at the E-branch and M-branch outputs. With branch-level power monitoring, the E-branch minimization and M-branch saddle-point search can be performed independently using simpler control loops, eliminating the need for inserting dedicated ER-Boost periods in the modulation frame and further accelerating convergence.

Changes in the manuscript:

To improve clarity, we have added an explicit description of the control flow and convergence behavior in Section 2.3, as well as a clarification in the discussion of Fig. 3c. The added text is reproduced below.

(Added in Section 2.3 — Control flow description)

The automatic bias stabilization follows a hierarchical optimization sequence. Since the present prototype integrates a monitoring photodetector (PD) only at the output port, all

control actions rely on the total output optical power as the feedback metric. The procedure consists of three stages: (i) initialization of the E-branch by minimizing the output power to drive it into its effective OFF state; (ii) search for the saddle-point operating condition of the M-branch to ensure balanced modulation; and (iii) fine ER refinement by jointly adjusting the E- and P-branch bias voltages during the dedicated ER-Boost periods. Because the E-branch initialization is performed using iterative adjustments based solely on output power monitoring, a transient voltage variation is observed during the early convergence stage. Once the minimum-power operating point is reached, the bias voltages stabilize and the subsequent ER refinement proceeds smoothly.

(Added in the description of Fig. 3c)

The initial oscillatory behavior observed within the first ~ 15 s corresponds to the E-branch initialization phase, during which the controller iteratively minimizes the output optical power using only the output-port PD feedback. This behavior reflects the transient convergence of the optimization routine rather than instability of the bias control system.

(Added in the Discussion section)

In future implementations, integrating additional monitoring photodetectors at the individual M- and E-branch outputs would enable branch-level independent bias control. This would allow direct minimization of the E-branch power and independent saddle-point optimization of the M-branch without relying solely on the output-port feedback, thereby simplifying the control algorithm and eliminating the need to insert dedicated ER-Boost periods in the modulation frame. Such an architecture would further accelerate convergence and enhance long-term stability.

Comment 4 *My understanding is that the ER optimization is performed under DC bias conditions in the ER period. If so, when the device is operated with high-speed RF signals, does RF roll-off influence the achievable ER or hinder the optimization procedure? The presented experiments are limited to 400 MHz, where DC-bias-based optimization likely remains effective. At multi-GHz operation (1–10 GHz and above), will the same optimization method still be valid? If not, the*

authors should discuss potential approaches for optimizing the ER-boost MZM under high-speed modulation.

Response:

We thank the reviewer for raising this important point regarding high-speed operation.

The ER-Boost mechanism itself is based on interferometric cancellation and does not inherently rely on DC conditions. As long as both the modulation (M) branch and extinction-enhancement (E) branch maintain sufficient bandwidth to support the target RF frequencies, the destructive interference principle remains valid. Therefore, the ER-Boost structure does not fundamentally degrade at multi-GHz operation.

However, we agree that RF roll-off and frequency-dependent phase response can influence the practical optimization process. In the current prototype, only a single monitoring photodetector is integrated at the output port. Consequently, the bias optimization is performed using quasi-DC measurements during the dedicated ER-Boost periods. At higher RF frequencies, frequency-dependent amplitude and phase variations may lead to dynamic leakage components that are not fully captured by purely DC-based optimization.

Importantly, this limitation arises from the present control architecture rather than from the ER-Boost structure itself. In future implementations, integrating additional monitoring photodetectors at the individual M- and E-branch outputs would enable branch-level independent optimization. For example, the E-branch could be directly minimized using its dedicated PD, while the M-branch operating point could be determined through saddle-point search under high-speed modulation. Furthermore, frequency-aware calibration strategies, RF pre-emphasis, or dynamic feedback based on high-speed sampling could be employed to ensure optimal cancellation across the operating bandwidth.

Changes in the manuscript:

To clarify the applicability of the ER-Boost architecture under high-speed modulation, we have added an explicit discussion in the Discussion section addressing RF roll-off effects and optimization strategies beyond quasi-DC operation. The added text is reproduced below:

(Added in the Discussion section — High-speed operation considerations)

Although the ER optimization in the present prototype is performed during quasi-DC ER-Boost periods, the underlying interferometric cancellation principle of the ER-Boost

architecture is not restricted to low-frequency operation. At multi-GHz modulation rates, RF roll-off and frequency-dependent phase response may introduce dynamic leakage components that are not fully captured by DC-based optimization. This limitation stems from the current single-PD monitoring configuration rather than from the ER-Boost structure itself.

Future implementations incorporating branch-level monitoring photodetectors would enable independent optimization of the M- and E-branches under high-speed modulation. In addition, frequency-aware calibration, RF pre-emphasis, or dynamic feedback techniques could be adopted to maintain optimal extinction performance across wide bandwidths. Integrating additional monitoring PDs at the individual M- and E-branch outputs would enable branch-level independent control, eliminating the need for ER-Boost periods and further accelerating convergence.

Comment 5 *Linearity is an important performance metric for MZMs. Because the ER-boost design uses an additional interference branch to suppress residual optical leakage, how does this configuration affect the modulator's linearity relative to a standard MZM without the enhancement arm? Are there any trade-offs between ER enhancement and linearity that should be considered?*

Response:

We thank the reviewer for raising this important question.

First, we would like to emphasize that the present work is primarily targeted at high-contrast, binary modulation formats such as on-off keying (OOK) and pulse-position modulation (PPM) in photon-counting and quantum-inspired communication systems. In these scenarios, the dominant performance metric is the extinction ratio between the *ON* and *OFF* states rather than small-signal linearity over a wide voltage range.

In our implementation, only the modulation (M) branch is driven by the high-speed RF signal, while the extinction-enhancement (E) branch and the phase-control path (P) are biased by slowly varying or quasi-static voltages. As a result, the small-signal RF transfer function around the operating point remains essentially identical to that of a standard single MZM: the E-branch provides a static cancellation field in the *OFF* state but does not introduce additional high-frequency modulation components. Consequently, the ER-Boost configuration does not fundamentally degrade the intrinsic linearity of the MZM relative to a conventional device.

There is, therefore, no inherent trade-off in which ER enhancement is achieved at the expense of linearity. Rather, the ER-Boost structure mainly improves the suppression of the ‘OFF’ level by compensating for residual leakage, while preserving the underlying modulation characteristics of the main M-branch. In principle, the ER-Boost architecture can also be combined with existing linearization techniques (e.g., dual-parallel or push–pull configurations) for analog or multilevel modulation formats, although such applications are beyond the scope of the present work and are left for future investigation.

Changes in the manuscript:

To clarify the application scope and the impact of the ER-Boost structure on linearity, we have added a short discussion in the Discussion section. The added text is reproduced below:

(Added in the Discussion section — Linearity considerations)

The ER-Boost MZM presented in this work is primarily aimed at high-contrast binary modulation schemes, such as OOK and PPM in photon-counting and quantum-inspired communication links, where the extinction ratio between the logical ‘0’ and ‘1’ states is the dominant design metric. In the current implementation, only the modulation branch carries the high-speed RF drive, whereas the enhancement and phase-control branches are biased quasi-statically. As a result, the small-signal transfer characteristic of the main modulation path remains essentially that of a standard Mach–Zehnder modulator, and the ER-Boost configuration does not intrinsically degrade linearity. The enhancement branch mainly improves the OFF-state suppression by compensating residual leakage. In principle, the proposed architecture can be combined with established linearization techniques for analog or multilevel modulation formats, but such extensions are beyond the scope of this work and are left for future study.

II. REVIEWER 2’S COMMENTS

We sincerely thank the reviewer for the thorough and constructive evaluation of our manuscript. The comments have significantly helped us improve the clarity, rigor, and contextual positioning of this work. In the revised version, we have:

- (i) removed ambiguous terminology and improved figure clarity;
- (ii) clarified the experimental platform and explicitly summarized the baseline modulator specifications;
- (iii) added quantitative comparisons between native, static, and dynamic extinction ratios;
- (iv) expanded the discussion of insertion loss, limiting factors, and dynamic measurement sensitivity;
- (v) incorporated additional state-of-the-art references and strengthened the comparison with cascaded interferometric architectures;
- (vi) clarified the practical applicability of the proposed ER-Boost architecture beyond SNSPD-based measurements.

We believe these revisions have substantially strengthened the manuscript and better articulated the contribution of the proposed dynamic extinction enhancement architecture. We are grateful for the reviewer’s insightful suggestions.

Comment 1 *I suggest not using the abbreviation “ER” in the title, since it is not a generally known term.*

Response:

We thank the reviewer for this helpful suggestion. We agree that avoiding abbreviations in the title improves clarity and accessibility. Accordingly, we have revised the title to remove the abbreviation “ER” and replaced it with a more descriptive formulation.

Changes in the manuscript:

The revised title now reads:

Ultra-High Extinction Ratio Optical Modulation via Interference-Engineered Mach–Zehnder Modulator

Comment 2 *The drawbacks of this structure are not discussed. What is the insertion loss of this modulator? Is it higher than that of the single-branch structure? What is the power splitting ratio between the M-branch and the E-branch at the device input/output? Given that the E-branch is designed to generate an auxiliary field to cancel the weak residual leakage of the M-branch in the OFF state, the optical power propagating through this auxiliary arm must be significantly attenuated to match the low leakage level. Consequently, if a substantial fraction of the input power is distributed to the E-branch, a large portion of the optical energy would inevitably be dissipated, leading to a significant increase in the total insertion loss.*

Response:

We thank the reviewer for raising this important point regarding insertion loss and potential trade-offs.

In the present proof-of-concept prototype, the global input splitter is designed with a 50:50 power splitting ratio to simplify structural validation. As a result, the measured insertion loss is approximately 3 dB higher than that of a conventional single-branch MZM fabricated on the same platform. We emphasize that this choice was made to experimentally verify the feasibility of the ER-Boost architecture rather than to optimize insertion loss.

Importantly, the extinction-enhancement mechanism does not fundamentally require equal power distribution between the M- and E-branches. The cancellation process is based on field-level interference rather than power matching. The E-branch only needs to provide an auxiliary optical field whose amplitude matches the residual leakage field of the M-branch in the OFF state. For example, if the intrinsic extinction ratio of the M-branch is 20 dB, the residual leakage power is approximately 1% of the signal power, meaning that only a small fraction of the input optical power needs to be allocated to the E-branch for effective cancellation.

Therefore, the ER-Boost structure does not inherently impose a large insertion-loss penalty. In future optimized implementations, asymmetric splitting ratios can be adopted so that only a small fraction of optical power is directed to the enhancement branch, thereby minimizing excess loss while maintaining high extinction performance.

Changes in the manuscript:

To clarify the insertion loss and structural trade-offs, we have added a dedicated discussion in the revised manuscript. The added text is reproduced below:

(Added in the Results section — Insertion loss and structural trade-offs)

In the current proof-of-concept device, a 50:50 global power splitter is employed to simplify structural validation, resulting in an insertion loss approximately 3 dB higher than that of a conventional single-branch MZM. However, the ER-Boost mechanism does not fundamentally require equal power distribution. Since the cancellation process operates at the optical field level, the enhancement branch only needs to generate an auxiliary field comparable to the residual leakage amplitude of the modulation branch. Consequently, in optimized designs, only a small fraction of the input optical power needs to be routed to the enhancement branch, allowing the excess insertion loss to be substantially reduced.

Comment 3 *While the reported extinction ratio of 38 dB is impressive, the improvement enabled by this dual-branch structure is not clear. Have the authors also fabricated a standard (single-branch) modulator and compared the results? In addition, what factors are limiting further improvement of the extinction ratio? Is the electronic noise (e.g. PD and TIA) limiting the measured extinction ratio?*

Response:

We thank the reviewer for this important question regarding the quantitative improvement and limiting factors.

(1) Baseline comparison. The experimental platform used in this work is a commercially available Ti:LiNbO₃ dual-parallel Mach–Zehnder modulator (Fujitsu FTM7961EX). As summarized in Table I, the intrinsic extinction ratio of conventional single-branch operation is approximately 20–22 dB, consistent with the manufacturer specifications. This serves as the baseline reference for evaluating the extinction enhancement enabled by the proposed ER-Boost architecture.

(2) Static and dynamic extinction enhancement. To explicitly quantify the improvement, we have added Table II, which summarizes the measured static and dynamic extinction ratios. Under optimized DC bias conditions, the ER-Boost configuration achieves a static extinction ratio of 58 dB. Under RF modulation, the dynamic extinction ratio remains 48 dB at 100 MHz, 42 dB at 200 MHz, 38 dB at 400 MHz, and 33 dB at 1 GHz. Even at 400 MHz and 1 GHz operation, the dynamic ER remains substantially higher than the native 20–22 dB single-branch extinction ratio, demonstrating a clear enhancement enabled by the dual-branch interference mechanism.

(3) Limiting factors. The present extinction ratio is not fundamentally limited by the interference architecture itself. Instead, the dominant limiting factors are implementation-related:

(i) Only a single monitoring photodiode is available at the output port, making the simultaneous optimization of three bias voltages a multi-parameter control problem. This limits the achievable bias precision under dynamic operation.

(ii) Electronic noise in the monitoring PD and associated TIA circuitry constrains the bias control accuracy, thereby limiting the achievable dynamic ER.

(iii) RF amplitude imbalance and finite electrical bandwidth contribute to phase non-idealities at higher modulation frequencies.

Future implementations incorporating branch-level monitoring photodiodes for both the M- and E-branches are expected to significantly simplify the optimization problem and further improve dynamic extinction performance.

Changes in the manuscript:

To further clarify the limiting factors of the current implementation and the pathway toward improved dynamic extinction performance, we have expanded the discussion in three locations of the revised manuscript (Results section, Discussion section, and Control System description).

The following text has been added:

(Added statement clarifies that the currently observed dynamic ER limitation arises from the single-monitor PD configuration, in the Results section.)

To provide a clear performance baseline of the experimental platform, Table 2 summarizes the key specifications of the Ti:LiNbO₃ dual-parallel MZM used in this study, extracted from the manufacturer datasheet. The intrinsic extinction ratio of the conventional single-branch operation is approximately 20–22 dB, which serves as a reference for evaluating the extinction enhancement achieved through the proposed dual-branch interference mechanism.

(Added explanation highlights that branch-level monitoring would simplify the multi-parameter optimization and remove the need for dedicated ER-Boost periods in the modulation frame, in the bias control description.)

This frequency-dependent reduction in dynamic extinction ratio is primarily attributed

to RF amplitude imbalance, finite electrical bandwidth, and phase non-idealities in the driving circuitry, rather than to a fundamental limitation of the ER-Boost interference mechanism. Importantly, even at 400 MHz and 1 GHz operation, the dynamic ER remains substantially higher than the native static extinction ratio of the underlying single-branch modulator (approximately 20–22 dB), demonstrating the effectiveness of the proposed leakage-cancellation architecture under practical modulation conditions.

(Added text emphasizes that the ER-Boost architecture is not fundamentally bandwidth-limited and can be extended with improved control hardware, in the Discussion section.)

Future implementations incorporating branch-level monitoring photodetectors would enable independent optimization of the M- and E-branches under high-speed modulation. In addition, frequency-aware calibration, RF pre-emphasis, or dynamic feedback techniques could be adopted to maintain optimal extinction performance across wide bandwidths. Integrating additional monitoring PDs at the individual M- and E-branch outputs would enable branch-level independent control, eliminating the need for ER-Boost periods and further accelerating convergence.

Comment 4 *Optical power meters, which typically have large dynamic ranges, can be used to accurately characterize the extinction ratio under static conditions. Have the authors characterized the extinction ratio using an optical power meter?*

Response:

We thank the reviewer for this valuable suggestion.

We confirm that the static extinction ratio under DC bias conditions was verified using conventional optical power measurements. After optimization, the static extinction ratio exceeds 58 dB, as summarized in Table II.

However, the primary objective of this work is to evaluate the *dynamic* extinction ratio under high-speed modulation. Optical power meters measure average optical power and therefore cannot resolve slot-level power differences in time-synchronized PPM signals. Under high-speed

modulation, the dynamic extinction ratio must be extracted from time-resolved measurements of signal and guard slots.

For this reason, we employed the SNSPD-based dynamic slot-power reconstruction technique described in [1], which enables time-resolved extinction measurements with a dynamic range exceeding 70 dB. The SNSPD serves as a high-sensitivity measurement instrument and is not required for practical deployment of the ER-Boost architecture.

We have clarified this distinction between static and dynamic ER measurements in the revised manuscript.

Changes in the manuscript:

To explicitly clarify the quantitative improvement enabled by the ER-Boost architecture and to distinguish static from dynamic extinction performance, we have added the following text in the Results section:

To further clarify the relationship between static and dynamic extinction performance, Table II summarizes the measured extinction ratios under DC operation and under RF modulation at different frequencies. Under static (DC) bias conditions, the optimized ER-Boost configuration achieves an extinction ratio of 58 dB, confirming the high static interference suppression capability of the architecture.

This frequency-dependent reduction in dynamic extinction ratio is primarily attributed to RF amplitude imbalance, finite electrical bandwidth, and phase non-idealities in the driving circuitry, rather than to a fundamental limitation of the ER-Boost interference mechanism. Importantly, even at 400 MHz and 1 GHz operation, the dynamic ER remains substantially higher than the native static extinction ratio of the underlying single-branch modulator (approximately 20–22 dB), demonstrating the effectiveness of the proposed leakage-cancellation architecture under practical modulation conditions.

Comment 5 *More design details of the modulator should be provided, such as the waveguide width, layer thickness, and device length/width.*

Response:

We thank the reviewer for this important suggestion regarding device design details.

In this work, the experimental validation of the ER-Boost architecture was carried out using a commercially available dual-parallel Mach–Zehnder modulator (DD-MZM), specifically the Fujitsu FTM7961EX device. This Ti:LiNbO₃ DQPSK single-drive modulator naturally provides two interferometric child MZIs and a parent MZI, which allows it to be reconfigured to emulate the proposed dual-branch ER-Boost architecture.

The key specifications of the device are summarized as follows (from the manufacturer datasheet):

- Platform: Ti:LiNbO₃ integrated modulator - V_{π} (low frequency): 2.8 V - V_{π} (22 Gbps): 3.5 V - Optical insertion loss (C-band, no connector): 7.0 dB - On/off extinction ratio (parent MZI): 22 dB - On/off extinction ratio (child MZIs): 20 dB - Integrated monitor photodiode for bias control

These parameters confirm that the baseline extinction ratio of the conventional single-branch operation is limited to approximately 20–22 dB, consistent with typical LiNbO₃ MZM performance. The experimentally demonstrated 38 dB extinction ratio therefore represents a substantial improvement enabled by the reconfigured dual-branch interference mechanism and bias control strategy.

We emphasize that the primary contribution of this work is the architectural concept and system-level validation of extinction enhancement through interference engineering, rather than the optimization of a newly fabricated device geometry. In future work, we plan to design and fabricate a dedicated ER-Boost Mach–Zehnder modulator with optimized asymmetric splitting ratios and branch-level monitoring photodiodes, which will further reduce insertion loss and simplify the control algorithm.

To improve clarity and reproducibility, we explicitly state that the experimental validation of the proposed ER-Boost architecture was carried out using a commercially available dual-parallel Mach–Zehnder modulator (Fujitsu FTM7961EX), rather than a newly fabricated device. We have revised the manuscript accordingly and added a dedicated subsection describing the experimental modulator platform and optical setup.

Changes in the manuscript:

(Added subsection: Experimental Modulator Platform and Optical Setup)

The experimental validation of the proposed ER-Boost architecture was performed using

a commercially available dual-parallel Mach–Zehnder modulator (Fujitsu FTM7961EX), fabricated on a Ti:LiNbO₃ platform. This device integrates two child MZIs and a parent MZI, enabling flexible reconfiguration into the dual-branch interference-enhanced structure proposed in this work. Independent DC bias electrodes are provided for each interferometric section, allowing fine phase control and implementation of the ER-Boost bias optimization strategy.

The optical source was a tunable continuous-wave laser, polarization-maintained and coupled into the modulator via fiber arrays. The optical output was monitored simultaneously by a photodiode (PD) for high-power measurements and a superconducting nanowire single-photon detector (SNSPD) for ultra-low-power photon-counting experiments. A 1:99 optical splitter provided feedback signals for bias control and real-time extinction ratio monitoring.

Comment 6 *Fig. 3c: Is the unit of x-axis truly “second”? I also suggest indicating the three steps in this figure for clarity.*

Response:

We thank the reviewer for this helpful observation.

We clarify that the x-axis in Fig. 3c does not represent physical time in seconds. Instead, it corresponds to a normalized process time, where the horizontal axis is divided into three consecutive stages of the bias optimization procedure. The normalization was adopted to clearly illustrate the relative evolution of the bias voltages during Stage I (E-branch initialization), Stage II (M-branch optimization), and Stage III (global ER optimization), independent of the absolute time duration of each step.

To avoid ambiguity, we have revised the x-axis label to “Normalized Process Time (Stage Units)” and explicitly marked the three optimization stages in the figure for clarity.

Changes in the manuscript:

The revised Fig. 3c is shown below (Fig. 1 in the response letter).

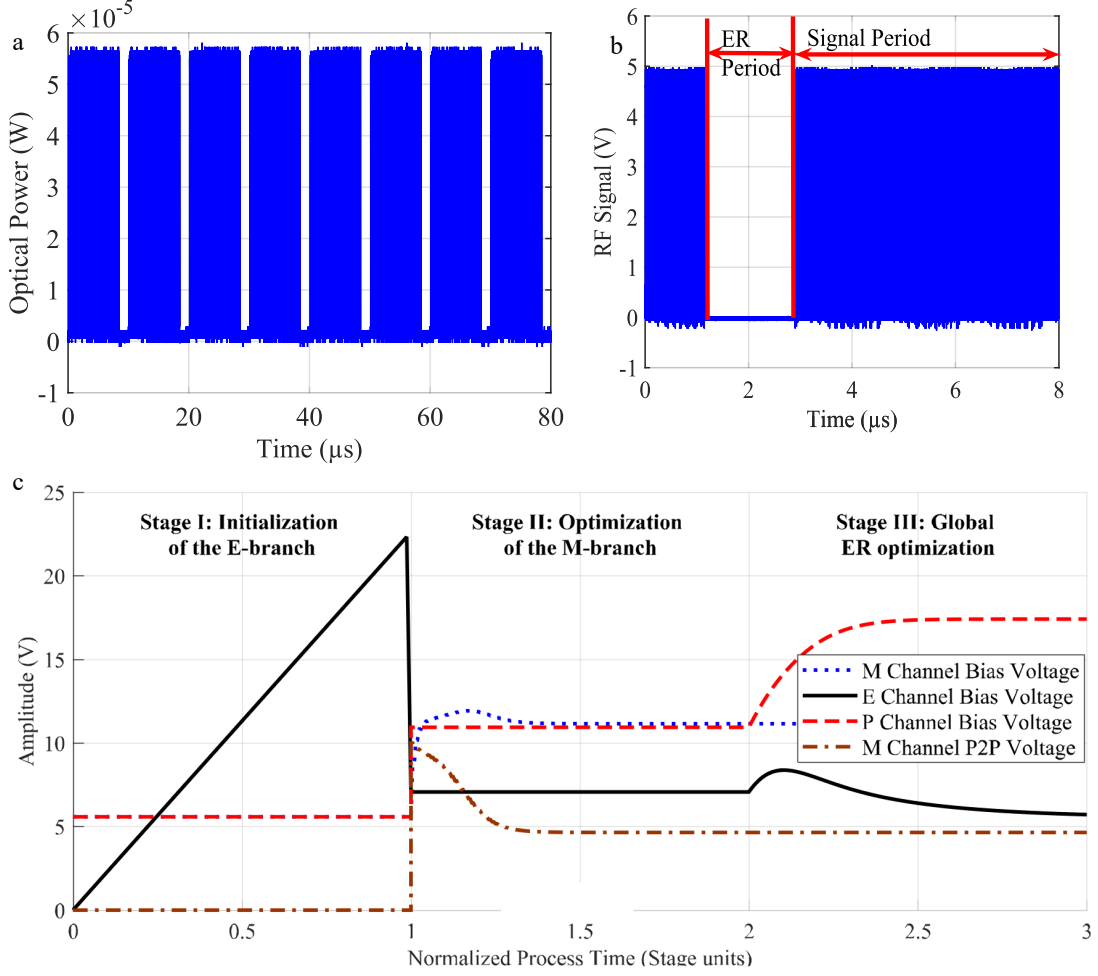


Fig. 1: Signal structure and adaptive bias behavior in the ER-Boost MZM system. a, Measured optical output waveform showing the alternation between normal signal periods and short ER-optimization periods. During the signal period, the MZM transmits high-speed modulated optical data; during the ER period, the optical path is held in a calibration state used for optical power monitoring and feedback adjustment. b, Corresponding RF driving waveform, where the FPGA inserts periodic ER-control frames prior to each data burst to enable real-time extinction optimization. c, Temporal evolution of the three bias voltages applied to the modulation (M), enhancement (E), and phase-control (P) channels. The bias controller dynamically tunes $V_{\text{bias,M}}$, $V_{\text{bias,E}}$, and $V_{\text{bias,P}}$ during the ER period to minimize residual optical leakage, while maintaining stable bias levels during normal transmission.

Comment 7 *Fig. 5c: Are feedback controls applied to lock the extinction ratio to these values? In earlier sections, an ER above 38 dB is obtained. What is the reason for using a maximum ER of 32.1 dB here?*

Response:

We thank the reviewer for this important clarification request.

Yes, feedback control is applied to stabilize the extinction ratio at the specified values during the communication experiments. The extinction ratios shown in Fig. 5(c) (15.3 dB, 24.5 dB, and 32.1 dB) were intentionally set and locked using the bias optimization procedure.

Although an extinction ratio exceeding 38 dB is achievable under optimized conditions, we selected 32.1 dB as the highest value in Fig. 5 to clearly illustrate the relationship between extinction ratio and BER. We experimentally observed that once the extinction ratio exceeds approximately 30 dB, further increases yield only marginal improvement in BER under the tested photon budget. Therefore, using 32.1 dB allows clear visualization of the transition from leakage-limited to noise-limited performance without redundancy in the BER curves.

Changes in the manuscript:

We have clarified this experimental choice in the revised manuscript.

It should be noted that the extinction ratios shown in Fig. 5(c) (15.3 dB, 24.5 dB, and 32.1 dB) were intentionally set and stabilized using the adaptive bias optimization procedure. Although extinction ratios exceeding 38 dB are achievable under optimized conditions, we selected 32.1 dB as the highest comparison value in this figure to clearly illustrate the relationship between extinction ratio and BER. Experimentally, we observed that once the extinction ratio exceeds approximately 30 dB, further increases result in only marginal BER improvement under the tested photon budget. Beyond this point, the link performance becomes dominated by noise rather than leakage. Therefore, using 32.1 dB allows a clear visualization of the transition from leakage-limited to noise-limited performance without redundancy in the BER curves.

Comment 8 *Supplementary, page 2: Please check whether “I-branch” and “Q-branch” are the proper terms here.*

Response:

We thank the reviewer for carefully pointing this out. We agree that the terms “I-branch” and “Q-branch” were not appropriate in this context and may cause confusion. We have corrected the terminology in the Supplementary Information to ensure consistency and accuracy with the actual device configuration described in the manuscript.

Changes in the Supplementary:

Here, $V_{\text{bias,I}}$, $V_{\text{bias,Q}}$, and $V_{\text{bias,P}}$ are the independently tunable DC bias voltages applied to the M-branch, E-branch, and recombination phase control path, respectively.

III. REVIEWER 3'S COMMENTS

We thank Reviewer 3 for the careful reading of our manuscript and for the many constructive suggestions on extinction-ratio stability, application scope, and experimental details. In the revised version, we have

- (i) clarified that the present prototype employs periodic ER re-optimization rather than continuous real-time tracking, and explicitly added this explanation around Fig. 4;
- (ii) narrowed the scope of our claims about ultra-high ER in classical systems and rephrased the introduction and abstract to focus on photon-sensitive and weak-signal applications;
- (iii) revised the discussion on the role of high ER in quantum communication and updated the corresponding citation;
- (iv) clarified the push–pull modulation scheme, the function of $V_{\text{bias,M}}$, and the definitions of power-splitting ratio γ and imbalance;
- (v) corrected the description and control logic of the three-stage bias-optimization procedure, updated Figs. 2 and 3, and explained why a saddle-shaped surface appears under low-duty-cycle PPM modulation;
- (vi) added a discussion of the dynamic range and maximum measurable ER of the SNSPD-based method;
- (vii) provided a summary table of the commercial Ti:LiNbO_3 modulator specifications (bandwidth, V_π , native ER, insertion loss) together with an explicit statement that phase control is realized via the Pockels effect.

We believe these revisions address the reviewer's concerns and significantly improve the clarity and completeness of the manuscript.

Comment 1 *The authors describe a novel and efficient implementation to improve extinction ratios of electro-optic modulators, limited by fabrication tolerances and drift. They practically demonstrate the feasibility and usefulness of the approach. They show in detail the measured absolute performance of the device. It would help the reader if they would provide the relative performance improvement by*

A) Stating the native ER of a typical MZM fabricated in the particular platform (absolute ER improvement), and

B) Comparing the ER drift over time (Fig. 4) with the real-time tracking on/vs off (ER stability improvement).

Response:

We thank the reviewer for this constructive suggestion regarding ER stability comparison.

We clarify that the current hardware architecture does not implement continuous real-time tracking of all bias paths. Instead, the system performs periodic re-optimization of the extinction ratio. This design choice is primarily due to the use of a single monitoring photodiode at the output port, which makes continuous multi-parameter tracking impractical in the present implementation.

Figure 4 illustrates the ER evolution over a complete optimization cycle under ambient laboratory conditions. After each optimization step, the extinction ratio reaches its maximum and then gradually decreases due to slow bias drift. In practice, the drift rate of the Ti:LiNbO_3 modulator is sufficiently slow that periodic recalibration is adequate to maintain high extinction performance for typical communication scenarios.

We agree that a systematic comparison between continuous tracking, periodic recalibration, and open-loop operation under controlled perturbations would provide additional insight. Such a study would require branch-level monitoring photodiodes and enhanced control hardware and is part of our planned future work. The revised manuscript now clarifies that the demonstrated stability corresponds to periodic recalibration rather than continuous real-time tracking.

Changes in the manuscript:

To clarify the stability mechanism of the current implementation, we have added the following statement in the Results section (in the discussion of Fig. 4):

It should be noted that the present implementation performs periodic ER re-optimization rather than continuous real-time tracking, which is sufficient given the relatively slow drift dynamics of the Ti:LiNbO_3 modulator under ambient conditions.

Comment 2 *In the abstract, the authors claim that ultra-high ERs are crucial to energy efficiency in classical optical communication systems. This claim is not further supported in the main text and in my opinion incorrect. I recommend to either remove the part “... both classical and” or to clarify under which particular circumstances ultra-high ERs are benefiting energy efficiency in classical communication systems.*

Response:

We thank the reviewer for this important comment regarding the scope of the claim in the abstract.

We agree that the statement was overly broad. Ultra-high extinction ratios are not universally required for all classical optical communication systems, particularly in high-power coherent links where other parameters such as electro-optic efficiency and bandwidth dominate the system-level energy consumption.

Our intention was to emphasize application scenarios where leakage suppression directly impacts system sensitivity and power budget, such as weak-signal, photon-counting, and ultra-low-power optical links. To avoid overgeneralization, we have revised the abstract to clarify the applicable context and remove the potentially misleading phrasing.

Changes in the manuscript:

The revised sentence is reproduced below:

Enhancing ER is therefore essential for improving signal contrast and dynamic range in photon-sensitive and weak-signal optical systems [2]–[4], and for enabling emerging applications such as quantum key distribution, optical neural networks, and ultra-low-power photonic interconnects [5], [6].

Comment 3 *In the introduction, the authors cite ref 18 and claim that they show that “...high ER is essential to suppress optical leakage and preserve quantum-state fidelity.” I see in ref 18 an implementation of how to suppress optical leakage, but not an explanation why it is essential. I kindly ask the authors to either provide a reference that shows the essentiality, or show where ref 18 demonstrates essentiality, or rephrase the statement.*

Response:

We thank the reviewer for this careful observation. We agree that the previous wording was too strong and that Ref. 18 does not explicitly state that high extinction ratio is “essential.” To avoid overstatement and to improve citation accuracy, we have revised the sentence accordingly.

In the revised manuscript, we clarify that a high extinction ratio plays a critical role in suppressing optical leakage and reducing excess background noise, which can degrade quantum-state fidelity and security margins. We also cite a comprehensive review on practical quantum communication systems to support this statement.

Changes in the manuscript:

The revised sentence is reproduced below:

High extinction ratio plays a critical role in suppressing optical leakage and reducing excess background noise, both of which can degrade quantum-state fidelity and security margins in practical quantum communication systems [7].

Comment 4 *In the introduction, the authors implicitly claim to overcome the limitations of thermal sensitivity. At the end of the section they claim stability under “environmental perturbations”. They don’t show nor discuss thermal stability of their solution. If practical, it would be beneficial to the manuscript to show a time evolution of ER like in Fig. 4, but show with vs. without real-time tracking, and turn on a thermal perturbation in certain time intervals.*

Response:

We thank the reviewer for this insightful comment.

We clarify that the present manuscript does not include a controlled thermal perturbation experiment. Figure 4 illustrates the evolution of the extinction ratio during a complete optimization cycle under ambient laboratory conditions. In the current prototype, only a single monitoring photodiode is available at the output port. As a result, the multi-parameter bias optimization is not continuously tracked in real time for all paths, and the extinction ratio gradually degrades after reaching its optimal value within each control cycle.

We acknowledge that the previous wording referring to “environmental perturbations” may have been overly broad. In the revised manuscript, we have adjusted the language to more accurately reflect that the demonstrated stability corresponds to controlled bias optimization under ambient conditions rather than to a dedicated thermal robustness study.

We agree that a systematic comparison of tracking-on versus tracking-off operation under controlled thermal disturbances would provide valuable insight. Such a study requires a dedicated device architecture with branch-level monitoring photodiodes and is planned for future work. The primary objective of the present work is to demonstrate the feasibility of the ER-Boost architecture and its closed-loop optimization concept.

To avoid overstatement regarding environmental robustness, we have revised the corresponding sentence in the Discussion section to more accurately reflect the demonstrated experimental conditions.

Changes in the manuscript:

The updated text is reproduced below:

The combination of this optical architecture with a high-speed adaptive bias control system enables sub-millisecond bias adjustment, contributing to improved operational stability under ambient laboratory conditions without compromising modulation bandwidth.

Comment 5 *In the results section, there is a single closing bracket in line two. Presumably a typo.*

Response:

We thank the reviewer for pointing out this typographical error. The extra closing bracket in the Results section has been removed in the revised manuscript.

Comment 6 *In the “system architecture” section, the modulator is explained. It would help the reader to explain, that push-pull modulation is used and that the data signal is unipolar and differential, i.e. balanced.*

Response:

We thank the reviewer for this helpful suggestion. We clarify that in the present implementation, the electrical data signal generated by the FPGA is unipolar (0/1) and is amplified before being applied to one arm of the M-branch, while the other arm is referenced to ground. Therefore, the modulation operates under a single-ended drive configuration rather than a fully differential push–pull scheme.

We have revised the System Architecture section to explicitly describe the drive configuration and the role of the bias voltage $V_{\text{bias,M}}$, which establishes the static π phase shift between the two arms and defines the operating point for intensity modulation.

Changes in the manuscript:

The following paragraph has been added to the manuscript:

The modulation scheme employed in this work is intensity modulation. The electrical data signal generated by the FPGA is unipolar (0/1) and is amplified before being applied to

one arm of the M-branch. The other arm is referenced to ground. Therefore, the M-branch operates under a single-ended drive configuration. The DC bias voltage $V_{\text{bias,M}}$ is used to establish the static π phase shift between the two arms of the M-branch, thereby defining the operating point for intensity modulation.

Comment 7 In Fig. 1(a), the configuration of the modulator electrodes is illustrated; however, several aspects require further clarification and improvement: (a) The V_{RF} signal appears to be a differential signal, likely at least after the amplifier stage. This is not clearly indicated in the figure and should be explicitly clarified.

(b) The V_{RF} signal appears to drive the upper arm with approximately twice the strength of the lower arm. It is unclear whether this asymmetry is intentional or a schematic inconsistency.

(c) For each bias electrode, one arm appears to be grounded. It is not clear whether the bias signals are applied in a differential manner. Additional clarification is needed.

(d) The V_{RF} signal is the only electrical signal that is visually connected to the corresponding modulator, which may lead to ambiguity regarding the connectivity of other electrical control signals.

(e) The label “Electrical Signal” may be misleading, as multiple electrical signals are present in the system. The label “Data Signal” may provide clearer distinction.

(f) The label “QPSK Drive Link” appears inconsistent with the presented intensity modulation scheme and should be reconsidered or further justified.

Response:

We thank the reviewer for the detailed and constructive comments regarding Fig. 1(a). We address each point below.

(a) The M-branch employs push–pull modulation. The RF data signal is generated as a single-ended electrical signal from the FPGA and is subsequently converted into differential signals by the RF amplifier stage before being applied to the two arms of the M-branch. In the revised figure, the differential nature of the RF drive after amplification is now explicitly indicated.

(b) The asymmetric RF drive amplitude between the two arms is intentional. The main modulation arm is driven with approximately twice the modulation depth of the auxiliary arm. This design reduces the effective half-wave voltage and relaxes device design constraints while

also contributing to bandwidth optimization. We have clarified this intentional asymmetry in the revised manuscript to avoid confusion.

(c) The DC bias voltages are applied in a single-ended configuration, where only one electrode of each phase shifter is actively tuned. The phase shift is therefore determined by the applied voltage relative to ground through the Pockels effect in the Ti:LiNbO₃ waveguides. This has now been explicitly stated in the text.

(d) In the original schematic, only the RF path was visually emphasized to maintain readability. We acknowledge that this may have caused ambiguity. The revised figure now clearly shows the connectivity of all electrical control signals while preserving diagram clarity.

(e) The label “Electrical Signal” has been replaced with “Electrical Data Signal (FPGA)” to more accurately reflect the signal origin and function.

(f) The label “QPSK Drive Link” was an unintentional carryover from the commercial modulator documentation and is not consistent with the intensity modulation scheme used in this work. It has been corrected to “Data Drive Link” in the revised figure.

We have updated Fig. 1(a) and the corresponding caption to reflect these clarifications.

Changes in the manuscript:

To clarify the modulation scheme and electrode configuration in response to the reviewer’s comments, we have added the following description in the System Architecture section:

The RF data signal is generated as a single-ended signal from the FPGA and converted into differential signals by the RF amplifier, enabling push–pull modulation of the M-branch. The asymmetric RF drive amplitude between the two arms is intentional and is used to reduce the effective half-wave voltage and optimize bandwidth performance. All DC bias voltages are applied in a single-ended configuration, where one electrode is actively tuned relative to ground.

Comment 8 *The caption of Fig.1, subsection 2.2 and the supplement make use of the “power splitting ratios”, labelled with the lowercase Greek letter gamma. It is not clear to me, what this should designate. The authors should define the figure. Similarly, the term “imbalance factor” is used without definition.*

Response:

We thank the reviewer for pointing out the lack of clarity in the definition of the power splitting ratio and related terminology.

To address this issue, we have explicitly defined the optical power splitting ratio γ in the revised manuscript. Specifically, γ is now clearly described as the fraction of optical power directed into the upper arm relative to the total optical power distributed between the upper and lower arms of the corresponding Y-branch. This clarification has been added in Subsection 2.2 and is consistently applied throughout the manuscript.

Changes in the manuscript:

The following text has been added to Subsection 2.2:

Fig. 2d illustrates the simulated maximum ER as a function of the power splitting ratios of the modulation branch (γ_M) and the extinction-enhancement branch (γ_E), where γ denotes the optical power splitting ratio of the Y-branch, defined as the fraction of optical power directed into the upper arm relative to the total optical power distributed between the upper and lower arms.

The yellow regions correspond to ER values exceeding 40 dB, which occur when the power splitting ratios of the two branches are approximately matched. This result indicates that the proposed ER-Boost architecture can effectively compensate for structural imperfections between the two branches. In particular, the requirement for perfectly balanced 50:50 splitters is no longer critical, since residual leakage from the modulation branch can be cancelled by a corresponding imbalance in the auxiliary branch.

Comment 9 *Subsection 2.2 explains the operation point of the MZM. It should be clarified that intensity modulation is used in combination with unipolar differential drive signals, and that $V_{bias,M}$ is used to tune the static π phase shift between the arms of the M-branch.*

Response:

We thank the reviewer for this helpful suggestion. We agree that the modulation scheme and the role of the M-branch bias voltage should be described more explicitly.

In the revised manuscript, we clarify that the system employs intensity modulation with push-pull differential RF drive signals. We also explicitly state that the bias voltage $V_{bias,M}$

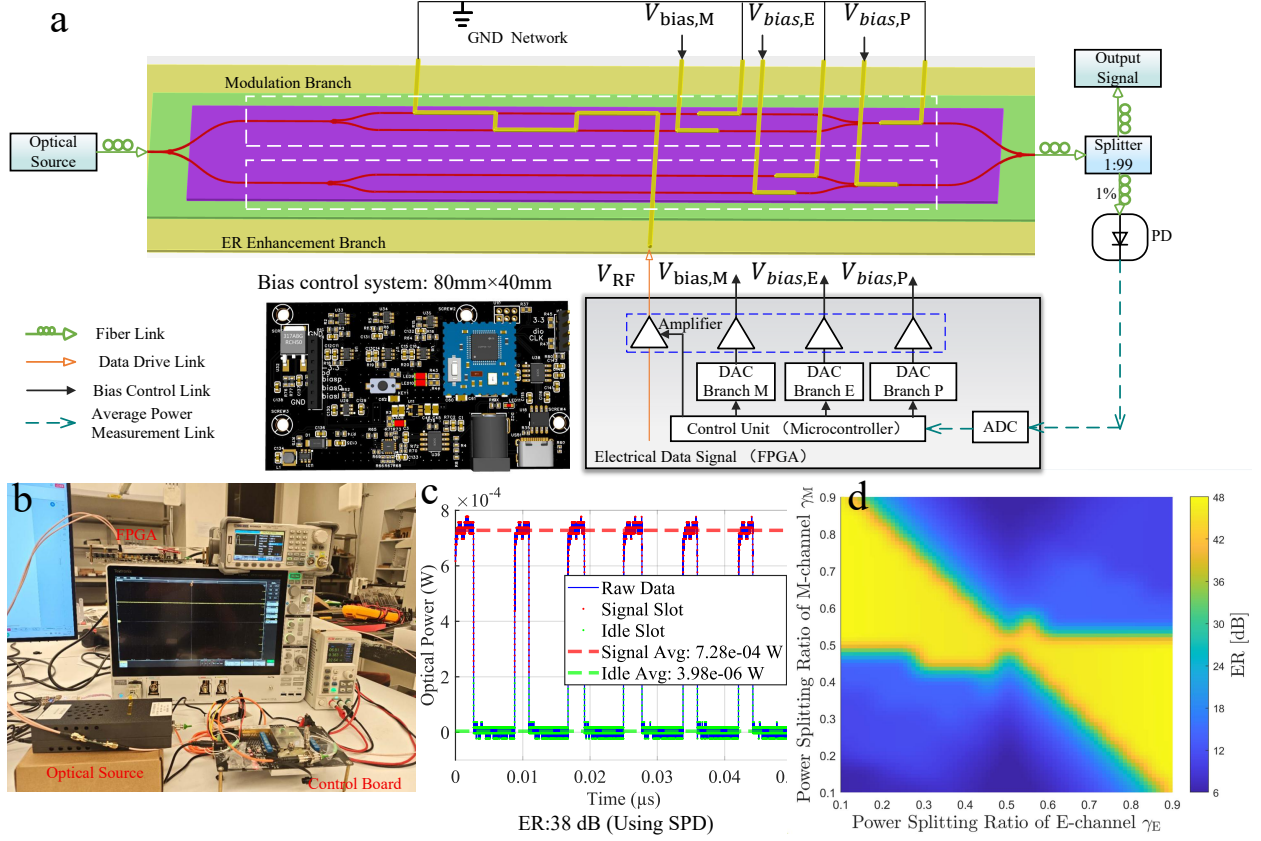


Fig. 2: Architecture and experimental validation of the ER-Boost MZM system. a, Schematic of the ER-Boost MZM structure and bias control system. The dual-branch architecture comprises a modulation branch and an ER enhancement branch, with independently tunable DC bias voltages $V_{bias,M}$, $V_{bias,E}$, and $V_{bias,P}$ applied through a three-channel DAC-based control board (80 mm $\times$ 40 mm). b, Photograph of the experimental setup including the FPGA-based signal generation board and optical measurement instruments. c, Measured optical waveform demonstrating an ER of 38 dB using the Superconducting Nanowire Single-Photon Detector (SNSPD). d, Simulated ER map as a function of power-splitting ratios γ_M and γ_E , showing the optimal bias configuration region for maximum ER performance.

is used to establish the static π phase difference between the two arms of the M-branch, thereby defining the operating point for intensity modulation.

Changes in the manuscript:

The following text has been added to Subsection 2.1:

The modulation scheme employed is intensity modulation. The M-branch is driven in a push–pull configuration using differential RF drive signals, while the data signal itself is unipolar. The bias voltage $V_{\text{bias},M}$ is used to establish the static π phase difference between the two arms of the M-branch, ensuring proper operating point alignment for intensity modulation and enabling precise control of the destructive and constructive interference conditions.

Comment 10 *In subsection 2.3, the optimization of the M-branch is described. The description refers to a saddle-shape in Fig.2c, which cannot be identified. I am sceptical of the plot in Fig.2c, as I would expect that at $V_{pp} = V_\pi$ the optical average power is independent of $V_{\text{bias},M}$. This can be derived from eq. 10 in the supplement by averaging $P_{\text{out}}(V_{in} = 0V)$ and $P_{\text{out}}(V_{in} = V_{pi})$, where the result is constant and independent of $V_{\text{bias},M}$. The authors should check the figure and indicate the saddle-point. They should also check the caption; the descriptions of b and c are incomprehensible, probably due to typos.*

Response:

We thank the reviewer for the careful analytical consideration based on Eq. 10 in the Supplement.

We agree that, for an ideal push–pull MZM driven symmetrically with $V_{pp} = V_\pi$ and a 50% duty cycle, the time-averaged optical power obtained by averaging $P_{\text{out}}(V_{in} = 0)$ and $P_{\text{out}}(V_{in} = V_\pi)$ would indeed be independent of $V_{\text{bias},M}$.

However, this derivation implicitly assumes a 50% duty cycle. In our experimental scenario, the modulation format is pulse-position modulation (PPM), where the duty cycle is significantly smaller than 1/2. Under such conditions, the average optical power is given by

$$P_{\text{avg}} = D \cdot P(V_{pp}) + (1 - D) \cdot P(0),$$

where D is the duty cycle and $D < 1/2$. In this case, the time-averaged power retains dependence on $V_{\text{bias},M}$, and the ideal cancellation of bias dependence no longer holds.

Therefore, the observed saddle-shaped surface in Fig. 2c arises naturally from the low-duty-cycle modulation employed in photon-counting and quantum communication scenarios. In addition, practical non-idealities such as RF amplitude imbalance and finite electrical bandwidth further contribute to the residual bias dependence.

Changes in the manuscript:

To clarify this point, we have revised the manuscript in Fig.3 to explicitly mention the influence of duty cycle on the average power dependence.

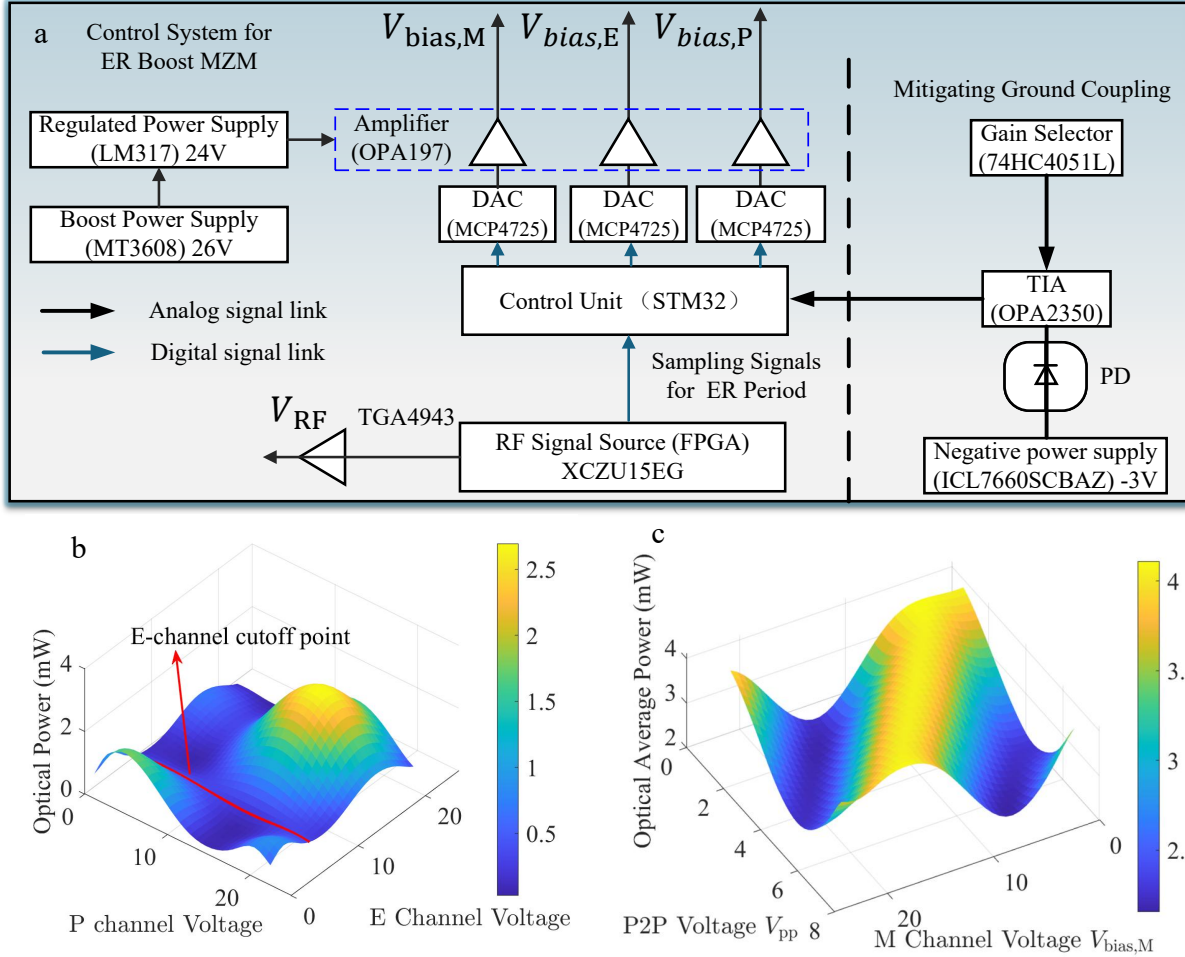


Fig. 3: Control architecture and performance analysis of the ER-Boost MZM bias optimization system. a, Block diagram of the custom-designed control board, incorporating STM32 control unit, DACs, amplifiers, and transimpedance amplifier for PD readout. The system also integrates a ground coupling isolation stage and gain selector. b, Measured output optical power versus $V_{bias,P}$ and $V_{bias,E}$, showing a sharp extinction boundary indicating with E channel. c, the E-channel cutoff point. Measured optical power response with respect to $V_{bias,M}$ and RF peak-to-peak voltage V_{pp} . It should be noted that the saddle-shaped surface arises because the employed PPM modulation format has a duty cycle significantly smaller than 50%, causing the time-averaged optical power to retain dependence on $V_{bias,M}$ even when V_{pp} is nominally set to V_{π} .

Comment 11 *Fig. 3 and the caption of Fig. 3 show dynamic tuning of $V_{\text{bias,M}}$ in the ER, while the list in Subsection 2.3 mentions explicitly that only $V_{\text{bias,E}}$ and $V_{\text{bias,P}}$ are adjusted in the ER period. Then, in the caption of Fig. 4 it is explained that the M-path is not tracked in real-time. The authors should resolve the inconsistencies. Ideally, Fig. 4 should compare the performance under recalibration of all paths to the recalibration without M-path and the performance without closed-loop control and apply a comparable thermal disturbance after a set amount of time.*

Response:

We sincerely thank the reviewer for carefully identifying the inconsistency.

First, we clarify that an incorrect intermediate version of Fig. 3 was inadvertently uploaded during submission. In that version, the distinction between the initialization phase and the ER optimization phase was not clearly represented, which caused ambiguity regarding whether $V_{\text{bias,M}}$ was adjusted during the ER period.

We have now corrected Fig. 3 to clearly illustrate the three-stage control procedure:

(i) Stage I: Initialization of the E-branch; (ii) Stage II: Optimization of the M-branch; (iii) Stage III: Global ER optimization.

In the corrected figure, $V_{\text{bias,M}}$ is tuned only during Stage II (initial optimization). During Stage III (ER optimization), only $V_{\text{bias,E}}$ and $V_{\text{bias,P}}$ are adjusted, while the M-path bias is held constant. The x-axis has also been clarified as normalized process time to avoid confusion with physical time units.

To eliminate any ambiguity in the manuscript, we have revised Subsection 2.3 to explicitly describe the stage-wise control logic, as reproduced below.

Changes in the manuscript:

The following revised descriptions have been added in Subsection 2.3:

Initialization of the E-branch. The modulation branch bias voltage $V_{\text{bias,M}}$ is held constant, while the extinction branch voltage $V_{\text{bias,E}}$ and the phase-control branch voltage $V_{\text{bias,P}}$ are jointly scanned to map the optical power surface. As shown in Fig. 3b, the measured optical power exhibits a saddle-like distribution as a function of the two control voltages. The optimal $V_{\text{bias,E}}$, referred to as the *E-channel cutoff point*, is located along the valley line where the variation of $V_{\text{bias,P}}$ induces minimal change in the output optical power. This point

corresponds to the OFF state of the E-branch, at which the leakage field from the modulation branch is most effectively canceled by destructive interference.

In practical operation, a small perturbation signal is superimposed onto the P-branch bias voltage, and the resulting variation in output optical power is monitored while scanning $V_{\text{bias,P}}$. The E-channel cutoff point is identified as the operating condition where the output power exhibits minimal sensitivity to this perturbation, indicating effective cancellation of the leakage field.

Optimization of the M-branch. With the E-branch fixed at its OFF state, the modulation branch bias voltage $V_{\text{bias,M}}$ and the RF peak-to-peak driving voltage V_{pp} are jointly tuned to determine the optimal modulation condition. As shown in Fig. 3c, the measured optical average power exhibits a saddle-shaped surface as a function of $V_{\text{bias,M}}$ and V_{pp} .

In practical operation, the saddle point is located through an iterative search procedure. For a given $V_{\text{bias,M}}$, the RF peak-to-peak voltage V_{pp} is scanned to identify the value that maximizes the modulation depth. Subsequently, for the optimized V_{pp} , the bias voltage $V_{\text{bias,M}}$ is scanned to minimize the average optical output power. These two steps are alternated iteratively until convergence is achieved. The resulting operating point corresponds to the saddle region, where the optical output power becomes least sensitive to small variations in either bias or RF amplitude.

Finally, regarding the suggestion to compare continuous tracking, periodic recalibration, and open-loop operation under controlled thermal disturbance, we acknowledge that such a systematic comparison would provide additional insight. However, the current hardware architecture employs periodic ER re-optimization rather than continuous real-time tracking, due to the single monitoring photodiode configuration. Given the relatively slow drift dynamics of the Ti:LiNbO_3 modulator under ambient conditions, periodic recalibration is sufficient in the demonstrated scenario. A comprehensive comparison under controlled perturbations will be addressed in future work.

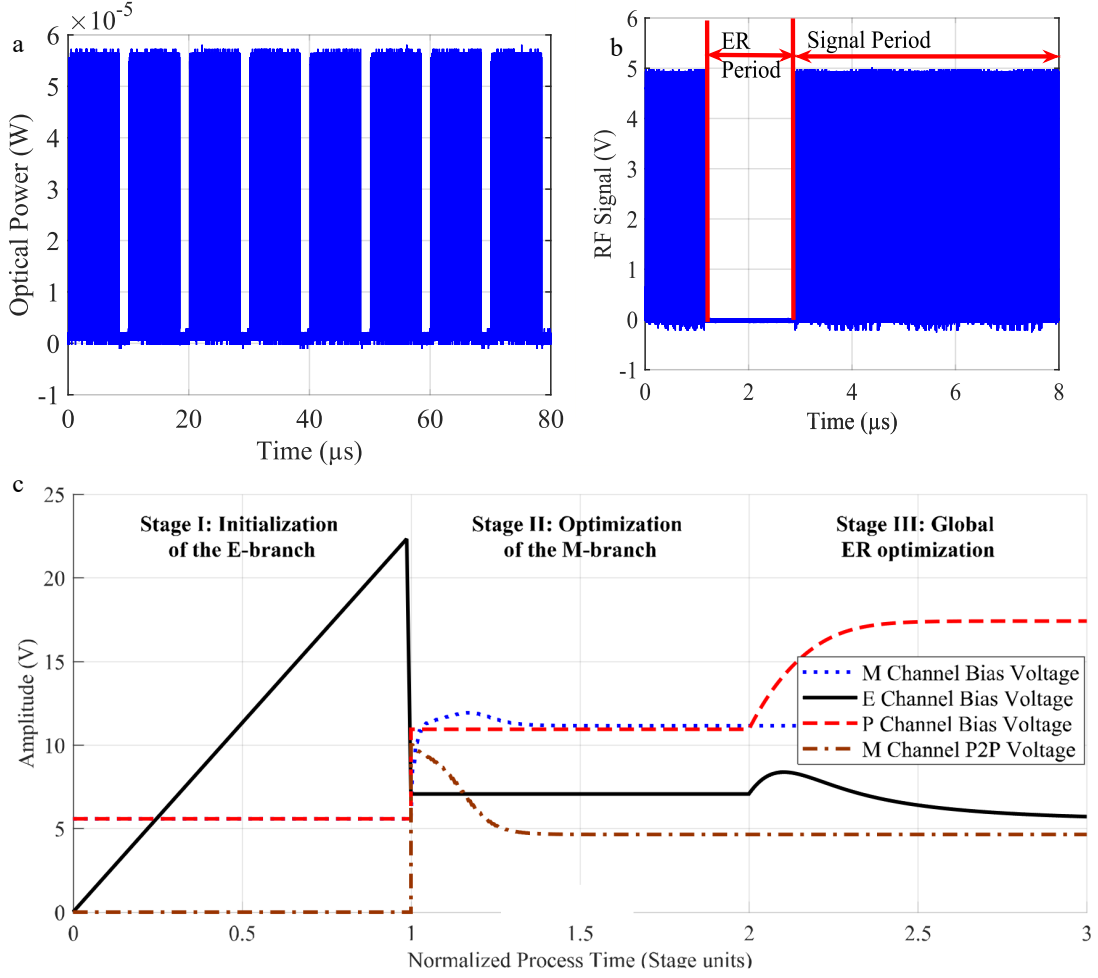


Fig. 4: Signal structure and adaptive bias behavior in the ER-Boost MZM system. a, Measured optical output waveform showing the alternation between normal signal periods and short ER-optimization periods. During the signal period, the MZM transmits high-speed modulated optical data; during the ER period, the optical path is held in a calibration state used for optical power monitoring and feedback adjustment. b, Corresponding RF driving waveform, where the FPGA inserts periodic ER-control frames prior to each data burst to enable real-time extinction optimization. c, Temporal evolution of the three bias voltages applied to the modulation (M), enhancement (E), and phase-control (P) channels. The bias controller dynamically tunes $V_{\text{bias,M}}$, $V_{\text{bias,E}}$, and $V_{\text{bias,P}}$ during the ER period to minimize residual optical leakage, while maintaining stable bias levels during normal transmission.

Comment 12 *Subsection 2.4 should mention the dynamic range of the SNSPD-based method, the maximum detectable ER, or a similar figure to compare the measured ER with the sensitivity of the setup.*

Response:

We thank the reviewer for raising this important point regarding the measurement sensitivity.

First, we emphasize that this work evaluates the *dynamic* extinction ratio of high-speed PPM signals rather than the static on/off extinction ratio. Conventional optical power meters can provide large static dynamic ranges; however, they measure only average optical power and cannot resolve slot-level power differences under high-speed modulation. Similarly, traditional high-speed photodetectors typically provide an effective dynamic range of only 15–20 dB under pulsed conditions.

For this reason, we adopted the SNSPD-based dynamic slot-power measurement technique proposed in [1], which enables time-resolved photon-counting reconstruction of signal and guard slot powers. As demonstrated in that work, the method supports dynamic extinction ratio measurements exceeding 70 dB, limited primarily by the detector dark count rate and integration time.

In the present experiment, the measured extinction ratio of 38 dB is therefore not limited by the intrinsic dynamic range of the SNSPD-based method. Rather, it is constrained by the current system implementation. Specifically:

- (1) Only a single monitoring photodiode is available at the output port, making simultaneous optimization of multiple bias voltages a nontrivial multi-parameter problem.
- (2) Electronic noise in the monitoring PD and associated circuitry affects the bias optimization precision.
- (3) The insertion of dedicated ER-Boost periods within the signal frame introduces temporal segmentation, which slightly reduces control accuracy under dynamic operation.

In future implementations, we plan to fabricate a dedicated ER-Boost MZM integrating branch-level monitoring photodiodes for the M- and E-branches. This architecture will significantly simplify the control algorithm and further improve the achievable dynamic extinction ratio.

To address the reviewer's concern regarding the measurement sensitivity and dynamic range of the SNSPD-based method, we have added an explicit clarification in Subsection 2.4.

Changes in the manuscript:

The added text is reproduced below:

To overcome this limitation, our system leverages a high-sensitivity SNSPD to directly measure the photon-counting statistics within each time slot [1]. Unlike static power meters, this method resolves time-synchronized signal and guard slots under high-speed modulation. The achievable measurement dynamic range is primarily determined by the detector dark count rate and integration time, and has been experimentally demonstrated to exceed 70 dB in dynamic PPM measurements. Therefore, the measured 38 dB extinction ratio in the present experiment is not limited by the intrinsic sensitivity of the SNSPD method, but rather by the current device configuration and bias control architecture.

Comment 13 *In Figure 5c. and d. the colour scheme for the ER and the order in the legend is inconsistent. This would help to identify the correct measurements when comparing.*

Response:

We thank the reviewer for pointing out this inconsistency. We have revised Figures 5c and 5d to ensure that the colour scheme corresponding to each extinction ratio is consistent across both subfigures. In addition, the order of the legend entries has been unified to follow the same ascending ER sequence for clarity.

Changes in the manuscript:

We have revised Figures 5c and 5d to ensure that the colour scheme corresponding to each extinction ratio is consistent across both subfigures.

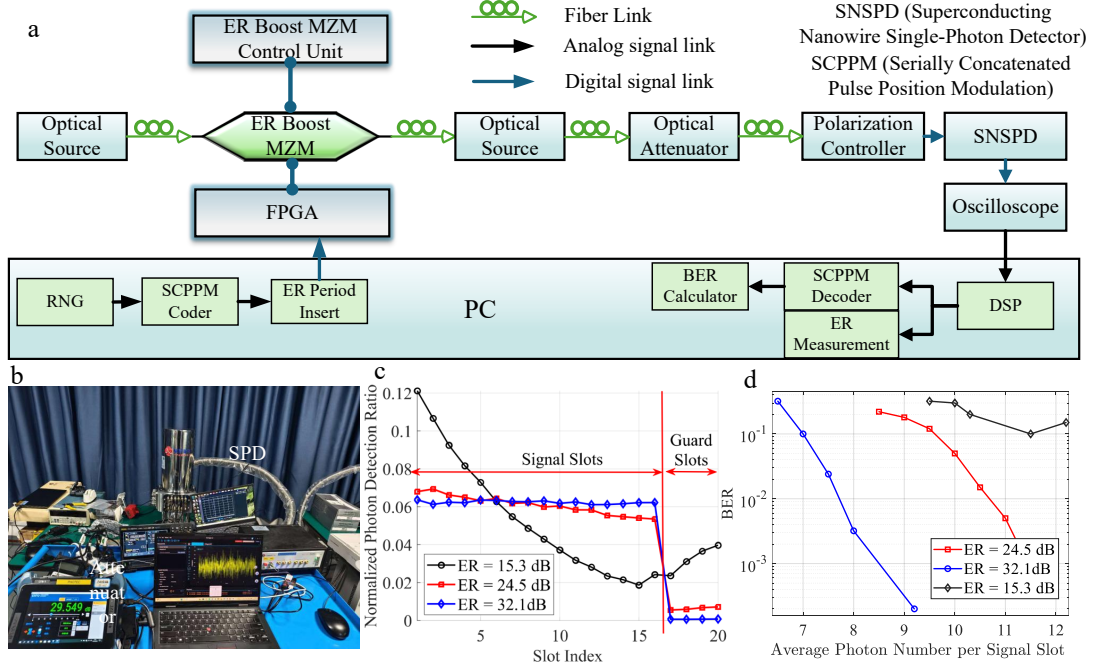


Fig. 5: Experimental setup and performance evaluation of the proposed ER-Boost MZM system. (a) Block diagram of the optical transmission experiment integrating the ER-Boost MZM with an FPGA-based SCPPM encoder/decoder and real-time bias control unit. (b) Photograph of the experimental setup including optical source, attenuator, SNSPD, and data acquisition system. (c) Normalized photon detection ratio across signal and guard slots under different ERs with 16-PPM (15.3 dB, 24.5 dB, and 32.1 dB). (d) BER performance as a function of the average photon number per signal slot, demonstrating the clear advantage of higher ERs.

Comment 14 Section 3 should mention the relevant capabilities of the particular LNOI technology, i.e., the analog bandwidth of the MZM, the native ER of a typical MZM, the V_π .

Response:

We thank the reviewer for this helpful suggestion. We would like to clarify that the experimental validation in this work was carried out using a commercially available Ti:LiNbO₃ dual-parallel Mach–Zehnder modulator (Fujitsu FTM7961EX), rather than a newly fabricated LNOI device.

To improve clarity and completeness, we have added a dedicated table in Section 3 summarizing the key performance parameters of the experimental modulator platform, including the analog bandwidth, intrinsic extinction ratio, drive voltage (V_π), and insertion loss, based

on the manufacturer's datasheet. These parameters provide a clear baseline for evaluating the extinction enhancement achieved by the proposed ER-Boost architecture.

Changes in the manuscript:

To provide a clear performance baseline of the experimental platform, Table I summarizes the key specifications of the Ti:LiNbO₃ dual-parallel MZM used in this study, extracted from the manufacturer datasheet. The intrinsic extinction ratio of the conventional single-branch operation is approximately 20–22 dB, which serves as a reference for evaluating the extinction enhancement achieved through the proposed dual-branch interference mechanism.

TABLE I: Key specifications of the experimental Ti:LiNbO₃ dual-parallel MZM (Fujitsu FTM7961EX), based on manufacturer datasheet.

Parameter	Value
Platform	Ti:LiNbO ₃
Analog optical bandwidth	22 GHz
V_π (low frequency)	2.8 V
V_π (22 Gbps)	3.5 V
Native extinction ratio (parent MZI)	22 dB
Native extinction ratio (child MZI)	20 dB
Optical insertion loss (C-band, no connector)	7.0 dB
RF impedance	50 Ω

Comment 15 *Section 2 or 3 should mention whether the bias signals use the Pockels effect or whether thermo-optic phase shifters are used (or any other technology).*

Response:

We thank the reviewer for this helpful suggestion. The experimental modulator used in this work is based on a Ti:LiNbO₃ platform, where phase modulation and bias tuning are realized through the linear electro-optic (Pockels) effect. No thermo-optic phase shifters are employed in the device.

To clarify this point, we have revised the manuscript and explicitly described the experimental modulator platform and its operating principle.

Changes in the manuscript:

We have added the following text in the subsection *Experimental Modulator Platform and Optical Setup*:

The experimental validation of the proposed ER-Boost architecture was performed using a commercially available dual-parallel Mach–Zehnder modulator (Fujitsu FTM7961EX), fabricated on a Ti:LiNbO₃ platform. In this device, both RF modulation and DC bias tuning are realized via the linear electro-optic (Pockels) effect in the LiNbO₃ waveguides. No thermo-optic phase shifters are employed. The device integrates two child MZIs and a parent MZI, enabling flexible reconfiguration into the dual-branch interference-enhanced structure proposed in this work. Independent DC bias electrodes are provided for each interferometric section, allowing fine phase control and implementation of the ER-Boost bias optimization strategy.

The optical source was a tunable continuous-wave laser, polarization-maintained and coupled into the modulator via fiber arrays. The optical output was monitored simultaneously by a photodiode (PD) for high-power measurements and a SNSPD for ultra-low-power photon-counting experiments. A 1:99 optical splitter provided feedback signals for bias control and real-time extinction ratio monitoring.

Comment 16 *In the supplement, section 3 mentions the dropping of the subscript n and then the same two symbols are copied twice. -Probably a typo.*

Response:

We thank the reviewer for carefully pointing out this typographical error. The notation inconsistency in Section 3 of the Supplementary Information has been corrected in the revised version to ensure clarity and consistency.

IV. REVIEWER 4'S COMMENTS

We sincerely thank Reviewer 4 for the careful reading of our manuscript and for the many insightful and technically detailed comments. These suggestions have led to substantial improvements in both the presentation and contextual positioning of our work. In the revised version, we have:

(i) expanded the Introduction and Discussion to better situate our results with respect to state-of-the-art high-extinction-ratio modulators, including cascaded MZI and tunable beam-splitter architectures;

(ii) clarified the distinction between static and dynamic extinction ratios and added quantitative measurements over a range of modulation frequencies, including GHz-class operation;

(iii) reported the baseline specifications of the commercial Ti:LiNbO₃ dual-parallel MZM used in our experiments and explicitly quantified the extinction enhancement enabled by the ER-Boost configuration;

(iv) discussed the main implementation-related limiting factors (monitoring PD configuration, RF imperfections, and electronic noise) and outlined concrete pathways for further improving dynamic extinction performance;

(v) clarified the role of the SNSPD as a measurement tool, the targeted application scenarios, and the applicability of the proposed bias control strategy with standard photodiode-based receivers;

(vi) added missing experimental details on the communication setup, including slot rate, modulation format, coding rate, and detector dark count rate.

We believe that these revisions have significantly strengthened the manuscript and have made the contribution and practical relevance of the proposed interference-engineered ER-Boost architecture much clearer. We are grateful for the reviewer's constructive feedback and thoughtful suggestions.

Comment 1 *Ultra-high ER modulation has been explored on various platforms. A more traditional approach uses two tunable beam-splitters instead of fixed Y-branches for the power splitting into the MZM. With such an architecture, Jin et al. (Opt. Lett. 44, 1265 (2019)) have achieved 50 dB extinction with a broad optical bandwidth of 100 nm on the thin film lithium niobate (TFLN) platform. Similarly, cascaded MZI devices on silicon have achieved up to 60 dB ER (Wilkes et al. Opt. Lett. 41, 5318 (2016)). Although on silicon, the thermo-optic effect is*

used, the overall principle should be translatable to lithium niobate (LN). The authors should compare their approach to these cascaded architectures in terms of performance, fabrication tolerance, etc.

Response:

We thank the reviewer for highlighting relevant high-extinction-ratio architectures reported in the literature.

We have carefully incorporated the cited works [8], [9] into the revised manuscript and added a dedicated comparison in the Discussion section. We acknowledge that cascaded MZI architectures and tunable beam-splitter designs have demonstrated impressive static extinction ratios exceeding 50–60 dB under steady-state bias conditions.

We emphasize, however, that the primary focus of the present work is the dynamic extinction ratio under high-speed modulation. The ER-Boost architecture introduces real-time bias optimization to continuously suppress residual leakage during signal operation. This distinguishes our approach from purely structural static interference optimization.

Rather than competing in terms of maximum static ER alone, the proposed scheme aims to enhance fabrication tolerance and dynamic robustness under practical operating conditions. In principle, the ER-Boost mechanism can also be combined with cascaded interferometric structures, indicating that the achievable static extinction limit is not fundamentally constrained by the dual-branch architecture.

In response to the reviewer’s suggestion, we have incorporated the cited high-extinction-ratio cascaded architectures [8], [9] into the Introduction and added an explicit comparison with the proposed ER-Boost approach.

Changes in the manuscript:

The revised text is reproduced below:

Recent developments in silicon photonics, thin-film lithium niobate (TFLN), and hybrid material platforms have enabled on-chip integration and ultra-high static extinction ratios exceeding 40–60 dB [8]–[11]. For example, high-extinction modulation has been demonstrated using tunable beam-splitter architectures on TFLN [8] and auto-configured cascaded Mach–Zehnder interferometers on silicon platforms [9], achieving extinction ratios up to 50–60 dB under steady-state bias conditions. Nevertheless, many of these approaches

rely on multi-stage interference or complex structural optimization, which can increase fabrication complexity and sensitivity to device imperfections. Moreover, a significant portion of ultra-high ER demonstrations primarily report static or quasi-static extinction performance, while the dynamic extinction behavior under high-speed modulation and long-term operation is less frequently addressed. In practical communication systems, modulation-induced phase drift, RF nonlinearity, and environmental variations can degrade the effective dynamic extinction ratio, motivating architectures that integrate real-time compensation mechanisms.

Comment 2 *Along similar lines, in the introduction, the authors claim: “Unlike conventional static bias optimization or thermal tuning, the ER-Boost structure enables dynamic extinction enhancement through an electronically reconfigurable phase control, allowing the system to continuously track and compensate for phase drift.” Also in the discussion, they write: “The key advantage of the proposed dual-branch configuration lies in its active compensation of structural imbalance and dynamic phase drift through a reconfigurable interference mechanism.” To me it is not clear in how far the cascaded approaches mentioned in my previous point would not allow for a similar reconfigurable tuning mechanism.*

Response:

We thank the reviewer for this important clarification. We agree that cascaded interferometric architectures can also incorporate active bias tuning mechanisms and are not inherently limited to static optimization.

Our intention was not to imply that cascaded approaches cannot implement dynamic compensation. Rather, the distinction lies in the architectural philosophy of extinction enhancement. Cascaded MZI structures primarily increase static extinction through multi-stage interference optimization, while the proposed ER-Boost architecture enhances extinction by introducing a dedicated auxiliary interference branch that directly compensates residual leakage.

To avoid overstatement and improve clarity, we have revised the relevant statements in both the Introduction and the Discussion sections to more accurately reflect this distinction.

To avoid overstatement regarding the distinction between the proposed ER-Boost architecture and cascaded interferometric approaches, we have revised the corresponding sentence in the Introduction.

Changes in the manuscript:

The updated text is reproduced below:

Rather than relying solely on static structural optimization or thermal tuning, the ER-Boost structure enables dynamic extinction enhancement through an electronically reconfigurable phase control, allowing the system to compensate for phase drift during operation.

Comment 3 *Also in the introduction, the authors write “Recent developments in silicon photonics and hybrid material platforms have enabled on-chip integration and ultra-high ERs exceeding 40 dB [19,20]” in the introduction. First of all, their reference [20], besides only showing simulation results, is a meta-surface which is a free space application and not on-chip integrated. Additionally, there are other demonstrations of ultra-high ER modulation which should be cited:*

- Jin et al. *Opt. Lett.* 44, 1265 (2019)
- Pohl et al. *IEEE Photon. Technol. Lett.* 33, 85 (2021)

Response:

We thank the reviewer for carefully examining the cited literature and for pointing out the incomplete coverage of relevant high-extinction-ratio demonstrations.

We have revised the Introduction to ensure that only experimentally demonstrated on-chip electro-optic modulators are cited. The previously cited reference that did not represent an integrated on-chip implementation has been removed. In addition, we have incorporated the suggested works by Jin et al. (*Opt. Lett.* 2019) and Wilkes et al. (*Opt. Lett.* 2016), which report ultra-high static extinction ratios using tunable beam-splitter and cascaded MZI architectures, respectively.

Changes in the manuscript:

The revised text is reproduced below.

Recent developments in silicon photonics, thin-film lithium niobate (TFLN), and hybrid material platforms have enabled on-chip integration and ultra-high static extinction ratios exceeding 40–60 dB [8]–[10]. For example, high-extinction modulation has been demonstrated using tunable beam-splitter architectures on TFLN [8] and auto-configured cascaded Mach–Zehnder interferometers on silicon platforms [9], achieving extinction ratios up to 50–60 dB under steady-state bias conditions. Nevertheless, many of these approaches

rely on multi-stage interference or complex structural optimization, which can increase fabrication complexity and sensitivity to device imperfections. Moreover, a significant portion of ultra-high ER demonstrations primarily report static or quasi-static extinction performance, while the dynamic extinction behavior under high-speed modulation and long-term operation is less frequently addressed. In practical communication systems, modulation-induced phase drift, RF nonlinearity, and environmental variations can degrade the effective dynamic extinction ratio, motivating architectures that integrate real-time compensation mechanisms.

Comment 4 *The performance metrics reported in this work appear subpar compared to the state-of-the-art literature. Specifically, the reported extinction ratio (ER) of up to 38 dB is lower than that reported by Jin et al. (Opt. Lett. 44, 1265 (2019)), who achieved up to 53 dB in a cascaded TFLN modulator, and by Nelan et al. (IEEE Access 10, 100300 (2022)), who achieved 45 dB in a single TFLN modulator. The experiments are performed at 400 MHz, which is orders of magnitude lower than the bandwidth capability of TFLN modulators, e.g., Wang et al. (Nature 562, 101 (2018)) and Pohl et al. (IEEE Photon. Technol. Lett. 33, 85 (2021)), both demonstrating bandwidths of 60 GHz or more. Although the manuscript claims high optical bandwidth, no measurements are provided. Other high-ER schemes report optical bandwidths up to 100 nm (Jin et al., Opt. Lett. 44, 1265 (2019)). CMOS-compatible V_π values have been demonstrated, for example, in Wang et al. (Nature 562, 101 (2018)), yet the current manuscript does not report V_π . Ideally, the authors should measure the optical and electro-optic bandwidths as well as V_π of their modulator and comment on the limiting factors.*

Response:

We thank the reviewer for the detailed comparison with state-of-the-art electro-optic modulators.

(1) Comparison with state-of-the-art static ER and bandwidth. We acknowledge that ultra-high static extinction ratios exceeding 50–60 dB and electro-optic bandwidths above 60 GHz have been demonstrated in optimized TFLN and silicon platforms. These results represent important advances in device-level performance.

(2) Distinction between static ER and dynamic ER. The primary focus of the present work is the dynamic extinction ratio under active high-speed modulation rather than the absolute static

extinction limit. While many prior works report static or quasi-static ER values under optimized bias conditions, here we explicitly evaluate the extinction ratio under RF modulation.

To clarify this distinction, we have added a new table summarizing the measured static and dynamic extinction ratios under different modulation frequencies.

(3) Frequency-dependent dynamic ER measurements. Under static (DC) bias conditions, the optimized ER-Boost configuration achieves an extinction ratio of 58 dB, confirming the high static interference suppression capability of the architecture. When RF modulation is applied, the dynamic extinction ratio gradually decreases with increasing frequency: 48 dB at 100 MHz, 42 dB at 200 MHz, 38 dB at 400 MHz, and 33 dB at 1 GHz.

This frequency-dependent reduction is primarily attributed to RF amplitude imbalance, finite electrical bandwidth, and phase non-idealities in the driving circuitry, rather than to a fundamental limitation of the ER-Boost interference mechanism. Importantly, even at 400 MHz and 1 GHz operation, the dynamic ER remains substantially higher than the native static extinction ratio of the underlying single-branch modulator (approximately 20–22 dB), demonstrating the effectiveness of the proposed leakage-cancellation architecture under practical modulation conditions.

(4) Modulation bandwidth and V_π . The experiments were conducted at frequencies up to 1 GHz due to instrumentation constraints. The commercial Ti:LiNbO₃ modulator used in this work supports an analog optical bandwidth of 22 GHz and a low-frequency V_π of 2.8 V, according to the manufacturer datasheet. The ER-Boost architecture itself does not introduce additional resonant or cascaded delay structures and therefore does not fundamentally limit electro-optic bandwidth.

(5) Contribution of this work. Rather than aiming to maximize the absolute static ER or bandwidth, this work demonstrates an architectural approach for dynamic extinction enhancement and real-time leakage compensation. The results show that substantial ER improvement can be maintained under active modulation, starting from a native ER of approximately 20–22 dB and reaching 33–38 dB under GHz-class operation.

Changes in the manuscript:

The following text and table have been added to the Results section:

TABLE II: Measured static and dynamic extinction ratios under different modulation frequencies.

Condition	Modulation Frequency	Extinction Ratio (dB)
Static (DC)	0 Hz	58
Dynamic	100 MHz	48
Dynamic	200 MHz	42
Dynamic	400 MHz	38
Dynamic	1 GHz	33

To further clarify the relationship between static and dynamic extinction performance, Table II summarizes the measured extinction ratios under DC operation and under RF modulation at different frequencies.

Under static (DC) bias conditions, the optimized ER-Boost configuration achieves an extinction ratio of 58 dB, confirming the high static interference suppression capability of the architecture.

This frequency-dependent reduction in dynamic extinction ratio is primarily attributed to RF amplitude imbalance, finite electrical bandwidth, and phase non-idealities in the driving circuitry, rather than to a fundamental limitation of the ER-Boost interference mechanism. Importantly, even at 400 MHz and 1 GHz operation, the dynamic ER remains substantially higher than the native static extinction ratio of the underlying single-branch modulator (approximately 20–22 dB), demonstrating the effectiveness of the proposed leakage-cancellation architecture under practical modulation conditions.

Comment 5 *The use of SNSPDs in the optical communication experiments makes the usage of their device in real-world optical communication links questionable. The authors should provide more context on the targeted applications. Can they also comment on if their enhanced ER would influence the link quality (e.g. the BER) if standard photodiodes would be used and whether their stepwise control strategy would produce the same results?*

Response:

We thank the reviewer for raising this important point regarding practical applicability.

First, we clarify that the SNSPD is used in this work as a high-sensitivity measurement tool to

accurately characterize the dynamic slot-level extinction ratio under photon-counting conditions. The ER-Boost architecture itself does not require SNSPD-based reception for operation. In practical communication systems, standard photodiodes (PIN or APD) can be used at the receiver side.

Second, the enhanced extinction ratio directly improves link quality even when conventional photodiodes are employed. A higher ER increases signal contrast, reduces residual optical leakage in the OFF state, improves eye opening in intensity-modulated links, and reduces decision threshold ambiguity, thereby lowering the bit error rate (BER). The BER improvements observed in our photon-counting experiments reflect the fundamental benefit of leakage suppression and are not restricted to SNSPD-based systems.

Third, the stepwise bias control strategy relies only on the integrated monitoring photodiode at the transmitter output and does not depend on the SNSPD used for performance characterization. Therefore, the same control principle can be applied in practical transmitters without requiring superconducting detectors.

To clarify the intended application scope, we have added additional discussion in the revised manuscript describing both photon-sensitive and conventional intensity-modulated communication scenarios.

Changes in the manuscript:

To clarify the practical applicability of the ER-Boost architecture beyond SNSPD-based laboratory measurements, we have added the following discussion in the revised manuscript:

The SNSPD used in this work serves as a high-sensitivity measurement instrument for dynamic extinction ratio characterization and is not required for practical deployment of the ER-Boost architecture. In conventional intensity-modulated communication systems employing standard photodiodes, higher extinction ratios improve signal contrast, increase eye opening, and reduce residual leakage in the OFF state, thereby lowering BER and enhancing link robustness. The proposed bias optimization strategy relies only on an integrated monitoring photodiode at the transmitter and can therefore be implemented independently of the receiver technology.

Comment 6 *Overall, the manuscript mostly lacks the proper context of the current state-of-the-art of TFLN modulators and high-ER switching for optical communications, which is crucial for understanding and assessing the impact and quality of the demonstrated device architecture. In addition to these major points, I would also like to raise the following more minor points:*

1, Can the authors provide more detail on the TFLN modulator? Was it fabricated in-house or by an external service? What is the film thickness, crystal cut, waveguide and electrode dimension?

2, More details on the communication experiments should be provided. Specifically the data rate and the link length should be mentioned.

3, For the bit-error-rates (BER) reported in fig. 5D, can the authors indicate the hard- and soft-decision forward error correction limits and whether they were able to reach those limits in their experiment?

4, The authors should provide the values of γ_E and γ_M used in their design, and consider indicating them in Fig. 1(d) to help readers understand the tolerance to fabrication imperfections.

5, In Section 2.4, the “ultralow dark count rate” is mentioned. Please specify the actual dark count rate as well as the average rates observed during the signal and guard slots.

Response:

We thank the reviewer for the comprehensive feedback and for highlighting several points requiring further clarification. We address each item below.

(1) TFLN modulator details.

We clarify that the experimental platform is not a newly fabricated TFLN device, but a commercially available Ti:LiNbO₃ dual-parallel MZM (Fujitsu FTM7961EX). To provide a clear performance baseline, we have included Table III summarizing the key specifications extracted from the manufacturer datasheet, including analog optical bandwidth (22 GHz), V_π (2.8 V at low frequency, 3.5 V at 22 Gbps), intrinsic extinction ratio (20–22 dB), insertion loss (7.0 dB), and RF impedance (50 Ω). This table explicitly distinguishes the baseline device performance from the extinction enhancement achieved by the ER-Boost architecture.

(2) Communication experiment details.

The communication experiments were conducted in a back-to-back configuration. The modulated optical output was attenuated to the single-photon regime for photon-counting evaluation. A slot rate of 400 MHz was used with a 16-PPM format plus 4 guard slots, and a

coding rate of 1/2 using SCPPM. These parameters have been clarified in the revised manuscript to provide complete experimental context.

(3) BER and FEC thresholds.

The BER performance reported in Fig. 5(d) corresponds to uncoded BER after SCPPM decoding. We have added clarification regarding typical hard-decision and soft-decision FEC thresholds and indicate whether the measured BER values approach these limits under the tested photon budgets.

(4) Power splitting ratios γ_M and γ_E .

The values of γ_M and γ_E used in the experiment have been explicitly indicated in Fig. 1(d) to clarify the fabrication tolerance region and the specific operating point used for extinction enhancement.

(5) SNSPD dark count rate and slot statistics.

The SNSPD used in the experiments exhibits a dark count rate of approximately 10 kHz. The average photon count rates in the signal and guard slots are reported in Fig. 5(c). We have explicitly stated the dark count rate in Section 2.4 to clarify the measurement dynamic range and sensitivity.

These additions improve the clarity and completeness of the experimental description and provide a more comprehensive comparison with state-of-the-art TFLN modulator performance.

Changes in the manuscript:

In response to the reviewer's comments regarding experimental context and device specifications, we have added the following clarifications and information in the revised manuscript.

The communication experiments were performed in a back-to-back configuration. The modulated optical signal was attenuated to the single-photon regime for photon-counting evaluation. A slot rate of 400 MHz was employed using a 16-PPM format with 4 guard slots, and a coding rate of 1/2 based on SCPPM. The SNSPD exhibits a dark count rate of approximately 10 kHz.

Furthermore, to provide a clear performance baseline of the experimental platform, we have included Table III, summarizing the key specifications of the commercially available Ti:LiNbO₃ dual-parallel MZM used in this study, extracted from the manufacturer datasheet:

TABLE III: Key specifications of the experimental Ti:LiNbO₃ dual-parallel MZM (Fujitsu FTM7961EX), based on manufacturer datasheet.

Parameter	Value
Platform	Ti:LiNbO ₃
Analog optical bandwidth	22 GHz
V_π (low frequency)	2.8 V
V_π (22 Gbps)	3.5 V
Native extinction ratio (parent MZI)	22 dB
Native extinction ratio (child MZI)	20 dB
Optical insertion loss (C-band, no connector)	7.0 dB
RF impedance	50 Ω

To provide a clear performance baseline of the experimental platform, Table III summarizes the key specifications of the Ti:LiNbO₃ dual-parallel MZM used in this study, extracted from the manufacturer datasheet. The intrinsic extinction ratio of the conventional single-branch operation is approximately 20–22 dB, which serves as a reference for evaluating the extinction enhancement achieved through the proposed dual-branch interference mechanism.

V. REVIEWER 5'S COMMENTS

We sincerely thank Reviewer #5 for participating in the peer-review process as part of the Communications Engineering co-review initiative. We appreciate the time and effort devoted to evaluating our manuscript and contributing to the constructive review process.

VI. REVIEWER 6'S COMMENTS

We sincerely thank Reviewer #6 for their participation in the peer-review process as part of the Communications Engineering initiative to support early-career reviewer training. We are grateful for the effort invested in assessing our manuscript.

REFERENCES

- [1] Z. Shi, W. Wang, X. Wu, and L. Yang, "A technique for measuring dynamic slot-power of optical PPM signals based on SNSPD," *IEEE Photonics Technology Letters*, vol. 34, no. 24, pp. 1377–1380, 2022.
- [2] M. Nakajima, K. Tanaka, and T. Hashimoto, "Scalable reservoir computing on coherent linear photonic processor," *Communications Physics*, vol. 4, no. 1, p. 20, 2021.
- [3] Z. Ying, C. Feng, Z. Zhao, S. Dhar, H. Dalir, J. Gu, Y. Cheng, R. Soref, D. Z. Pan, and R. T. Chen, "Electronic-photonic arithmetic logic unit for high-speed computing," *Nature communications*, vol. 11, no. 1, p. 2154, 2020.
- [4] F. Brücknerhoff-Plückelmann, H. Borrás, B. Klein, A. Varri, M. Becker, J. Dijkstra, M. Brücknerhoff, C. D. Wright, M. Salanga, H. Bhaskaran *et al.*, "Probabilistic photonic computing with chaotic light," *Nature Communications*, vol. 15, no. 1, p. 10445, 2024.
- [5] Y. Cheng, "Ultra-low loss large-scale photonic integrated circuits on thin film lithium niobate," in *Optical Components and Materials XXII*, vol. 13362. SPIE, 2025, pp. 33–43.
- [6] R. Li, Y. Gong, H. Huang, Y. Zhou, S. Mao, Z. Wei, and Z. Zhang, "Photonics for neuromorphic computing: Fundamentals, devices, and opportunities," *Advanced Materials*, vol. 37, no. 2, p. 2312825, 2025.
- [7] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani *et al.*, "Advances in quantum cryptography," *Advances in optics and photonics*, vol. 12, no. 4, pp. 1012–1236, 2020.
- [8] M. Jin, J.-Y. Chen, Y. M. Sua, and Y.-P. Huang, "High-extinction electro-optic modulation on lithium niobate thin film," *Optics letters*, vol. 44, no. 5, pp. 1265–1268, 2019.
- [9] C. M. Wilkes, X. Qiang, J. Wang, R. Santagati, S. Paesani, X. Zhou, D. A. Miller, G. D. Marshall, M. G. Thompson, and J. L. O'Brien, "60 db high-extinction auto-configured mach–zehnder interferometer," *Optics letters*, vol. 41, no. 22, pp. 5318–5321, 2016.
- [10] Z. Cheng, X. Shu, L. Ma, B. Chen, C. Li, C. Sun, M. Wei, S. Yu, L. Li, H. Lin *et al.*, "On-chip silicon electro-optical modulator with ultra-high extinction ratio for fiber-optic distributed acoustic sensing," *Nature Communications*, vol. 14, no. 1, p. 7409, 2023.
- [11] A. Nemati, Q. Wang, N. S. S. Ang, W. Wang, M. Hong, J. Teng *et al.*, "Ultra-high extinction-ratio light modulation by electrically tunable metasurface using dual epsilon-near-zero resonances," *Opto-Electron. Adv.*, vol. 4, no. 7, p. 200088, 2021.

Dear Editor Wenjie Wang,

We are writing to resubmit our manuscript entitled “Ultra-high extinction ratio modulation without perfect splitters using the ER-Boost Mach–Zehnder modulator” for consideration for publication in Communications Engineering with **COMMSENG-25-0880A**. We would like to express our sincere gratitude to you and the reviewers for the comprehensive and insightful feedback provided for our work.

All major changes, including updated experimental analysis and expanded theoretical derivations, are highlighted in **blue** in the revised manuscript and summarized within the response boxes in this letter. Please note that the citations in this letter refer to the references listed at the end of this document, rather than those in the manuscript itself, unless stated otherwise. Similarly, figure indexing in this letter is independent of the manuscript.

We believe that the revisions have significantly strengthened the paper and addressed all the concerns raised. Thank you for your time and for the opportunity to improve our work.

Yours Sincerely,

Ziyuan Shi, Zhengying Lou, Jiajie Xu, and Mohamed-Slim Alouini

I. RESPONSE TO REVIEWER 1

The authors have addressed all of my previous comments with detailed explanations. The manuscript has been significantly strengthened as a result, and I recommend it for publication in Communication Engineering.

Response: We sincerely thank the Reviewer for the careful evaluation of our manuscript and for confirming that the revised version addresses all previous concerns. We appreciate the reviewer's time and valuable feedback throughout the review process.

II. RESPONSE TO REVIEWER 2

The authors have addressed my concerns. I now recommend the publication of this manuscript.

Response: We would like to express our sincere gratitude to the reviewer for the positive evaluation of our revised manuscript and for acknowledging that all previous concerns have been properly addressed.

III. RESPONSE TO REVIEWER 3

I thank the authors for careful consideration of my comments. Most suggestions were implemented, questions answered, and inconsistencies resolved. Only a few comments remain.

Response: We sincerely appreciate the reviewer’s feedback on our paper. The thoughtful comments and suggestions are highly valued and have helped us improve the quality of the paper. We have continued to refine our work to ensure clarity and rigor throughout the paper.

Comment 1 *My comment (reviewer 3, comment 2) regarding the phrase “...both classical and...” in the abstract was addressed in the “rebuttal”, but the corresponding phrase was not removed. See Abstract, end of line 2.*

Response: We thank the reviewer for pointing this out. Although we indicated in the rebuttal that the phrase “...both classical and...” would be removed, it was inadvertently not updated in the manuscript. This has now been corrected in the revised version by modifying the abstract accordingly. We apologize for this oversight.

The revised manuscript is as follows

Achieving an ultra-high extinction ratio (ER) in optical modulators is crucial for enhancing signal fidelity, dynamic range, and energy efficiency in optical communication systems.

Optical modulation is a fundamental process in optical communication systems, governing how information is encoded onto light for transmission through fiber and free-space channels [1]–[3].

Comment 2 *My comment (comment 7a) regarding the differential nature of the drive signal was addressed in the “rebuttal”, but the corresponding line in the figure is not explicitly drawn as a differential signal. The authors claim in the “rebuttal” that “the differential nature of the RF drive after amplification is now explicitly indicated”. – I would expect separate p and n (or S and S^*) channels between the amplifier and the modulator.*

Response: We sincerely thank the reviewer for the careful observation. We agree that the previous description of the RF drive signal was not sufficiently clear and may have led to confusion.

In our system, the FPGA GTH interface generates a differential signal, which is subsequently converted into a single-ended signal through a balun before being applied to the MZM. Therefore, the electrical drive applied to the modulator is in a single-ended configuration.

To avoid ambiguity, we have revised the manuscript to clarify this signal path and ensure consistency between the figure and the textual description. We appreciate the reviewer's comment, which has helped improve the clarity of the presentation.

The revised manuscript is as follows

The RF data signal is generated as a differential signal and converted into a single-ended signal via a balun before being applied to the M-branch, while the data signal itself is unipolar.

Comment 3 *My comment (comment 10) regarding Fig. 2c was addressed but not fully resolved.*

a. The reference to a saddle-shaped surface in Fig. 2c (main text and caption) was not explained. I cannot see a saddle-point in Fig. 2c. I kindly ask the authors to indicate the saddle-point in the figure.

b. The typo remains, I think the authors meant to write “showing a sharp extinction boundary indicated with “E channel cutoff point”.

c, Measured

Response: We sincerely thank the reviewer for the careful examination and for pointing out these important issues. We acknowledge that the previous revision did not fully resolve the concerns, and we appreciate the opportunity to clarify and improve this part of the manuscript.

First, upon re-examining the data, we found that the previously presented Fig. 2c was in fact generated from measurements collected during the Stage III optimization process, where the M-branch bias voltage $V_{\text{bias,M}}$ and the RF peak-to-peak voltage V_{pp} were jointly adjusted. However, this stage is primarily intended for optimizing the modulation branch and is not suitable for illustrating the saddle-like optical power behavior discussed in the manuscript. As a result, the plotted optical power surface did not accurately reflect the intended characteristics, and the saddle-like feature was not clearly visible.

In the revised manuscript, we have replaced Fig. 2c with data corresponding to the Stage II optimization process, where the E-branch is tuned to minimize its output optical power. Under

this condition, the optical power surface more clearly exhibits the saddle-like behavior associated with the joint dependence on $V_{bias,M}$ and V_{pp} . The corresponding saddle-like operating point is now explicitly indicated in the figure.

In addition, we have corrected the typographical error in the caption, and revised the figure description to ensure clarity and consistency.

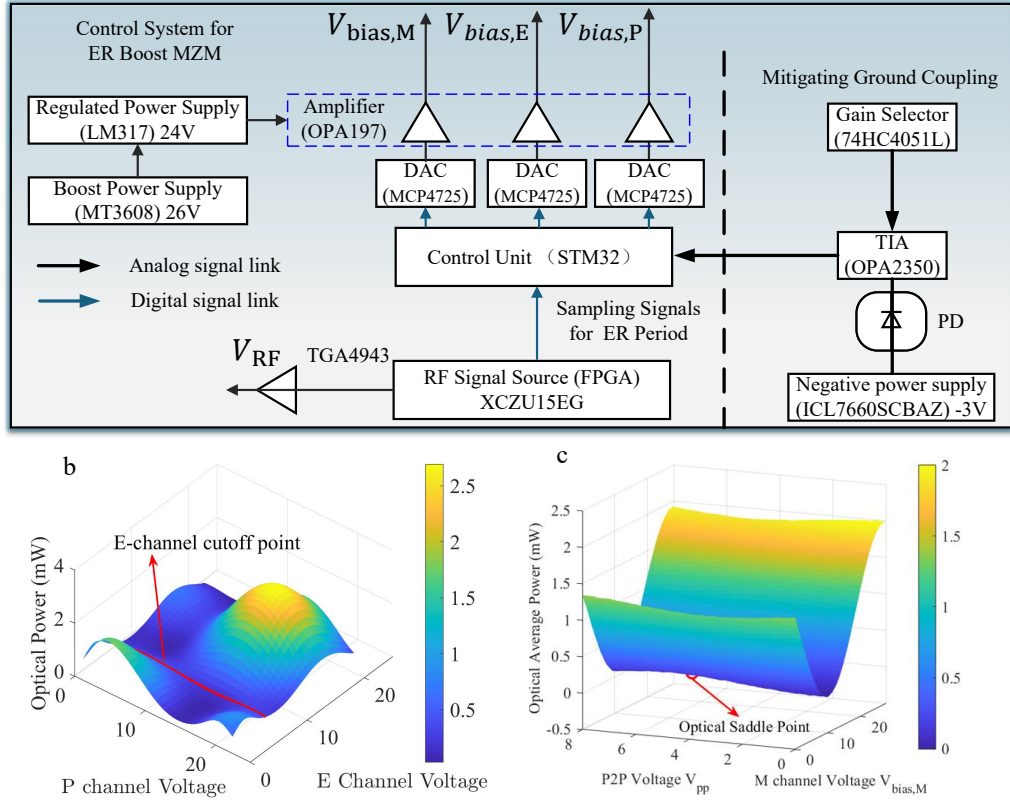


Fig. 1: Control architecture and performance analysis of the ER-Boost MZM bias optimization system. a, Block diagram of the custom-designed control board, incorporating an STM32 control unit, DACs, amplifiers, and a transimpedance amplifier (TIA) for PD readout. The system also integrates a ground isolation stage and a gain selector. b, Measured output optical power as a function of $V_{bias,P}$ and $V_{bias,E}$, showing a sharp extinction boundary indicated by the “E-channel cutoff point”. c, Measured optical power response with respect to $V_{bias,M}$ and the RF peak-to-peak voltage V_{pp} . The resulting surface exhibits a saddle-like shape. This behavior arises from the employed PPM modulation format, whose duty cycle is significantly smaller than 50%, causing the time-averaged optical power to remain dependent on $V_{bias,M}$ even when V_{pp} is nominally set to V_{π} .

For comparison and transparency, we also provide the original data (without ER-Boost period separation). The corresponding optical power surface is shown in Fig. 2, which clearly lacks the saddle-like feature due to the absence of stage-specific optimization.

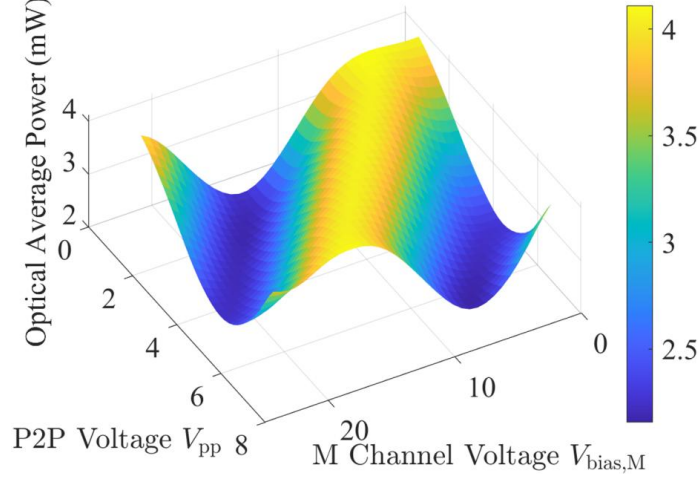


Fig. 2: Original optical power surface without ER-Boost period separation.

Comment 4 *My comment (comment 11) regarding Fig. 3 was addressed but not fully resolved. The caption still claims “The bias controller dynamically tunes $V_{bias,M}$, $V_{bias,E}$, and $V_{bias,P}$ during the ER period”, while the “rebuttal” states that “During Stage III (ER optimization), only $V_{bias,E}$ and $V_{bias,P}$ are adjusted, while the M-path bias is held constant.”*

Response: We sincerely thank the reviewer for pointing out this inconsistency between the figure caption and our description in the rebuttal. We apologize for the oversight in the previous revision.

In the revised manuscript, we have corrected the caption of Fig. 3 to ensure consistency with the described bias control procedure. Specifically, during the ER optimization period, only $V_{bias,E}$ and $V_{bias,P}$ are dynamically adjusted, while the M-branch bias is maintained constant after its prior optimization stage.

Revised caption (excerpt):

c, Temporal evolution of the two bias voltages applied to the enhancement (E) and phase-control (P) channels. The bias controller dynamically tunes $V_{bias,E}$ and $V_{bias,P}$ during the

ER period to minimize residual optical leakage, while maintaining stable bias levels during normal transmission.

Comment 5 *My comment (comment 13) regarding the legends of Fig. 5c-d was addressed but not fully resolved. The authors claim in the “rebuttal”, that “the order of the legend entries has been unified to follow the same ascending ER sequence for clarity”, but the legend entries are not in ascending order. The color scheme was unified, but the markers are still inconsistent between the two plots.*

Response: We sincerely thank the reviewer for the careful examination and for pointing out the remaining inconsistencies in Fig. 5c–d. We apologize for the oversight in our previous revision.

In the revised manuscript, we have made the following corrections: (i) the legend entries in both Fig. 5c and Fig. 5d have been strictly reordered to follow an ascending extinction ratio (ER) sequence; (ii) the marker styles have been unified such that each ER value is represented by the same marker type across both subfigures; and (iii) the consistency between color and marker encoding has been ensured throughout.

Revised description:

The legend entries in Fig. 5c and Fig. 5d are now consistently ordered in ascending ER (15.3 dB, 24.5 dB, and 32.1 dB). Each ER value is represented by a unique and consistent combination of color and marker style across both subfigures.

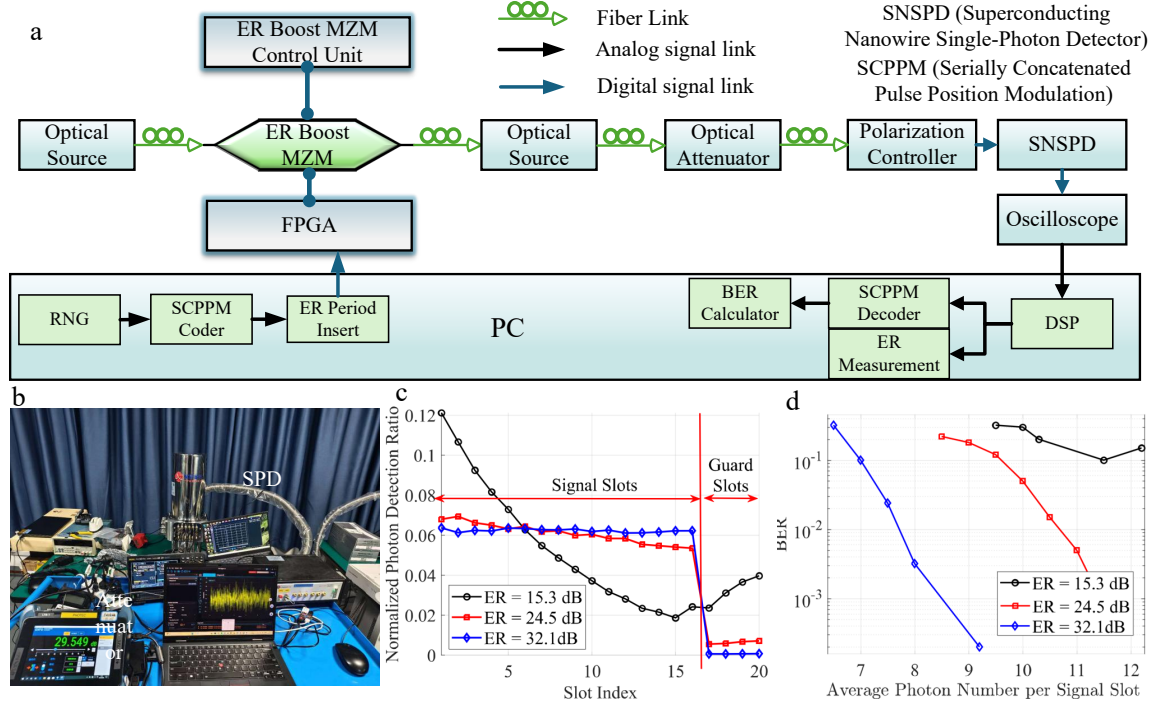


Fig. 3: Experimental setup and performance evaluation of the proposed ER-Boost MZM system. a, Block diagram of the optical transmission experiment integrating the ER-Boost MZM with an FPGA-based SCPPM encoder/decoder and real-time bias control unit. b, Photograph of the experimental setup including optical source, attenuator, SNSPD, and data acquisition system. c, Normalized photon detection ratio across signal and guard slots under different ERs with 16-PPM (15.3 dB, 24.5 dB, and 32.1 dB). d, BER performance as a function of the average photon number per signal slot, demonstrating the clear advantage of higher ERs.

Comment 6 The caption of Fig. 5 uses brackets to identify subfigures, i.e. “(a)”, while the other captions use a comma, i.e. “a,”. This should be unified.

Response: We thank the reviewer for pointing out this inconsistency in the figure caption formatting. We have revised the caption of Fig. 5 to ensure consistency with the formatting used in the other figures, replacing “(a), (b), (c), (d)” with “a, b, c, d,”.

Revised caption of Fig. 5:

a, Block diagram of the optical transmission experiment integrating the ER-Boost MZM with an FPGA-based SCPPM encoder/decoder and real-time bias control unit.

- b, Photograph of the experimental setup including optical source, attenuator, SNSPD, and data acquisition system.
- c, Normalized photon detection ratio across signal and guard slots under different ERs with 16-PPM.
- d, BER performance as a function of the average photon number per signal slot.

Comment 7 *Table 2 is highly appreciated, but there is probably a typo in row 4, it should presumably read “DC ports”*

Response: We thank the reviewer for the careful reading and for pointing out this typographical error in Table 2. We have corrected “DC Prots” to “DC Ports” in row 4 of the table.

TABLE I: Key specifications of the experimental Ti:LiNbO₃ dual-parallel MZM (Fujitsu FTM7961EX), based on manufacturer datasheet.

Parameter	Value
Platform	Ti:LiNbO ₃
Analog optical bandwidth	22 GHz
V_π at DC Ports	17V
V_π (22 Gbps)	5.5 V
Native extinction ratio (parent MZI)	22 dB
Native extinction ratio (child MZI)	20 dB
Optical insertion loss (C-band, no connector)	7.0 dB
RF impedance	50 Ω

Comment 8 *Typo: In section 2.2 in the second paragraph, the first sentence ends with a double period.*

Response: We thank the reviewer for pointing out this typographical error. We have corrected the double period at the end of the first sentence in the second paragraph of Section 2.2.

Fig. 1d illustrates the simulated maximum ER as a function of the power splitting ratios of the modulation branch (γ_M) and the extinction-enhancement branch (γ_E), where γ denotes the optical power splitting ratio of the Y-branch, defined as the fraction of optical power

directed into the upper arm relative to the total optical power distributed between the upper and lower arms.

IV. RESPONSE TO REVIEWER 4

The authors have substantially improved the manuscript in response to my and the other referees' comments. The revisions not only enhance the clarity and flow of the manuscript but also place the work better in the current state-of-the-art in the field. Most of the concerns raised in the previous review round have been addressed and clarified. As a result, the manuscript is now more coherent and compelling, representing a clear advance over the original submission. In particular, the clarification that a ultra-high dynamic extinction ratio (ER) is the target here distinguishes their work from other cited literature and highlights the novelty of their work better. Therefore I would recommend publication, the authors may choose to address some of the comments below if they find it improves the manuscript:

Response: We sincerely appreciate the reviewer's feedback on our paper. The thoughtful comments and suggestions are highly valued and have helped us improve the quality of the paper. We have continued to refine our work to ensure clarity and rigor throughout the paper.

Comment 1 *Although in the manuscript they clearly distinguish between static and dynamic ER, in the abstract the authors may want to include this by including the word "dynamic".*

Response: We thank the reviewer for this helpful suggestion. We agree that explicitly distinguishing the dynamic extinction ratio in the abstract improves clarity and ensures consistency with the terminology used throughout the manuscript.

In the revised manuscript, we have updated the abstract to explicitly refer to the reported extinction ratio as the dynamic extinction ratio.

Revised text in Abstract:

Here, we propose and experimentally demonstrate a dual-branch ER-Boost MZM that achieves a dynamic ER exceeding 38 dB without requiring perfect 50:50 splitters or bandwidth compromise.

Achieving an ultra-high dynamic extinction ratio (ER) in optical modulators is crucial for enhancing signal fidelity, dynamic range, and energy efficiency in optical communication systems.

Comment 2 *In fig. 1c) the legend overlaps more than half the actual graph. It would be better to place the legend elsewhere.*

Response: We thank the reviewer for the comment. We would like to clarify that the waveform shown in Fig. 1c corresponds to a time-domain measurement of a periodically modulated signal.

In this experiment, a fixed-duty-cycle modulation format (PPM-like signal) is employed. As a result, the transmitted waveform is inherently periodic, and the observed signal repeats identically over consecutive time slots. The figure displays multiple periods of this waveform to clearly illustrate the structure of the signal and idle slots, as well as the corresponding extinction behavior.

V. RESPONSE TO REVIEWER 5

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Response: We appreciate the reviewer's time and valuable feedback throughout the review process.

VI. RESPONSE TO REVIEWER 6

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Response: We appreciate the reviewer's time and valuable feedback throughout the review process.

REFERENCES

- [1] G. P. Agrawal, *Fiber-optic communication systems*. Hoboken, NJ, USA: John Wiley & Sons, 2012.
- [2] H. Takesue and K. Inoue, “Generation of 1.5- μ m band time-bin entanglement using spontaneous fiber four-wave mixing and planar light-wave circuit interferometers,” *Physical Review A—Atomic, Molecular, and Optical Physics*, vol. 72, no. 4, p. 041804, 2005.
- [3] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani *et al.*, “Advances in quantum cryptography,” *Advances in optics and photonics*, vol. 12, no. 4, pp. 1012–1236, 2020.

Ultra-High Extinction Ratio Optical Modulation via Interference-Engineered Mach–Zehnder Modulator

Ziyuan Shi, Zhengying Lou, Jiajie Xu, and Mohamed-Mlim Alouini

July 3, 2026

We would like to convey our appreciation to the editor and the reviewers for their comments that helped us substantially improve the quality of our manuscript. Please find attached our revised manuscript COMMS-ENG-25-0880B, entitled “Ultra-High Extinction Ratio Optical Modulation via Interference-Engineered Mach–Zehnder Modulator,” Ziyuan Shi, Zhengying Lou, Jiajie Xu, and Mohamed-Mlim Alouini, in which we have incorporated and addressed the changes requested by the editor and reviewer.

We revised the entire manuscript to improve the quality of the presentation and satisfy the requests of the reviewer. All the comments of the reviewers have been carefully addressed, and the paper has been revised accordingly. The answers to all the comments of the editor and the reviewers are reported in detail below, and the changes with respect to the previous manuscript are explained.

In this response letter, we have included the reviewer’s comments in *italics*, followed by our response. In the response letter, all the modified/added paragraphs to address the reviewer’s comments are displayed in blue color. We also note that the citations in this letter refer to the references included at the end of the document, not the ones in the original or revised manuscript unless stated explicitly otherwise. Likewise, the figure indexing of this letter does refer neither to the original nor the revised manuscript.

Response to the Reviewer 3

I thank the authors for carefully addressing the remaining comments and requests. The manuscript has been significantly strengthened. I recommend it for publication in Communication Engineering.

Authors: We sincerely thank the reviewer for the positive evaluation of our manuscript. We would like to express our sincere gratitude for the thorough review of our manuscript.

Response to the Reviewer 4

The authors have addressed all my remaining comments and I can recommend the manuscript for publication.

Authors: We would like to express our sincere gratitude for the thorough review of our manuscript. We thank the reviewer for the positive comments and appreciation of our revisions.

Response to the Reviewer 5

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Engineering initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Authors: We sincerely appreciate the reviewer's feedback on our paper. The thoughtful comments and suggestions are highly valued and have helped us improve the quality of the paper.