nature portfolio

Peer Review File

Thin-Film Lithium Niobate Waveguide-Integrated p-i-n Tungsten Selenide Photodiodes with 10Gbit/s Data Reception

Corresponding Author: Professor Shun Wang

Parts of this Peer Review File have been redacted as indicated to remove third-party material.

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)
This manuscript reports a WS2-integrated lithium niobate (LN) waveguide photodetector capable of 10 Gbit/s data reception. The materials integration and device fabrication appear to have been carried out with reasonable care, and the experimental execution is generally solid. However, in my view, the conceptual advance and overall level of novelty do not meet the standard expected for Nature Communications.
My main concern lies in the motivation for this specific device platform and wavelength range, which is not yet sufficiently convincing. The manuscript does not clearly establish what unique opportunity or unmet need is addressed by implementing a visible-wavelength photodetector on an LN waveguide platform. In the broader context of LN photonics, the most established and compelling strength of LN lies in high-speed modulation, typically in the C-band. From the perspective of monolithic integration, a more persuasive direction would therefore be high-speed photodetection in the same spectral window.
In parallel, visible-band photodetection and interconnect-related applications are already being actively pursued on other material platforms, particularly GaN-based technologies, where substantially higher aggregate data capacities (512 Gbit/s) have been demonstrated in multi-channel systems (i.e. Avicena). Against this background, the manuscript does not yet provide a sufficiently compelling system-level or application-level rationale for choosing the WS2/LN platform for visible-band high-speed detection. For this reason, I believe the work may be more suitable for a specialized photonics journal, such as Optics Express, rather than Nature Communications.
Specific comments
1. Although the device is presented as a waveguide-integrated photodetector, the influence of the waveguide geometry on detector performance is not adequately addressed. The Methods section only states that a width of 450 nm was chosen to satisfy the single-mode condition, but there is no meaningful discussion of how the optical mode overlaps with the WS2 layer, nor of how parameters such as etch depth, sidewall angle, and modal confinement affect absorption, responsivity, or bandwidth. Without such analysis, it is difficult to assess whether the reported device performance is close to optimal or simply reflects one unexamined design point.
2. In the Methods, the authors state that Ar plasma ICP-RIE was used to achieve “smooth, vertical sidewalls.” This point requires clarification. In LNOI processing, achieving truly vertical sidewalls is nontrivial. The authors should provide supporting evidence, such as cross-sectional SEM images, sidewall angle measurements, or a more detailed description of the etching conditions and process optimization, to substantiate this claim.
3. In Fig. 1(a), the LN waveguide is illustrated with nearly vertical sidewalls. This representation is potentially misleading, as perfectly vertical sidewalls are generally difficult to realize in practical LNOI fabrication. The schematic should be revised to more faithfully reflect the actual fabricated geometry, or the authors should provide clear experimental evidence showing that such sidewall profiles were indeed achieved.
4. It is unclear whether a SiO2 upper cladding is present above the LN waveguide. This is an important structural detail and should be explicitly stated. If an oxide cladding is used, its thickness should be given, and the authors should explain how sufficient optical-field interaction with the WS2 absorber is maintained through the cladding. If no upper cladding is present, then the etched LN topography may introduce a substantial step height and sloped sidewall geometry, which raises concerns regarding the mechanical continuity and conformal transfer of the WS2 film. In that case, the authors should discuss whether cracking, delamination, or strain-induced performance variation could occur, and whether the bending or deformation of the WS2 layer affects the detector characteristics.

5. The system-level measurement conditions are not sufficiently documented. The manuscript should clearly specify the optical source power, as well as the optical power coupled into the chip. In addition, if the transmitter is based on direct modulation of the optical source, the signal quality before entering the photodetector should also be characterized and reported. For example, the authors should clarify whether the optical waveform, extinction ratio, eye quality, or any relevant transmitter-side performance metrics were measured prior to detection. At present, these key details are missing, making it difficult to assess whether the reported system performance is limited by the detector itself or by the quality of the input optical signal.

6. In Fig. 4(f)–(h), the eye diagrams are shown without a time scale. This makes it difficult to evaluate the signal quality and to interpret the data quantitatively. A horizontal time axis, or at least the time window per division, should be included.

7. The grating coupler performance is not sufficiently documented. What is the coupling efficiency of the grating couplers, and was it measured experimentally? In Supplementary Fig. 1(f), the grating structure appears to show visible damage or deformation. The origin of this feature should be explained, and the authors should comment on whether it affects insertion loss, coupling stability, or device-to-device reproducibility.

Reviewer #2

(Remarks to the Author)
Weijia Tang et al. present a WSe2 photodiode integrated on a TFLN waveguide platform. The PD is targeted for operation in the visible regime and is tested at 638nm. The paper does basic photodetector charactzerization as well as bandwidth and data communication experiments. Furthermore, they show in detail how the doping and contacts are designed to achieve an effective p-i-n structure. The characterization is done well and extensive, but i suggest the following revisions for the paper to be considered for publication in Nat. Comm. :

- 1) There is quite a bit of repetition within the paper due to the structure that was chosen which makes the paper harder to understand. The paper goes from vision, to PD performance, to contact engineering, to FET characterization and back to PD operation. I suggest to the authors reconsider the structure to remove repetitions and make the structure clearer. Personally, I would also refrain from the quite often repeating use of the words like "ultra", "extremely" or "outstanding" which don't add anything to the manuscript in my opinion. Additionally, there are several typos and grammatical errors that should be fixed.
- 2) The specific detectivity evaluation: this is done purely by calculation with the dark current and responsivity. It would be much more correct to measure the frequency dependent noise spectral density. Additionally, in Figure 1d the authors report the specific detectivity for varying optical input powers. Here, there would be shot noise occuring from the photocurrent and not just noise from the dark current further reducing specific detectivity. I recommend to the authors to look into: doi.org/10.1038/s41467-023-37635-1 and doi.org/10.1038/s41566-025-01759-1
- 3) I'm bothered by the EQE analysis and figure. Figure 1e shows data points that clearly exceed 100%, but the authors report a "near-unity EQE" of 98%. Where does this missmatch come from? Also, could the authors show the optical absorption of the WSe2 PD? Figure 4d shows a reference output that would allow for such an analysis. Are there mode simulations for the TFLN waveguides with the WSe2 on top that would theoretically match the measured absorption and EQE? Along these lines: could the authors comment on the grating and propagation losses that they compensate? Ultimately, this is an internal responsivity and not a system responsivity that is reported.
- 4) With what optical input power did the authors perform the data communication experiements? According to Suppl. Figure 4b, the PD saturates and is restricted to the output of a few uA. Is this really enough to achieve these data rates? Or was there a trans impedance amplifier used to boost the signal electrically?
- 5) On the achieved bandwidth: could the authors illustrate the p-i-n channel of the PD, how long it is and where absorption takes place? What are the resulting transit times and carrier velocities? Are these values realistic for WSe2? This analysis would provide a more physical interpretation of the achieved bandwidth.
- 6) The authors heavily emphasize visible light communications for "short-reach high-density ultra-low-power-consumption" applications. However, most TFLN development and existing integrated photonic ecosystems operate in the infrared telecommunications bands. Could the authors comment on this? Why the switch to visible PDs?
- 7) The SOTA comparison: The comparison on the TMDC mobilities is done well. The SOTA comparison in Table 1 on the hand seems very selective and incomplete. What do the authors want to state with this SOTA table? What is the advantage of the presented work over established technologies which are not listed at all? It seems to me that Si-PIN and GaAs-PIN diodes can easily outperform the here presented work. The conclusion

Reviewer #3

(Remarks to the Author)
This study reports on the integration of WSe2 flakes onto thin-film lithium niobate (TFLN) waveguides to realize high-performance integrated photodetectors. The authors demonstrate an impressive external quantum efficiency (EQE) of approximately 98%, alongside a high responsivity and a 3-dB bandwidth of 1.12 GHz, achieved through a lateral p-i-n junction architecture. While the device fabrication and the integration of 2D materials with TFLN waveguides represent significant technical advancement, several aspects regarding the physical interpretation of the efficiency and the device's operational mechanisms require further justification. Therefore, we believe the article requires a major revision to clarify these points before it can be considered for publication in Nature Communications.
Major comments:
1. The authors claim that the high dielectric constant of the TFLN substrate effectively screens ionized impurity scattering. However, considering the intrinsic ferroelectric nature of lithium niobate, a significant gate hysteresis is expected in the IDS-

VGS characteristics. The current Figure 3 lacks bidirectional sweep data, which is essential to evaluate the interfacial quality and the reliability of the extracted mobility. The authors should provide a hysteresis analysis to support their claim of a high-quality junction with minimal trap states.
2. The authors report a near-unity EQE of 98%, but the reference point for the input optical power is unclear. It is critical to distinguish whether these values are external (measured at the fiber) or internal (de-embedded for coupling/propagation losses). Please provide a detailed power budget and clarify exactly where the input power was measured to ensure the performance is not overestimated through excessive de-embedding
3. The 98% EQE claim is based on the assumption that all attenuated optical power is absorbed by the WSe2. However, the 30 nm thick flake on a 450 nm waveguide inevitably causes mode mismatch and scattering losses at the edges. Without distinguishing these non-absorption losses from actual absorption, the reported efficiency is likely misinterpreted. The authors must provide numerical or experimental evidence to quantify scattering losses and prove that they are not being misinterpreted as photocurrent-generating absorption
4. The contact potential difference (CPD) variations reported for p- and n-type regions are relatively small (35 mV and 44 mV, respectively). Given the high carrier mobilities and bandwidths claimed, one would expect a more significant Fermi level shift. Furthermore, the n-type region was characterized after Ar plasma exposure, which could introduce surface states or adsorbates in ambient conditions. How can the authors distinguish the intrinsic doping effect from surface-damage-induced potential changes in their KPFM analysis?
5. The authors emphasize the role of the MoOx buffer layer by citing a work function of 6.7 eV for MoO3. While the results are promising, a few points require further clarification to strengthen the manuscript's physical interpretation. First, since the work function of molybdenum oxide is highly sensitive to the O/Mo ratio, it is recommended to provide XPS data to confirm the stoichiometry of the evaporated film. The reported KPFM shift of 35 mV appears relatively small for a 6.7 eV layer, raising questions about whether the actual work function is lower due to oxygen vacancies or if the 3 nm film is discontinuous. Providing high-resolution TEM or AFM data would help confirm the uniformity of this ultra-thin interface and further validate the high device performance
6. The waveguide has a total thickness of 300 nm with an etch depth of approximately 135 nm, creating a substantial step height of ~165 nm for the 30 nm thick WSe2 flake to conform over. Given that this step height is more than five times the flake's thickness, significant localized strain is inevitable. How did the authors ensure uniform interfacial contact and account for the resulting strain-induced variations in the device's electronic and optical properties?
7. The manuscript lacks a clear definition of the intrinsic (i) region's width. This dimension is critical as it determines the trade-off between junction capacitance and carrier transit time. Without specifying this width, it is difficult to evaluate the physical origin of the reported 1.12 GHz bandwidth. The authors should provide the dimensions and explain how the i-region was optimized.

Minor comments:
8. The manuscript contains several typographical errors that should be carefully reviewed. (e. g. “peek shift” in page 7). Additionally, some acronyms are not defined on first mention, for example, FLP or AFM.
9. While Ar plasma treatment aligns the Fermi level, the resulting Se vacancies can create trap states that lead to FLP, potentially limiting gate tunability. The authors should clarify if these traps induce persistent photoconductivity or sub-threshold degradation.
10. The reported dark current values show inconsistencies between the figures. For instance, between Supplementary Figures 3a and 5a. Can you clarify this?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The revised manuscript has improved, and the newly added cross-sectional characterization of the heterogeneous waveguide is particularly valuable. Nevertheless, several important issues remain unresolved. At this stage, I am not yet in a position to support publication, and I would like to evaluate a further revised version before making a final recommendation.

1. I remain unconvinced by the authors’ response regarding the motivation for operating in the visible wavelength range. Current optical I/O technologies are predominantly developed for the O- or C-band, supported by mature light-source, waveguide, coupling, and packaging ecosystems. By comparison, operation in the visible range presents substantial challenges, particularly in terms of efficient light sources, waveguide loss, sidewall-scattering sensitivity, and system integration. The present manuscript does not yet establish a sufficiently compelling system-level advantage that would justify moving optical I/O into this wavelength regime. The authors should therefore substantially strengthen the motivation in the main text and clearly explain what unique benefit the visible band provides, which application scenario requires it, and whether this benefit outweighs the associated technological challenges. Without such clarification, the broader significance and practical relevance of the proposed direction remain questionable.

2. Figure 1a remains confusing and requires substantial revision. First, the label “Fiber Edge-Coupling” is inaccurate; “fiber-to-chip edge coupling” or simply “edge coupling” would be more appropriate. Second, the signal-flow direction and the ordering of the components are unclear. If the intended signal direction is from the fiber to the chip, it is not evident why the signal first passes through the demultiplexer, particularly because the input is illustrated as broadband. The functions of several other elements are also not identified. For example, the components electrically connected to the XPU through wire

bonds should be explicitly labelled. Although this is intended as a conceptual illustration, it should still present a physically consistent architecture with unambiguous signal directions, component functions, and interconnections. In its current form, the figure raises more questions than it resolves.

3. There appears to be an inconsistency between the simulation geometry in Extended Data Fig. 1b and the experimentally characterized waveguide cross-section. The simulated waveguide still appears to have vertical sidewalls, whereas the newly provided characterization indicates a sidewall angle of approximately 70°. The authors should clarify how the sidewall angle is defined and whether the experimentally measured geometry was incorporated into the simulation. If not, the simulations should be repeated using the realistic waveguide cross-section, and the authors should discuss whether the nonvertical sidewalls affect the modal properties, confinement, coupling behavior, or other key performance metrics reported in the manuscript.

Reviewer #2

(Remarks to the Author)
I thank Weiijia Tang et al. for the revisions. I'll go through them by each point:
-1) The readability of the paper is much improved with the changes made to the manuscript.
-2) The new discussion on the specific detectivity is good.
-3) The repeated measurements, inclusion of error bars and averaging to evaluate the EQE is a great improvement. The value of 95.6% is already more realistic than the value of 98%, considering that the absorption is not 100%. However, the new Figure 1e still shows EQE above 100% without error bars on the data point that go below 100%. One issue could be deviations between the calibration grating and channel and the actual device leading to further uncertainty. Or could the authors bias the device higher to show that their EQE flattens out?
-4) On the photocurrent: 11.8 uA at an optical input power of 193.5 uW means the device is operated at the maximum that is reported in Figure 1c and is the saturated output of device. This corresponds to a responsivity of ~61 mA/W which means the device is operated at a point where it has diminished efficiency. However, it seems this was necessary to perform the data measurements. This seems to be a major drawback of the device that only ~10 uA of photocurrent can be produced and the maximum achievable photocurrent should be discussed in the manuscript.
On the SNR: I'm well aware of SNR and NPDR. The limiting factor when doing data transmission with such low RF powers can be the scope. E.g., a 6 GHz Keysight oscilloscope has on the highest sensitivity an input noise floor of ~ -61 dBm (or ~4 uA rms). This noise floor is typically higher if you go to higher bandwidth oscilloscopes. Just as a rough estimation, with ~10 uA the ideal RF power is P = I^2 * 50 Ohm = 5 nW = -53 dBm, so not even an order of magnitude above the noise floor.
The papers the authors cite further point towards my question regarding an amplifier:
"For context, similar photocurrent levels have been reported to achieve Gbit/s data rates: 30 μA (corresponding to NPDR ≈ 85 mW-1) for 3 Gbit/s (Nat. Photonics 2015, 9, 247-252) …"
-> "… of the measurement instruments (including the current amplifier and bias tee)"
" …4 μA (NPDR ≈ 600 mW-1) for 1 Gbit/s rate (ACS Photonics 2018, 5, 1846-1852)"
-> "The generated electrical signal was read out with a microwave probe, amplified with a 50 Ω RF power amplifier, and finally detected by a digital storage oscilloscope."
Depending on the measurement equipment, I see that that an amplifier free operation is possible to achieve data transmission. Nevertheless, I would disagree with the statement that "the μA level photocurrent is sufficient for the real time oscilloscope (RTO) to sample the signal". Or could the authors show data comm where the device is not operated in a regime where the responsivity is diminished? This would be at around ~1 uA output current?
Again, I would ask the authors to include a complete experimental description of the data measurements, i.e., include the operation condition on the device, especially the optical input power. Additionally, as mentioned above, the authors should discuss the maximum photocurrent output which is a major drawback of the device.
-5) I thank the authors for the inclusion of this discussion; this is an interesting addition.
-6) The new discussion on the motivation of short-reach applications in the introduction is done well.
-7) The clarification on the scope of the SOTA table helps.

Reviewer #3

(Remarks to the Author)
The authors have provided a thorough and careful response, which I believe has significantly strengthened the manuscript.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I thank the reviewer for their efforts in revising the manucript.

A minor issue that should be addressed is in Introduction, "For emerging chip-scale OIO, multiple device platforms are currently being actively explored, including silicon photonics, thin-film lithium (TFLN) photonics and vertical-cavity surface-emitting laser (VCSEL)-based systems [1, 5, 6]", the VCSEL-based system should not be listed along with silicon photonics and TFLN platforms, since they are different materials systems. The author may say "III-V materials systems."

I can recommend its publicaion after correcting the abovementioned sentence.

Reviewer #2

(Remarks to the Author)
I thank the Author's for their detailed answers. The manuscript is now in a status where everything seems to be present such that a reader can come to their own conclusion. I recommend the manuscript for publishing. Yet, I still want to add two comments where I'll leave the decision with the editor:

- "Upon verification, we acknowledge that the measurement conditions were not clearly described in our previous response,where we confused the configurations for small-signal frequency response measurements with those for optical communication tests. We apologize for this confusion."
--> Such a statement does not spark a lot of confidence in the presented results, making me question what else is wrongly described and presented. Omitting a 40dB amplification is not a small thing, as the results would otherwise not have been achieved. And answering my remark by insisting it is possible and citing other works that contradict the author's own statement adds to the point where confidence in the author's understanding and description of results is further lowered.

- While a careful reader might notice and draw their own conclusions, I still find statements like the one in the abstract potentially misleading: "The device features a record-high measured bandwidth of 1.12 GHz and at least 10 Gbit/s data rate in visible light. Simultaneously, the integration of TFLN waveguide overcomes the traditional speed-responsivity trade-off, achieving a responsivity of ~492 mA·W-1 ..."
--> 10 Gbit/s data rate was achieved, but not simultaneously with a responsivity of 0.5 A/W. The responsivity was clearly diminished by operating the device at high optical input powers. In the conclusion, the same statement is made using "alongside" instead of "simultaneously". And yes, the device itself achieves these parameters (all on in device), but not with the same operation conditions. While re-reading this I also noticed that the abstract states at least 10 Gbit/s, while the conclusion states up to 10 GBit/s.

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/

Manuscript ID: NCOMMS-26-024174-T

Title: Thin-Film Lithium Niobate Waveguide-Integrated p-i-n Tungsten Selenide Photodiodes with 10Gbit/s Data Reception

Point by Point response to the Reviewer's Comments

We sincerely appreciate the reviewers for the constructive comments and insightful suggestions, which have greatly improved the quality and rigor of our manuscript. We have carefully addressed all the raised concerns with detailed point-by-point responses below. For easy identification, all revisions in the main text and Supplementary Information are highlighted in blue. We hope the revised manuscript meets the publication standard.

Reviewer #1 (Remarks to the Author):

This manuscript reports a WS_2-integrated lithium niobate (LN) waveguide photodetector capable of 10 Gbit/s data reception. The materials integration and device fabrication appear to have been carried out with reasonable care, and the experimental execution is generally solid. However, in my view, the conceptual advance and overall level of novelty do not meet the standard expected for Nature Communications.

My main concern lies in the motivation for this specific device platform and wavelength range, which is not yet sufficiently convincing. The manuscript does not clearly establish what unique opportunity or unmet need is addressed by implementing a visible-wavelength photodetector on an LN waveguide platform. In the broader context of LN photonics, the most established and compelling strength of LN lies in high-speed modulation, typically in the C-band. From the perspective of monolithic integration, a more persuasive direction would therefore be high-speed photodetection in the same spectral window.

In parallel, visible-band photodetection and interconnect-related applications are already being actively pursued on other material platforms, particularly GaN-based technologies, where substantially higher aggregate data capacities (512 Gbit/s) have been demonstrated in multi-channel systems (i.e. Avicena). Against this background, the manuscript does not yet provide a sufficiently compelling system-level or application-level rationale for choosing the WS_2/LN platform for visible-band high-speed detection. For this reason, I believe the work may be more suitable for a specialized photonics journal, such as Optics Express, rather than Nature Communications.

We thank the reviewer for the thoughtful assessment. We fully acknowledge the reviewer's concerns and would like to provide a systematic response addressing the choice of the visible wavelength band, the positioning relative to GaN-based technologies, and the overall conceptual novelty of our work.

1. Response regarding the visible wavelength band and TFLN platform

We appreciate the reviewer's valuable criticism. We fully agree that thin-film lithium niobate (TFLN) has demonstrated mature and outstanding performance in electro-optic modulation at the telecom C-band. However, the visible-band on-chip lithium niobate optoelectronic devices targeted in our work address a fundamentally distinct application, specifically chip-to-chip and on-chip optical interconnects (OIO), which cannot be covered by existing C-band TFLN systems. These two scenarios differ fundamentally in application, transmission distance, and system requirements, as detailed below.

1.1) Distinct application scenarios and transmission distances, filling a technology gap

Visible-band on-chip lithium niobate optoelectronic devices are directed toward a fundamentally distinct on-chip optical input-output (OIO) scenario, which cannot directly adopt the TFLN-based C-band optical module architecture with external modulation mode. Therefore, visible LN photonics is difficult to be evaluated using the criteria developed for long-haul C-band optical communications. As the reviewer correctly pointed out, TFLN has received the most attention for high-speed modulation in the C-band, with devices exhibiting modulation bandwidths exceeding 100 GHz (*Chip* 2022, **1**, 100029, *Nat. Commun* 2026, **17**, 1138). However, current mainstream TFLN modulators and detectors are developed for medium- to long-haul fiber communication (2 km-100 km), operating in the C-band. In contrast, our work targets on-chip and chip-to-chip short-reach visible-light interconnects with transmission distances of only a few centimeters. This is a new optical routing scheme for next-generation high-density heterogeneous integrated chips, which does not compete with conventional long-haul C-band fiber communication but rather represents a new expansion direction for lithium niobate photonics, addressing a clear and unmet engineering need.

Compared with conventional telecom wavelengths, the visible band presents distinct advantages for chip-scale optical interconnects: (i) Stronger optical mode confinement derived from shorter wavelengths allows for sub-wavelength waveguide dimensions, tighter bending radii, and higher routing density, maximizing photonic integration space. (ii) Higher photon energy scales down the voltage and power required for optoelectronic conversion, paving the way for ultra-low-power high-performance

computing. These merits make the visible spectrum a preferred choice for short-reach, high-density on-chip networks.

In recent years, pioneering work from leading groups has demonstrated that TFLN is actively expanding into the visible spectrum, and visible-wavelength LN devices are recognized as an important direction in the field. Benefiting from the broad transparency window of LN (approximately 400-5000 nm) and its excellent electro-optic coefficient, TFLN modulators have also made significant progress in the visible band. For example, visible-light modulators have achieved an electro-optic bandwidth exceeding 35 GHz with a voltage-length product as low as 0.55 V·cm (*Nat. Commun* 2023, **14**, 1496). The Lončar group at Harvard further demonstrated a visible-light TFLN modulator with an edge-coupling loss of <1 dB/facet and an electro-optic bandwidth of >50 GHz (*Nat. Commun* 2024, **15**, 10459). The group of Qiang Lin at the University of Rochester also realized high-speed modulation across the full visible range from 400 to 700 nm, achieving the lowest modulation efficiency figure of merit for lithium niobate traveling-wave modulators (*Optica* 2023, **10**, 125-126). These works directly prove that the TFLN platform is inherently capable of high-performance active optoelectronic devices in the visible band, and that visible-wavelength LN devices represent an actively pursued frontier with significant potential, rather than a niche or unimportant extension.

1.2) Monolithic photodetectors are better suited for on-chip interconnect scenarios

On one hand, developing C-band TFLN detectors following existing approaches would not meet the miniaturization and low-power requirements of high-density short-reach interconnects. On the other hand, the multi-band capability of TFLN modulators enables TFLN-based high-speed photonic integrated chips to support wavelength-division multiplexing and heterogeneous optical interconnects, greatly enhancing bandwidth density and system flexibility. Clearly, a critical bottleneck for a complete TFLN on-chip integrated system is the heterogeneous integration of high-speed photodetectors. This provides a clear entry point for our work: by integrating a WSe_2-based high-speed p-i-n photodetector onto a TFLN waveguide, we demonstrate a complete visible-light photonic link dedicated to short-reach chip interconnects, achieving 10 Gbit/s data reception. This result fills a gap in on-chip integrated visible-light high-speed detection using novel low-dimensional materials and represents an important step toward a more complete on-chip optical interconnect system.

From a longer-term perspective, a complete on-chip transceiver system requires three core modules: light sources, modulators, and detectors. The emission bands of transition metal dichalcogenides (TMDCs) such as WSe_2 naturally fall in the visible to

near-infrared range (1-2 eV), and room-temperature electroluminescence has been widely reported (*Adv. Mater*. 2024, **36**, 2310498, *Nat. Commun*. 2024, **15**, 6799). More importantly, TMDCs can be heterogeneously integrated onto arbitrary substrates including TFLN via van der Waals forces, offering unparalleled process flexibility, low cost, and freedom from lattice mismatch compared with conventional semiconductors. In recent years, wafer-scale growth and integration of two-dimensional(2D) material systems have been advancing rapidly toward compatibility with existing semiconductor processes (*Science* 2025, **390**, 6771, *ACS Nano* 2026, **20**, 2104-2113). The rapid progress in wafer-scale high-quality two-dimensional (2D) materials, combined with their inherent van der Waals integration capability, lays a critical foundation for ultimately realizing truly monolithic fully integrated transceiver chips on the 2D material/TFLN platform.

2. Response regarding comparison with GaN-based technologies

We acknowledge that GaN-based visible photonics has made breakthrough progress in on-chip light sources and emitter arrays. However, the GaN approach and our WSe_2/LN detector address different parts of the optical interconnect link; they are complementary and serve different functions, with no substitution relationship. The key distinctions are as follows.

2.1) The core innovation of Avicena's GaN-based multi-channel system lies in the visible emission side

Regarding the GaN micro-LED multi-channel 512 Gbit/s system mentioned by the reviewer, we have carefully read the relevant reports on Avicena's LightBundle^{TM}eKit. The 512 Gbit/s data rate is achieved through spatial parallel multiplexing of 256 active channels, with a per-channel data rate of ~3.5 Gbit/s. The core innovation is a micro-LED-based ultra-parallel light source array that solves the on-chip light emission problem. In contrast, our work focuses on a monolithically integrated high-speed optical interconnect platform. Avicena's work primarily highlights the advantages of the GaN platform in high-density on-chip light sources, while challenges remain in the monolithic integration of a full-link closed-loop system that includes high-speed optical signal modulation, transmission, and detection. These are precisely the core scientific issues that the present work aims to address.

2.2) Different system-level objectives

We acknowledge the advantage of the Avicena approach in total bandwidth (512 Gbit/s). However, the total bandwidth and single-channel data rate represent different technical dimensions. 256-channel parallel operation relies on spatial multiplexing, while the single-channel rate reflects the intrinsic capability limit of the

material system and device physics. More importantly, the visible-band TFLN platform, leveraging its strong Pockels effect, enables ultra-broadband, low-power electro-optic modulation, and high-density integration, together providing a mature technological foundation for on-chip high-speed interconnect per land. In addition, TFLN platform allows future on-chip optical signal processing functions such as dynamic routing, wavelength-division multiplexing, and signal shaping. In contrast, current GaN micro-LED platforms primarily focus on high-spatial-parallel interconnect lands with SerDes or Mux/DeMux. Thus, the two approaches represent distinct technological pathways, with different emphases and advantages.

2.3) This work focuses on the visible-light receiving and detection side using WSe_2-integrated TFLN

The core contribution of this work is the integration of a WSe_2 lateral homojunction p-i-n photodetector onto a TFLN waveguide, providing a key device building block toward monolithic integration of modulation, transmission, and detection at visible band. Although our design achieves a single-channel data rate of 10 Gbit/s, which is higher than the ~3.5 Gbit/s per-channel rate reported in the Avicena scheme, it should be emphasized that the two studies pursue fundamentally different objectives. Avicena's approach focuses on scaling total system bandwidth through massive spatial parallelization with electrical CMOS SerDes or Mux/DeMux, whereas our work concentrates on enhancing single-channel data transmission and explores the potential for full transceiver integration on the TFLN platform, taking advantage of its broad transparency window in the visible range, its strong Pockels electro-optic effect, and the heterogeneous integration capability of two-dimensional materials.

3. Novelty and Revisions

Therefore, the novelty of our work does not lie in demonstrating an isolated visible photodetector, but in establishing a missing building block toward a monolithically integrated visible TFLN photonic platform for chip-scale optical interconnects. We respectfully believe that this system-level significance distinguishes the present work from incremental device optimization and substantially broadens the application scope of lithium niobate integrated photonics beyond conventional long-haul telecom communications.

We further thank the reviewer for the valuable suggestions. To address the reviewer's concerns and strengthen the motivation and system-level discussion, we have, in the revised manuscript, clearly distinguished the application scenarios and technical requirements of long-haul C-band TFLN devices from those of short-reach visible-light on-chip interconnects, and clearly articulate the unique miniaturization and low-power

advantages of visible-light interconnects. In addition, we have supplemented sufficient simulation results (please refer to our supplementary materials) to further demonstrate the overall characteristics and advantages of the WSe_2/LN detector platform.

In summary, we respectfully request that the reviewer reconsider the conceptual novelty and unique application value of this work. The revised manuscript will fully and clearly address all the concerns raised in this review. In the following sections, we provide point-by-point responses to each specific question.

Specific comments

1. Although the device is presented as a waveguide-integrated photodetector, the influence of the waveguide geometry on detector performance is not adequately addressed. The Methods section only states that a width of 450 nm was chosen to satisfy the single-mode condition, but there is no meaningful discussion of how the optical mode overlaps with the WS_2 layer, nor of how parameters such as etch depth, sidewall angle, and modal confinement affect absorption, responsivity, or bandwidth. Without such analysis, it is difficult to assess whether the reported device performance is close to optimal or simply reflects one unexamined design point.

We thank the reviewer for raising the concern regarding the influence of waveguide geometry on the detector performance. In response, we have performed additional structural characterization and three-dimensional optical simulations to systematically evaluate the impact of key design parameters—including waveguide width, etch depth, sidewall angle, mode overlap, and WSe_2 length—on device performance. The results are summarized in Figures R1.1 and R1.2. A detailed discussion is provided below to demonstrate that the current design is well optimized and approaches the performance limit.

Single-mode propagation and waveguide dimensions: The optical micrograph in Figure R1.1a shows that the WSe_2 coverage region is approximately 400 μm away from both the input and output grating couplers. Since the grating couplers excite the bare TFLN waveguide rather than the WSe_2-covered region, the coupling efficiency is unaffected by the WSe_2 integration. Therefore, stable single-mode operation must be ensured in the regions without WSe_2 coverage (i.e., from the grating coupler to the onset of the WSe_2 layer and along the entire reference waveguide). The ~130 nm etch depth was determined from the grating coupler optimization (Supplementary Figure 1), and the subsequent mode solver simulations were performed based on this fixed value. Figure R1.1(b) presents the simulated relationship between waveguide width and dispersion. Mode solver analysis at 638 nm over a range of ridge widths (0-3μm) confirms that the waveguide supports only the TE_{00} fundamental mode for ridge widths

up to 450nm, with the mode field well confined within the ridge region (Supplementary Figure 1b). These results confirm that the 130 nm etch depth also satisfies the single-mode condition, demonstrating that the grating coupler and waveguide designs are inherently aligned. The grating coupler efficiency (-6.7 dB) was calibrated using an on-chip reference waveguide. Because the WSe_2 flake is positioned far from the grating couplers, the coupling efficiency remains unaffected by WSe_2, ensuring stable and reproducible optical power coupling.

Mode overlap and absorption saturation: Based on the actual waveguide geometry (Figure R1.2a), which features a sidewall angle of approximately 70° and a ~10° tilt transition of the WSe_2 flake covering only the waveguide top, a 3D finite-difference time-domain (FDTD) model was established. Figure R1.2b presents the simulated absorption fraction as a function of WSe_2 length, showing that the absorption saturates above 99% for a length of 12 μm. To ensure complete absorption with a safety margin, the WSe_2 flake length in the actual device was chosen as 16 μm. Therefore, the absorption of incident light by WSe_2 is sufficient and does not limit device performance. Mode analysis reveals that the waveguide supports two dominant *TE* modes at 638 nm, with the optical energy of both modes concentrated in the waveguide top region and exhibiting high overlap integral with the 30-nm-thick WSe_2 layer (Figures R1.2c and R1.2d). This is a key factor enabling the near-unity EQE.

Assessment of intermodal dispersion: Although the reviewer did not explicitly ask about intermodal dispersion, multiple-mode excitation could in principle introduce group-delay differences, potentially limiting the bandwidth for high-speed communication. From the simulation results, the effective index difference between the two *TE*-modes is Δn_{eff} =0.135. With an absorption length of L=16 μm, the resulting time delay between the two modes is $\Delta t = \Delta n_{eff} \times L/c$ =7.2*f*s, which is far smaller than the bit period of 100 ps for a 10 Gbps signal. Therefore, the impact on eye diagram opening and bit error rate is negligible.

Optical power measurement and calibration: Since the WSe_2 photodetector is located at the center of the device waveguide, the optical path before reaching the WSe_2 region includes one grating coupler and half of the waveguide length. To calibrate the total loss of this section, we employed on-chip reference waveguides. Specifically, multiple reference waveguides with identical total length and layout to the device waveguide (without WSe_2 but subjected to the same fabrication processes) were fabricated on the same chip. By measuring their output powers and averaging the results, the insertion loss from the input grating coupler to the midpoint of the waveguide (i.e., half of the total length) was determined to be 6.7 dB. The optical power incident on the WSe_2 region is therefore directly obtained by subtracting this calibrated loss from the

fiber input power. This method, which relies on the measured output of reference waveguides as a benchmark, effectively cancels the influence of grating coupler and waveguide propagation losses without requiring complex corrections for sidewall scattering. It has been widely adopted in the literature (*ACS Photonics* 2022, **9**, 859-867, *Nat. Photonics* 2015, **9**, 247-252. *Adv. Photonics Res.* 2023, **4**, 2300045, *Light-Sci Appl* 2020, **9**, 29)

In summary, the device performance reported in this work represents a carefully optimized design within the current fabrication constraints: the large distance between the grating couplers and the WSe_2 flake ensures stable coupling efficiency; the 450-nm ridge width guarantees single-mode propagation before reaching WSe_2 with the mode field confined to the ridge; the 70° sidewall angle and the ~10° gentle tilt of WSe_2 suppress additional scattering; and the optimized WSe_2 length achieves absorption saturation while rendering intermodal dispersion negligible. The corresponding simulation results and discussions have been incorporated into the revised manuscript for a comprehensive evaluation. We thank the reviewer for the valuable comments.

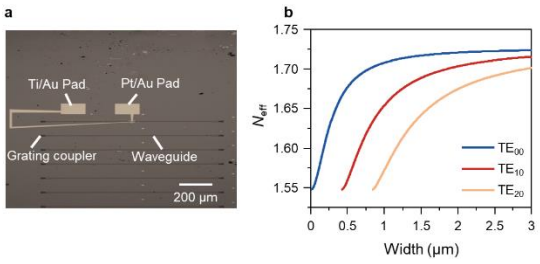


Figure R1.1 Optical image of the WSe_2 waveguide-integrated photodetector. a, Optical image of the device, showing the test structure along with reference waveguide structure. Scale bar: 100μm. b, Simulated dispersion relation with ridge waveguide.

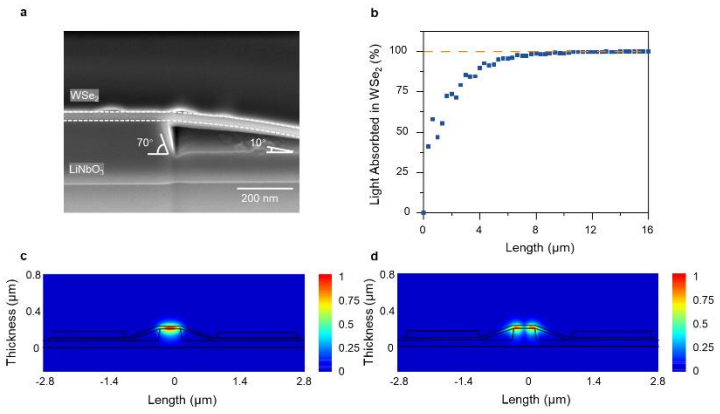


Figure R1.2(Revised Supplementary Fig.3) a, Cross-section scanning electron microscope (SEM) image of WSe_2 on TFLN waveguide. b, Simulated absorbance of WSe_2 as a function of propagation length. c,d, Simulated mode profile between WSe_2 and TFLN waveguide for TE (c) fundamental mode and (d) second-order mode.

2. In the Methods, the authors state that Ar plasma ICP-RIE was used to achieve "smooth, vertical sidewalls." This point requires clarification. In LNOI processing, achieving truly vertical sidewalls is nontrivial. The authors should provide supporting evidence, such as cross-sectional SEM images, sidewall angle measurements, or a more detailed description of the etching conditions and process optimization, to substantiate this claim.

We thank the reviewer for pointing out that the description of "smooth, vertical sidewalls" in the original manuscript was not sufficiently precise. We agree that achieving truly vertical (90°) sidewalls in LNOI using Ar ICP-RIE is challenging. Following the reviewer's suggestion, we performed focused ion beam (FIB) milling on the fabricated waveguides and acquired cross-sectional scanning electron microscopy (SEM) images (Figure R1.2a). Cross-sectional SEM analysis reveals that the actual sidewall angle is approximately 70° with respect to the vertical direction, rather than the ideal 90°. This indicates that the term "vertical" in the original manuscript was not rigorously used. We have therefore revised the manuscript to replace the qualitative description "vertical sidewalls" with the experimentally measured sidewall angle (~70°) and have included the corresponding cross-sectional SEM image.

3. In Fig. 1(a), the LN waveguide is illustrated with nearly vertical sidewalls. This representation is potentially misleading, as perfectly vertical sidewalls are generally difficult to realize in practical LNOI fabrication. The schematic should be revised to more faithfully reflect the actual fabricated geometry, or the authors should provide clear experimental evidence showing that such sidewall profiles were indeed achieved.

We thank the reviewer for the careful observation. Accordingly, we have revised the schematic to more faithfully reflect the actual fabricated geometry, as shown in Figure R1.3.

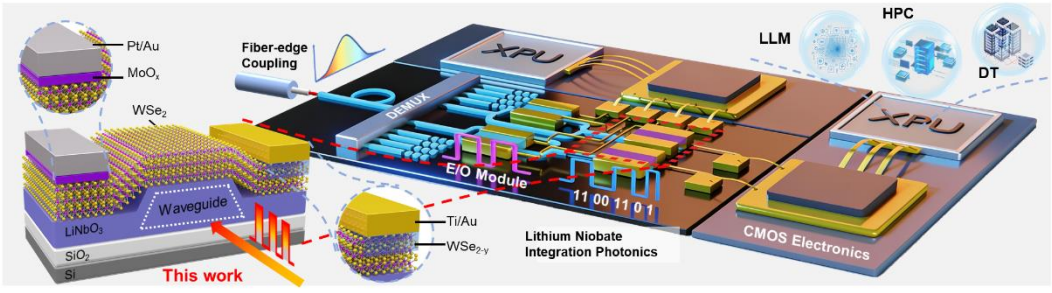


Figure R1.3(Revised Fig.1a) Architecture of a chip-to-chip optical interconnect, where the detector is highlighted and detailed in the inset. (Please refer to the manuscript for detailed captions)

4. It is unclear whether a SiO_2 upper cladding is present above the LN waveguide. This is an important structural detail and should be explicitly stated. If an oxide cladding is used, its thickness should be given, and the authors should explain how sufficient optical-field interaction with the WS_2 absorber is maintained through the cladding. If no upper cladding is present, then the etched LN topography may introduce a substantial step height and sloped sidewall geometry, which raises concerns regarding the mechanical continuity and conformal transfer of the WS_2 film. In that case, the authors should discuss whether cracking, delamination, or strain-induced performance variation could occur, and whether the bending or deformation of the WS_2 layer affects the detector characteristics.

We thank the reviewer for raising this important structural detail. In this work, no SiO_2 upper cladding was deposited on the LN waveguide. The WSe_2 flake was directly transferred onto the bare TFLN waveguide via a dry transfer method to maximize the optical interaction between the waveguide mode and the WSe_2 flake. Therefore, the concern regarding optical-field penetration through an oxide cladding does not apply to the present device. As shown in the revised Supplementary Figure 2(b), the TFLN waveguide has etched trenches approximately 2 μm wide on both sides. The WSe_2 flake straddles the waveguide ridge and the trenches, landing on the unetched platforms outside the trenches. The cross-sectional SEM image (Figure R1.2a) reveals that the WSe_2 does not conformally contact the vertical sidewalls; instead, it bridges the trench with a gentle tilt of about 10° from the waveguide top to the lateral platforms. Furthermore, the suspended WSe_2 region sits closely atop the waveguide top surface, and absorption is predominantly confined to the WSe_2 layer, as confirmed by the 3D FDTD simulations presented in Fig. R1.2. This suspended configuration avoids direct contact with the sharp waveguide edges, thereby minimizing localized strain accumulation. No cracking or delamination was observed on the WSe_2 flake, consistent with observations in other 2D material transfers onto etched waveguides (*Adv. Photonics Res.* 2023, **4**, 2300045).

To evaluate the local strain, Raman line-scan analysis was performed across the WSe_2 flake from the waveguide platform to the trench region (step size: 0.2 μm). As shown in Figure R1.4, the Raman signal intensity in the trench region is significantly lower than that on the waveguide platform. This intensity reduction is attributed to laser defocusing over the etched trench, which reduces the collection efficiency of the Raman scattering signal. Across the entire scan range, however, no measurable shift of the E_{2g} peak is observed relative to the WSe_2 on the unetched platform (variation < 0.05 cm^{-1}), indicating that the WSe_2 flake experiences negligible strain in the bridging region. The suspended structure effectively releases the intrinsic strain introduced during the transfer process.

To further quantify the physical impact of this minor strain, we compare with the work of Shen et al., where a bandgap reduction of approximately 50-75 meV per 1% uniaxial tensile strain in WSe_2 was determined (*ACS Nano* 2016, **10**, 4712-4718). Using the reported Raman strain coefficient of WSe_2 (approximately -1.5 cm^{-1}/% strain, *ACS Appl. Mater. Interfaces* 2024, **16**, 49602-49611), the local strain estimated from the E_{2g} peak shift is less than 0.033%, yielding an estimated bandgap change of less than 2.5 meV. This value is far below the typical bandgap shifts resolvable in the aforementioned-literature. Therefore, this strain level is insufficient to produce any observable impact on the optical absorption or detection performance of the device.

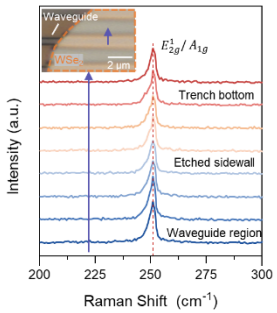


Figure R1.4(Revised Supplementary Fig.4) Raman line-scan of WSe_2 across the waveguide, etched sidewall, and trench bottom, acquired with a 488 nm laser at power of 0.15 mW (100× objective, NA=0.8). The inset is the optical image with the scan path indicated by the purple solid line.

5. The system-level measurement conditions are not sufficiently documented. The manuscript should clearly specify the optical source power, as well as the optical power coupled into the chip. In addition, if the transmitter is based on direct modulation of the optical source, the signal quality before entering the photodetector should also be characterized and reported. For example, the authors should clarify whether the optical waveform, extinction ratio, eye quality, or any relevant transmitter-side performance metrics were measured prior to detection. At present, these key details are missing, making it difficult to assess whether the reported system performance is limited by the detector itself or by the quality of the input optical signal.

We thank the reviewer for pointing out that the system-level measurement conditions require more detailed documentation. In this work, a 638-nm single-mode laser diode (LD) with a launched optical power of tens of milliwatts was used as the light source, and high-speed signals were generated through direct electrical modulation. The output light passed through two polarizers (to ensure *TE* polarization for optimal grating coupler efficiency), a fiber collimator, and a polarization-maintaining fiber, and was then coupled into the on-chip waveguide via a grating coupler. The optical power at the fiber output was measured to be approximately 905.5 μW. After accounting for the 6.7

dB insertion loss of the on-chip grating coupler, the optical power inside the chip was approximately 193.6 μW. The calibration method for the on-chip optical power is described in our response to Comment 1.

For responsivity and EQE measurements, a continuous-wave (DC) light source was used. These measurements do not involve modulation extinction ratio or eye quality, as they depend only on the absorbed optical power and the DC photocurrent. For bandwidth measurements, the laser diode was modulated with a frequency-swept signal, and the amplitude of the detected electrical signal was recorded as a function of modulation frequency. In this configuration, the extinction ratio does not affect the position of the 3 dB roll-off point and therefore does not compromise the accuracy of the bandwidth extraction. We also measured the electro-optic modulation bandwidth of the 638 nm directly modulated laser diode, which was approximately 1.18 GHz (Figure R1.5), comparable to the 1.12 GHz bandwidth of the photodetector. Because the optical transmitter was a commercially available laser diode, its output optical eye diagram was not independently characterized. Instead, the measured modulation bandwidth was used to evaluate its dynamic performance, which is the key parameter governing the present system-level transmission experiment. Therefore, the demonstrated 10 Gbit/s data transmission reflects the combined response of the transmitter and receiver rather than the intrinsic performance of the photodetector alone.

For bandwidth-limited optical communication tests, the photodetector delivers high responsivity and low noise, providing sufficient signal-to-noise ratio (SNR) to mitigate inter-symbol interference arising from the 4-level pulse amplitude modulation (PAM4) signals exceeding the system's 3-dB bandwidth. Combined with piecewise linear equalizer (PLE, see Supplementary Note 6) at the receiver digital signal processing, the distortion is compensated and the reliable symbol recovery is achieved. In addition, the time-domain optical waveform at the receiver side is shown in Figure R1.6, which provides a direct view of the full-sequence received optical signal. To highlight finer details, magnified views of the waveform segments at 4 Gbps, 6 Gbps, and 10 Gbps are also presented. The equalized time-domain optical waveform shows four distinct levels, which proves that we can restore the signal quality very well. Under these conditions, clearly open 10-Gbps eye diagrams were obtained at the receiver (the corresponding eye diagrams are shown in Figures 4f-h). The eye opening and low jitter indirectly indicate that the input signal quality is sufficient for error-free transmission. The successful recovery after equalization indicates that the received signal retains sufficient SNR and linearity for reliable decoding.

We have incorporated the above details and a discussion of the laser diode's bandwidth limitations into the revised Methods section and Supplementary Information. We thank the reviewer for the careful review.

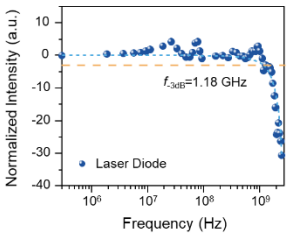


Figure R1.5(Revised Supplementary Fig.15) Broadband frequency response of the 638nm laser diode.

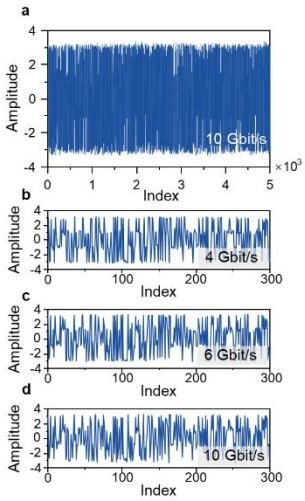


Figure R1.6 Temporal waveforms of the transmitter-side optical signal measured at optical output. **a,** Full pseudo-random binary sequence (PRBS-7) sequence at 10 Gbit/s over 2.3 × 10^4 sampling points. **b-d,** zoomed-in waveforms at 4, 6, and 10 Gbit/s over 0-300 sampling points.

6. In Fig. 4(f)-(h), the eye diagrams are shown without a time scale. This makes it difficult to evaluate the signal quality and to interpret the data quantitatively. A horizontal time axis, or at least the time window per division, should be included.

We thank the reviewer for the careful observation. We have added the time scale to the corresponding figures. For the eye diagrams at 4 Gbps, 6 Gbps, and 10 Gbps, the time scales are 125 ps, 62.5 ps, and 50 ps, respectively.

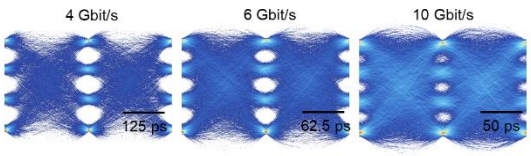


Figure R1.7 (Revised Fig.4f-h) Eye diagrams at data rates of 4 Gbit/s, 6 Gbit/s and 10 Gbit/s.

7. The grating coupler performance is not sufficiently documented. What is the coupling efficiency of the grating couplers, and was it measured experimentally? In Supplementary Fig. 1(f), the grating structure appears to show visible damage or deformation. The origin of this feature should be explained, and the authors should comment on whether it affects insertion loss, coupling stability, or device-to-device reproducibility.

We thank the reviewer for the comments. Regarding the coupling efficiency, we have obtained an average coupling efficiency of 21.7% (-6.7 dB) through experimentally measured using multiple on-chip reference waveguides. The detailed calibration method has been described in our response to Comment 1 and in the Methods section of the revised manuscript, and is therefore not repeated here.

Regarding the morphology of the grating coupler shown in Supplementary Figure 1(f), we note that this SEM image was taken from the output-side grating coupler after WSe_2 transfer and electrode fabrication. The non-ideal features are transfer residues (e.g., residual PMMA) rather than deformation or damage of the grating structure itself. To avoid possible contamination effects during optical measurements, the cleaner grating coupler on the opposite side was used as the input port. Cross-check measurements using multiple reference waveguides, yielded a consistent average coupling efficiency of -6.7 dB, indicating that these transfer residues do not measurably affect the coupling efficiency or its reproducibility. The SEM image of the input-side grating coupler (Figure R1.8) further confirms that the grating structure remains intact, with well-defined features and no evidence of structural damage. We have included the above explanation in the revised Supplementary Information.

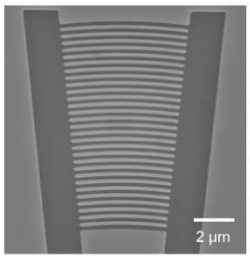


Figure R1.8(Revised Supplementary Fig.1f) SEM image of the TFLN input-side grating coupler.

Reviewer #2 (Remarks to the Author):

Weijia Tang et al. present a WSe_2 photodiode integrated on a TFLN waveguide platform. The PD is targeted for operation in the visible regime and is tested at 638nm. The paper does basic photodetector characterization as well as bandwidth and data communication experiments. Furthermore, they show in detail how the doping and contacts are designed to achieve an effective p-i-n structure. The characterization is done well and extensive, but i suggest the following revisions for the paper to be considered for publication in Nat. Comm. :

We thank the reviewer for the careful review and constructive suggestions. In response to the concerns raised, we have carefully revisited the manuscript structure, revised the specific detectivity evaluation with measured noise spectral density, clarified the EQE analysis with additional experimental and simulation evidence, provided the bandwidth analysis with detailed carrier transport calculations, and refined the SOTA comparison with a clearly defined scope. All the related changes have been incorporated into the revised manuscript and Supplementary Information. We provide point-by-point responses to each comment below.

- 1) There is quite a bit of repetition within the paper due to the structure that was chosen which makes the paper harder to understand. The paper goes from vision, to PD performance, to contact engineering, to FET characterization and back to PD operation. I suggest to the authors reconsider the structure to remove repetitions and make the structure clearer. Personally, I would also refrain from the quite often repeating use of the words like "ultra", "extremely" or "outstanding" which don't add anything to the manuscript in my opinion. Additionally, there are several typos and grammatical errors that should be fixed.

We appreciate the reviewer's valuable comment and agree that the original manuscript structure can be further optimized for better readability and clearer logical progression. In response, we have comprehensively polished the manuscript structure, strengthened section-to-section logical connections, removed redundant descriptions, and revised wording and typos throughout the text.

Structural logic clarified and introduction revised

To strengthen the structural coherence of the manuscript, we have revised the last paragraph of the Introduction to explicitly outline the step-by-step verification logic that underpins the main text. Specifically, the revised paragraph is as follows, with the changes highlighted:

"First, the static optoelectronic characteristics reveal a high-quality p-i-n junction featuring efficient carrier separation, highlighting its potential for high-speed photoresponse. Since the photogenerated carrier drift process is limited by mobility, which is mainly influenced by the contact and the dielectric interfaces, we then separately characterized the two interfaces. At WSe_2/metal contact, the effectiveness of doping is verified through material characterization. Specifically, the n-type doping is achieved through Se vacancy defect engineering, which also enhances fermi level pinning (FLP) to enable n-type Ohmic contact. In addition, a high work function molybdenum oxide (MoO_x) buffer layer is used to break Fermi level pinning and facilitate p-type doping, thereby achieving p-type Ohmic contact. At WSe_2/LN interface, FET characterizations confirm that the high dielectric constant of LN can suppress ionized impurity scattering, thereby greatly enhancing the mobility and solidifying the prerequisite for high-frequency response. Finally, the dynamic optoelectronic performance, including GHz bandwidth and 10 Gbit/s eye diagrams, is systematically demonstrated. Meanwhile, the integration of TFLN waveguide breaks the inherent trade-off between responsivity and response speed, achieving responsivity with an approaching-100% external quantum efficiency (EQE)."

Transitional sentences added

To improve logical clarity, we have added transitional sentences to strengthen the logical flow below sections。

At the end of Section 2, we added:

"Having established effective p- and n-type doping, the carrier transport properties are further determined by fabricating field-effect transistors (FETs) on the same TFLN substrate following the same self-aligned process."

At the end of Section 3, we added:

"The combination of high mobility and low Ohmic contact confirms the performance advantages of the fabricated p-i-n junction diode, which is expected to significantly enhance high-frequency performance, particularly the response speed. These properties lay the foundation for the dynamic characterizations described in the following section."

Section titles revised

We have optimized the section titles to better reflect the logical progression:

1) Static optoelectronic characteristics of TFLN waveguide-integrated p-i-n WSe_2 photodetector

2) Doping and characterization of p- and n-type terminals

3) Electrical characteristics of p-/n-type WSe_2 FETs on TFLN

4) Dynamic optoelectronic characteristics of TFLN waveguide-integrated p-i-n WSe_2 photodetector

Redundant descriptions eliminated

We have removed repetitive descriptions in the main text as follows:

In the concluding paragraph of Section 1, we revised the repeated summary of device performance.

Old:

"The device performance, including…., validates a high-quality p-i-n junction with minimal trap states. This performance arises from the specialized lateral architecture and self-aligned fabrication process, which offers two key advantages."

Revised:

"The device performance validates a high-quality p-i-n junction, stemming from the specialized lateral architecture and self-aligned fabrication process, which provides two key advantages."

In Section 3, we removed the redundant opening sentence:

"The effectiveness of MoO_x buffer layers and Se-vacancy defect engineering for p-type and n-type polarity control has been demonstrated at the material level, respectively."

as its content had already been covered in the preceding section.

Wording and typos corrected

We have also removed all subjective adjectives such as "ultra," "extremely," and "outstanding" throughout the text, and carefully proofread the entire manuscript for grammatical errors and typos, for example, correcting "peek shift" to "peak shift". We thank the reviewer for the valuable suggestions.

- 2) The specific detectivity evaluation: this is done purely by calculation with the dark current and responsivity. It would be much more correct to measure the frequency dependent noise spectral density. Additionally, in Figure 1d the authors report the specific detectivity for varying optical input powers. Here, there would be shot noise occurring from the photocurrent and not just noise from the dark current further reducing specific detectivity. I recommend to the authors to look into: doi.org/10.1038/s41467-023-37635-1 and doi.org/10.1038/s41566-025-01759-1

We thank the reviewer for the valuable suggestions. We agree that an accurate evaluation of detector performance should be based on experimentally measured frequency-dependent noise spectral density, i.e., $D^* = R\sqrt{AB}/\sqrt{i_{noise}^2}$, where A is the device area, R is the responsivity, B is the bandwidth, and i_{noise}^2 represents the total noise current power, including 1/f noise, shot noise, generation-recombination (g-r) noise and thermal noise. This suggests that using only the dark-current shot noise ($i_{noise} \approx i_{shot} = \sqrt{2eI_d}$) for estimation may lead to an overestimation of D^*. The i_{noise}^2 can be obtained from the current noise power density spectra $S(f)$ using the formula of $i_{noise}^2 = \int_0^B S(f)df$. we have conducted the time-resolved dark current with applying a constant V_{ds}=-4V (Figure R2.1a). By taking the Fourier transform of the dark current trace, the $S(f)$ as a function of frequency is obtained as shown in Figure R2.1b. The spectrum exhibits a clear frequency dependence in the low-frequency regime($S \propto 1/f$) with a slope of -1.1, indicating that the dark current noise is dominated by 1/f noise.

Based on the measured noise spectrum, we extracted the value at f=1 Hz from the fitted line, giving $S(f)$= 1.1×10^{-29} A^2/Hz (Figure R2.1b). Substituting this into the formula $D^* = R\sqrt{A\Delta f}/\sqrt{S(f)}$ yields D^*=8.3×10^{10} Jones (@638nm, -4V, 1Hz). This value is close to the previously estimated shot-noise-limited detectivity D^*_{shot} =1.21×10^{11}. To facilitate a fair comparison with previously reported photodetectors, whose detectivity values are generally evaluated without considering photocurrent-induced noise, the performance comparison in the revised Supplementary Figure 16 is still based on detectivity evaluated under dark-state conditions. However, instead of using the previously adopted D^*_{shot}, we now use the D^* calculated from the experimentally measured dark-state noise spectrum, which provides a more rigorous evaluation of the detector performance. Nevertheless, following the reviewer's suggestion, we also measured the time-resolved photocurrent under varying incident optical powers (0.24 nW to 1.1 μW) and obtained the corresponding noise spectral densities via Fourier transform (Figure R2.2). Based on the measured $S(f)$, D^* is recalculated at each optical power level (Figure R2.2h). The results show that at the lowest incident power (0.24 nW), D^* is approximately 1.0×10^{10} (@638nm, -4V, 1Hz), and decreases with increasing incident optical power. Furthermore, we extracted the relationship between $S(f)$ at 1 Hz and the photocurrent (I_{ph}) (Figure R2.2i). The fitting reveals that $S(f) \propto I_{ph}^{1.83}$, which is close to the quadratic dependence expected from the Hooge formula for 1/f noise (exponent of 2). This indicates that at low frequencies (1 Hz), 1/f noise is the dominant noise mechanism (*Adv Phys-X* 2017, **2**, 428-449).

Moreover, since the 1/*f* noise power increases with photocurrent faster than the photocurrent itself (i.e., noise power $\propto I_{ph}^{1.83}$while the signal power $\propto I_{ph}$), the signal-to-noise ratio decreases with increasing optical power, leading to a corresponding decrease in D^*.

In accordance with the reviewer's suggestion, we have added the detailed analysis of the time-resolved current measurements under different optical powers and the corresponding noise spectral densities, along with the recalculated D^* values, in the Supplementary Information. The relevant data and discussions have also been incorporated into the main text as appropriate. We thank the reviewer for the constructive feedback.

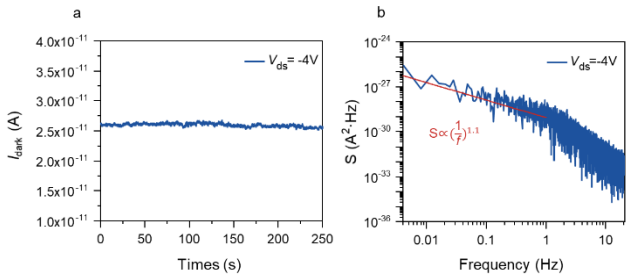


Figure R2.1(Revised Supplementary Fig.7) a, The time-resolved dark current at V_{ds} = -4V. b, The noise power density as a function of frequency at V_{ds} = -4V.

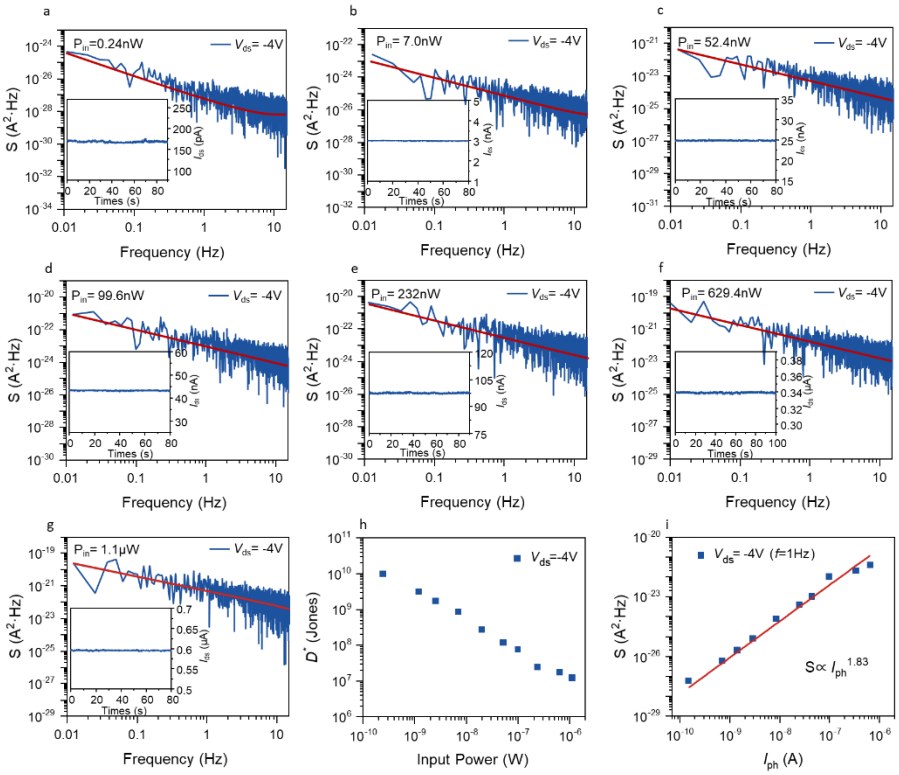


Figure R2.2(Revised Supplementary Fig.8) **a**,**b**,**c**,**d**,**e**,**f**,**g**, The noise power density as a function of frequency at V_{ds} = -4V under incident powers from 0.24 nW to 1.1 μW. The insets show the time-resolved photocurrents at V_{ds} = -4V. **h**, D^* as a function of input optical power. The values of $D^*$are derived from S(f=1HZ) measured at V_{ds} = -4V. **i**, S(f=1Hz) plotted against I_{ph}. The red solid line represents a power-law fit $S(f) \propto I_{ph}^{1.83}$.

- 3) I'm bothered by the EQE analysis and figure. Figure 1e shows data points that clearly exceed 100%, but the authors report a "near-unity EQE" of 98%. Where does this mismatch come from? Also, could the authors show the optical absorption of the WSe_2 PD? Figure 4d shows a reference output that would allow for such an analysis. Are there mode simulations for the TFLN waveguides with the WSe_2 on top that would theoretically match the measured absorption and EQE? Along these lines: could the authors comment on the grating and propagation losses that they compensate? Ultimately, this is an internal responsivity and not a system responsivity that is reported.

We thank the reviewer for pointing out the EQE analysis and the issue of data points exceeding 100% in Figure 1e. We first clarify that the reported EQE is the on-chip EQE, calculated after correcting for the grating-coupling and waveguide propagation losses as described below. In the measurement system, a 638 nm single-mode laser diode was used as the light source. The light passed through two polarizers (to ensure *TE* polarization), a fiber collimator, and a polarization-maintaining fiber. This coupling process is sensitive to environmental vibrations, temperature fluctuations, and mechanical drift of the fiber coupling stage, which introduces fluctuations in the optical power coupled into the fiber. Such fluctuations become more pronounced at low incident optical powers. To address this, we performed repeated measurements across the incident power range and added error bars to Figure 1d and 1e. The error bars clearly show that the scatter is larger at low optical powers. After averaging the data points in the linear-increased photocurrent region, the revised EQE is 95.6%, which has been updated accordingly in the main text. This value reflects a more conservative estimate.

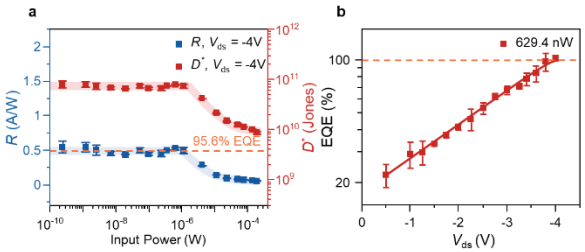


Figure R2.3(Revised Fig. 1d,e) **d,** R and D^* as a function of incident light power with reverse bias voltage of -4 V. The solid lines are guides to the eye. The orange dash line corresponds to R at 100% EQE. **e**, EQE versus V_{ds} measured at a fixed incident power

of 629.4 nW. The red solid line is a guide to the eye. The orange dashed line marks 100% EQE.

The efficient absorption of WSe_2 is demonstrated through both experimental and simulation approaches. First, multiple waveguides with identical length and grating coupler design were fabricated on the same chip. One of them serves as the device waveguide (with WSe_2 transfer), while the others serve as reference waveguides (without WSe_2 but subjected to the same subsequent fabrication processes), as shown in Figure R2.4a. Since the WSe_2 device is located at the center of the waveguide, the optical path from the input grating coupler to the WSe_2 region corresponds to half of the total waveguide length. Using the reference waveguides, the average insertion loss of one grating coupler and half of the waveguide length was calibrated to be 21.3% (6.7 dB). This calibration was performed under relatively high incident optical power to minimize measurement fluctuations, ensuring good repeatability of the calibration results. The optical power incident on the WSe_2 region (labeled in Figure 1d of the main text) is obtained by subtracting this calibrated loss from the fiber input power.

In addition, based on the actual geometry of WSe_2 on the waveguide (Figure R2.4b), with a sidewall angle of approximately 70° and WSe_2 covering only the top with a ~10° tilt transition, three-dimensional FDTD simulations were performed. The simulated absorption fraction as a function of WSe_2 length (Figure R2.4c) shows that absorption saturates above 99% for a length of 12 μm. To ensure complete absorption with a safety margin, a WSe_2 length of 16 μm was chosen in the actual device. In the linear region where the photocurrent scales linearly with incident power (0.24 nW to 1.1 μW), no output optical signal was detected at the device output (residual power below the detection limit of the power meter, 100 pW), confirming that the absorption of incident light by WSe_2 is sufficient. Figures R2.4d and R2.4e show the two dominant *TE* modes supported by the TFLN (thin-film lithium niobate) waveguide after WSe_2 transfer at 638 nm. The mode profiles indicate that the optical energy is concentrated in the top region of the waveguide, exhibiting high overlap with the 30-nm-thick WSe_2 layer. This result theoretically supports the feasibility of achieving near-unity EQE.

Based on the reviewer's concerns regarding optical power coupling and mode simulations, the relevant discussions and results have been incorporated into the revised manuscript. We thank the reviewer for the careful review.

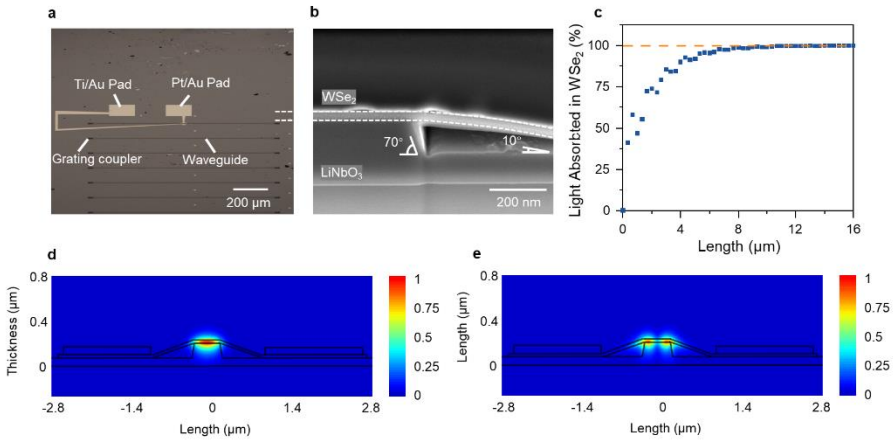


Figure R2.4 a, Optical image of the device, showing the test structure along with reference waveguide structure. b, Cross-section scanning electron microscope (SEM) image of WSe_2 on TFLN waveguide. c, Simulated absorbance of WSe_2 as a function of propagation length. d,e, Simulated mode profile between WSe_2 and TFLN waveguide for TE (d) fundamental mode and (e) second-order mode.

- 4) With what optical input power did the authors perform the data communication experiments? According to Suppl. Figure 4b, the PD saturates and is restricted to the output of a few µA. Is this really enough to achieve these data rates? Or was there a trans impedance amplifier used to boost the signal electrically?

We thank the reviewer for the question. In this work, the optical power before entering the WSe_2 region was 193.5 µW, and the corresponding photocurrent was 11.8 µA. The data rate in high-speed optical communication is determined not by the absolute photocurrent, but by the detector bandwidth and the available signal-to-noise ratio (SNR). Since the signal scales with the photocurrent (or responsivity), whereas the noise is closely related to the dark current, the normalized photocurrent-to-dark-current ratio (NPDR = R/I_{dark}) is widely adopted as a practical figure of merit for evaluating the communication capability of photodetectors. For context, similar photocurrent levels have been reported to achieve Gbit/s data rates: 30 µA (corresponding to NPDR ≈ 85 mW^{-1}) for 3 Gbit/s (*Nat. Photonics* 2015, **9**, 247-252), and 4 µA (NPDR ≈ 600 mW^{-1}) for 1 Gbit/s rate (*ACS Photonics* 2018, **5**, 1846-1852). In contrast, our device exhibits an NDPR of 2.32×10^6 mW^{-1}, which is several orders of magnitude higher than those reported in these previous demonstrations, providing favorable conditions for achieving a high system SNR. To further optimize the system SNR, a piecewise linear equalizer (PLE) was employed at the receiver to compensate for nonlinear distortions in the link through adaptive compensation based on piecewise linear basis functions, effectively improving the received SNR and eye diagram quality.

No transimpedance amplifier (TIA) was used for additional electrical signal amplification. Under conditions where dark current, thermal noise, and shot noise are well controlled, the μA-level photocurrent is sufficient for the real-time oscilloscope (RTO) to sample the signal, and clear time-domain waveforms can be observed on the RTO. In the eye diagram measurements, the electrical signals directly output from the detector were acquired and processed by the PLE, yielding clearly open eye diagrams. This experimentally confirms that the SNR level is sufficient to achieve 10 Gbit/s data transmission without the need for additional electrical amplification.

- 5) On the achieved bandwidth: could the authors illustrate the p-i-n channel of the PD, how long it is and where absorption takes place? What are the resulting transit times and carrier velocities? Are these values realistic for WSe_2? This analysis would provide a more physical interpretation of the achieved bandwidth.

We thank the reviewer for the question. In this work, the i-region width of the p-i-n photodiode is 1.9 μm, and the absorption and separation of photogenerated carriers occur within this region. According to the standard carrier transport model, the carrier drift velocity is $v = \mu E = \mu \frac{V}{d}$, and the transit time is $\tau = \frac{d}{v} = \frac{d^2}{\mu V}$, where μ is the carrier mobility, E is the electric field in the i-region, V is the voltage across the i-region, and d is the i-region width. The carrier mobility in the channel was extracted from FETs. The peak mobility was 382 $cm^2/V{\cdot}s$ and the average mobility was 275 $cm^2/V{\cdot}s$. These mobility values are consistent with the reported maximum mobility of WSe_2 FETs (350 $cm^2/V{\cdot}s$, *Sci. Rep.* 2015, **5**, 8979). Theoretical calculations also show that the maximum mobility of WSe_2 can reach 400 $cm^2/V{\cdot}s$ (*Nano-Micro Lett.* 2017, **9**, 50, *Phys. Rev. B* 2013, **87**, 115418). It should be noted that the above peak mobility was obtained under gate bias (accumulation or inversion conditions), representing the upper limit of the device mobility. Under the operating condition of the photodiode (zero gate bias), the carrier mobility is approximately 128 $cm^2/V{\cdot}s$ (Supplementary Fig. 13), which is the value used for the following transit-time analysis. Although this operating mobility is lower than the peak value, it is still higher than those reported for most transition metal dichalcogenides (TMDC)-based transistors (Figures 3d and 3h in the main text), ensuring a sufficiently high carrier drift velocity.

At a mobility of 128 $cm^2/V{\cdot}s$, with an applied voltage of V =2V, and an i-region width of d=1.9μm, the carrier drift velocity is calculated to be $v=\mu{\cdot}E\approx1.35\times10^6$cm/s. This drift velocity is well below the reported saturation drift velocity of WSe_2 (4.2×10^6cm/s, *Nano. Lett* 2023, **23**, 664-6672), and falls within a reasonable range. The corresponding carrier transit time is $\tau = 141\text{ps}$, giving a theoretical 3-dB bandwidth limit of 3.1 GHz. It should be noted that this value is only the theoretical upper limit of the photodiode

response speed. Due to the voltage division caused by contact resistance and series resistance, the actual transit time will be longer than 141 ps, resulting in a bandwidth lower than the theoretical value.

In the voltage range from 0 to 2 V, the bandwidth increases linearly with voltage, indicating that carrier drift process plays an important role in determining the response speed in this range. Since the drift velocity at 2 V has not yet reached saturation, further increasing the voltage should continue to improve the bandwidth linearly. However, this expectation is inconsistent with the experimental results. When the applied voltage was increased from 2 V to 4 V, no corresponding increase in bandwidth was observed, suggesting that factors other than carrier transit begin to dominate the measured response. Considering that the modulation bandwidth of the laser diode used in this work is only 1.18 GHz (Figure R2.5), the experimentally measured bandwidth of 1.12 GHz is therefore likely limited by the modulation bandwidth of the light source rather than by the intrinsic carrier transit process within the photodetector. Consequently, the intrinsic transit-time-limited bandwidth of the photodetector is expected to exceed the experimentally measured value.

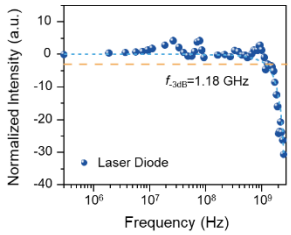


Figure R2.5(Revised Supplementary Fig.15) Broadband frequency response of the 638nm laser diode.

- 6) The authors heavily emphasize visible light communications for "short-reach high-density ultra-low-power-consumption" applications. However, most TFLN development and existing integrated photonic ecosystems operate in the infrared telecommunications bands. Could the authors comment on this? Why the switch to visible PDs?

Thank you for your comments. We fully acknowledge that mainstream TFLN research and industrial ecosystems are well-established for telecom C/O-band long-reach fiber communications (from 2 km to 100 km). We fully agree that telecom-band TFLN will remain the dominant platform for long-reach optical communications. However, our visible-wavelength WSe_2/TFLN photodetector is not intended to replace this ecosystem, but rather to explore a complementary application space. Visible-light integration enables ultra-short-reach, high-density, ultra-low-power chip-to-chip optical interconnects for next-generation AI compute packages, which cannot be efficiently

addressed by infrared telecom-band devices. Benefiting from the broad transparency window of the TFLN platform (approximately 400-5000 nm), we believe that, from the perspective of TFLN-based integrated photonic systems, multi-band compatibility will greatly enhance system flexibility and application versatility. It is precisely because of this advantage that visible-wavelength TFLN photonics has been extensively explored by the academic community in recent years. We provide several specific examples below.

In the visible band, the TFLN platform has already demonstrated a solid foundation for both passive and active devices. Visible-light modulators with an electro-optic bandwidth exceeding 35 GHz and a voltage-length product as low as 0.55 V·cm have been realized (*Nat. Commun* 2023, **14**, 1496), and broadband visible-to-near-infrared modulators with edge coupling loss below 1 dB/facet and bandwidth exceeding 50 GHz have also been demonstrated (*Nat. Commun* 2024, **15**, 10459). These examples demonstrate that visible wavelengths offer the potential for superior integration density and lower power consumption for on-chip links, benefiting from stronger optical mode confinement at shorter wavelengths, as similarly validated by the full-spectrum visible TFLN modulator reported by Lin Qiang's group at the University of Rochester (*Optica* 2023, **10**, 125-126). However, to realize a fully functional TFLN on-chip integrated system, the key bottleneck lies in heterogeneously integrated high-speed photodetectors. This provides a clear entry point for our work: we target this system-level gap by integrating a TMDC-based p-i-n photodetector onto a TFLN waveguide and demonstrate 10 Gbit/s data reception, taking an important step toward a more complete on-chip interconnect system.

In summary, our motivation is not to shift TFLN photonics away from the telecom band, but to extend its capability toward emerging short-reach visible-light interconnects, considering the high-density low-power-consumption requirements of OIO. The broadband transparency and electro-optic functionality of TFLN, together with heterogeneous integration of TMDC materials, provide a complementary technological pathway. We have incorporated the relevant discussion into the revised manuscript, and we hope this better clarifies the system-level value of our work. We thank the reviewer for the valuable comments.

- 7) The SOTA comparison: The comparison on the TMDC mobilities is done well. The SOTA comparison in Table 1 on the hand seems very selective and incomplete. What do the authors want to state with this SOTA table? What is the advantage of the presented work over established technologies which are not listed at all? It seems to me that Si-PIN and GaAs-PIN diodes can easily outperform the here presented work. The conclusion

We appreciate the reviewer for pointing out that the original Table 1 did not clearly define the comparison scope, which may indeed give the impression of a selective benchmark. We clarify the core design logic of SOTA Table as follow.

1. Design objective and defined scope of the SOTA table

The central goal of this work is to demonstrate a high-speed visible-band on-chip photodetector based on heterogeneous integration of a two-dimensional TMDC (WSe_2) on a TFLN waveguide platform. The comparison in the SOTA table is therefore strictly limited to the specific research direction of visible-band integrated photodetectors based on emerging materials (such as 2D materials and perovskites) via heterogeneous integration. The purpose of the table is to benchmark our device against representative emerging-material visible-band integrated photodetectors on similar heterogeneous integration platforms, thereby evaluating the progress of this specific research direction.

We acknowledge that the comparison boundary was not clearly stated in the original manuscript, which may have caused misunderstanding. In the revised version, we will explicitly define the scope of the comparison in the table title, caption, and main text discussion to eliminate any ambiguity.

2. Rationale for not directly comparing with Si-PIN and GaAs-PIN

It has been demonstrated that Si PIN and GaAs PIN photodiodes can achieve excellent bandwidth and responsivity in the visible band (*IEEE Photonics Technol. Lett.* 1996, **8**, 833-835, *J. Lightwave Technol.* 2024, **42**, 8302-8307), and we fully agree with the reviewer that these mature technologies outperform the present device in many performance metrics. Although Si-PIN and GaAs-PIN are technologically mature and high-performance photodetectors, the objective of our work is different. Our study focuses on heterogeneous integration of emerging materials on a TFLN photonic platform rather than on maximizing the performance of discrete photodiodes. The specific reasons are as follows.

1) Distinct material and device pathways: This work focuses on heterogeneous integration of emerging materials (2D materials, perovskites, etc.) to explore epitaxy-free, low-temperature, and simple transfer processes for monolithic integration of multifunctional photonic devices on TFLN. It is not intended to compete directly with mature bulk semiconductor photodiodes, but rather to explore a different integration route for future photonic integrated circuits.

2) Integration compatibility limitations: Unlike atomically thin WSe_2, which can be directly transferred onto pre-patterned TFLN waveguides through van der Waals integration, conventional Si and GaAs photodiodes generally require additional

heterogeneous integration approaches (e.g., wafer bonding or flip-chip assembly) for integration with TFLN photonic circuits.

Therefore, although Si PIN and GaAs PIN represent the current state of the art in mature photodetector technologies, they were not included in Table 1 because the purpose of this table is to compare representative emerging visible-band heterogeneous integrated photodetectors rather than mature semiconductor photodiodes. We appreciate the reviewer's suggestion and agree that the original presentation could give the impression that established technologies were intentionally omitted. We have revised the manuscript to clarify the scope and purpose of the comparison.

Reviewer #3 (Remarks to the Author):

This study reports on the integration of WSe_2 flakes onto thin-film lithium niobate (TFLN) waveguides to realize high-performance integrated photodetectors. The authors demonstrate an impressive external quantum efficiency (EQE) of approximately 98%, alongside a high responsivity and a 3-dB bandwidth of 1.12 GHz, achieved through a lateral p-i-n junction architecture. While the device fabrication and the integration of 2D materials with TFLN waveguides represent significant technical advancement, several aspects regarding the physical interpretation of the efficiency and the device's operational mechanisms require further justification. Therefore, we believe the article requires a major revision to clarify these points before it can be considered for publication in Nature Communications.

We thank the reviewer for the careful review and constructive comments. We have carried out additional experiments and analyses, including bidirectional gate sweeps, noise spectral density measurements, XPS and PL analyses, Raman strain mapping, and 3D FDTD simulations, to address the concerns regarding the physical interpretation of the efficiency and the device operational mechanisms. These supplementary data and discussions have been incorporated into the revised manuscript and Supplementary Information. We provide detailed point-by-point responses to each comment below.

Major comments:

1. The authors claim that the high dielectric constant of the TFLN substrate effectively screens ionized impurity scattering. However, considering the intrinsic ferroelectric nature of lithium niobate, a significant gate hysteresis is expected in the I_{ds}-V_{gs} characteristics. The current Figure 3 lacks bidirectional sweep data, which is essential to evaluate the interfacial quality and the reliability of the extracted mobility. The authors should provide a hysteresis analysis to support their claim of a high-quality junction with minimal trap states.

We thank the reviewer for pointing out the lack of bidirectional sweep data in Figure 3 and for raising the concern regarding hysteresis. In response, we have performed additional experiments and provide the following analysis.

Bidirectional gate voltage sweeps were performed on both n-type and p-type WSe_2 FETs, and the corresponding I_{ds} -V_{gs}, characteristics and extracted mobilities are shown in Figures R3.1 and R3.2. The results show that the hysteresis mainly affects the gate voltage position at which the peak mobility is reached, while the peak mobility itself does not show a clear dependence on the sweep direction. This indicates that the ferroelectric hysteresis has negligible influence on the extracted peak mobility, consistent with previous reports (*ACS Nano* 2025, **19**, 8985-8996, *Ferroelectrics* 2019,

550, 1-11). More importantly, although the FET characterization is used to evaluate the transport properties of WSe_2, the photodiode itself operates without an applied gate bias (V_{gs}=0V). Therefore, any gate-sweep hysteresis observed in the FET characterization does not influence the static or dynamic photodetection measurements.

In addition, we provide multiple lines of evidence for the high-quality junction and low defect density. First, the ideality factor extracted from the I_{ds}-V_{ds} curve is n=1.17 (Figure 1b of the manuscript), which is very close to the ideal value of 1, indicating that the forward current is dominated by minority carrier diffusion and that generation-recombination current due to defect states is negligible. Second, the linear dependence of photocurrent on incident optical power (Supplementary Figure 6) further suggests that defect-related nonlinear recombination or trapping channels are minimal (*Adv. Sci.* 2017, **4**, 1700323).

Furthermore, we measured the time-resolved dark current and obtained the noise power spectrum of the p-i-n device at zero gate bias via Fourier transform. As shown in Figure R3.3, the noise power spectrum exhibits a slope of approximately 1.1 in the low-frequency regime, characteristic of 1/*f* noise. No generation-recombination (g-r) noise plateau, which typically appears as a low-frequency plateau followed by a $1/f^2$ roll-off, was observed over the entire measurement range (*ECS Transactions*, 2016, **75**, 111-120). These observations indicate a relatively low density of electrically active trap states at the WSe_2/TFLN interface, supporting the reliability of the extracted mobility as well as efficient carrier transport in the photodiode. The bidirectional sweep data and the above discussion have been incorporated into the revised manuscript. We thank the reviewer for the careful review.

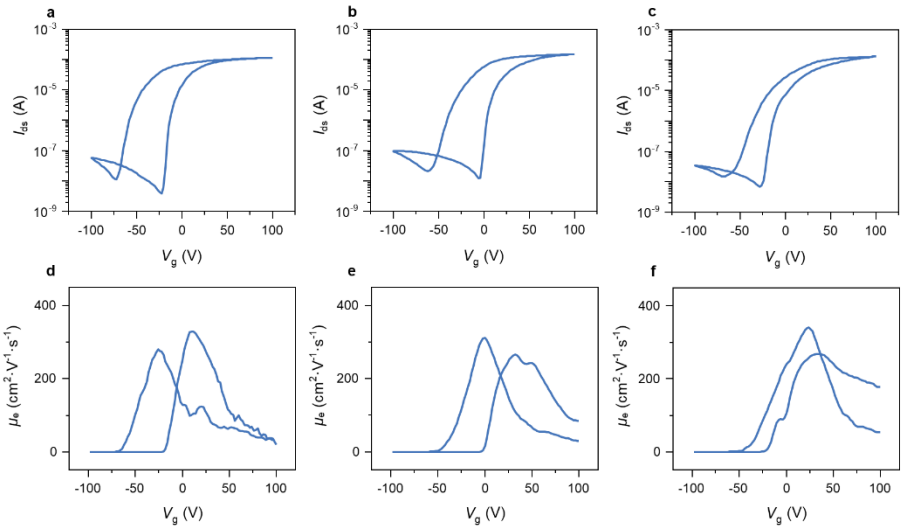


Figure R3.1(Revised Supplementary Fig.13) a,c,e, Transfer characteristics measured at V_{ds} of 5 V from n-type WSe_2 FETs. b,d,f, Carrier mobility as a function of gate voltage for n-type FETs.

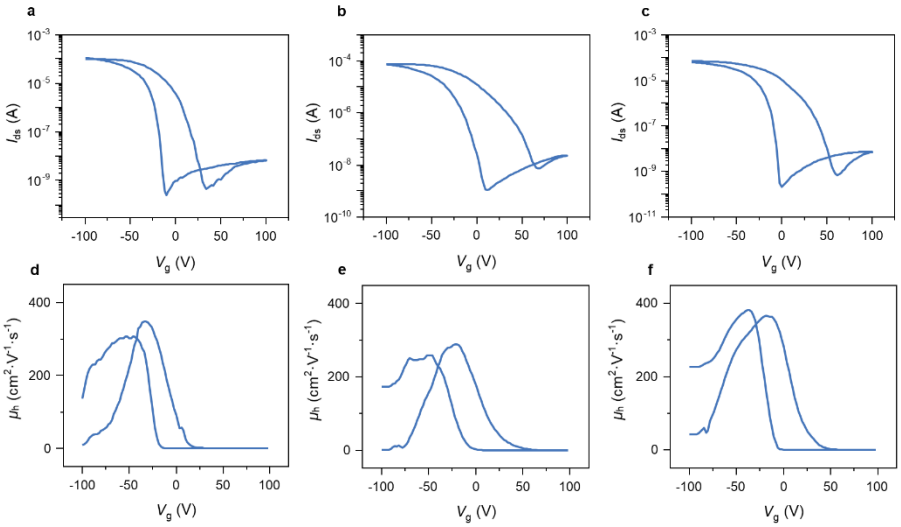


Figure R3.2(Revised Supplementary Fig.13) a,c,e, Transfer characteristics measured at V_{ds} of 5 V from p-type WSe_2 FETs. b,d,f, Carrier mobility as a function of gate voltage for p-type FETs.

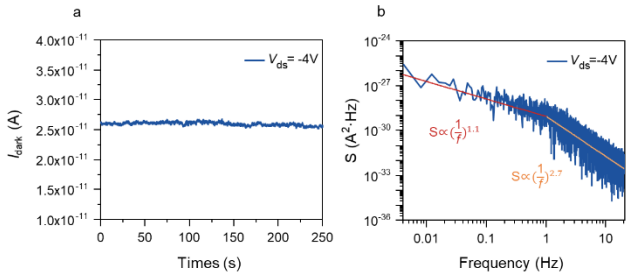


Figure R3.3(Revised Supplementary Fig.7) a, The time-resolved dark current at V_{ds} = -4V. b, The noise power density as a function of frequency at V_{ds} = -4V.

2. The authors report a near-unity EQE of 98%, but the reference point for the input optical power is unclear. It is critical to distinguish whether these values are external (measured at the fiber) or internal (de-embedded for coupling/propagation losses). Please provide a detailed power budget and clarify exactly where the input power was measured to ensure the performance is not overestimated through excessive de-embedding

We thank the reviewer for pointing out the ambiguity regarding the reference point for EQE. We confirm that the reported EQE is an on-chip external quantum efficiency (EQE), rather than a fiber-to-device external efficiency (*Nat. Commun*, 2017, **8**, 14323, *Nat. Commun* 2022, **13**, 6362). The input optical power is referenced to the optical power propagating in the waveguide immediately before the WSe_2 active region, after

calibration of the grating-coupler and propagation losses. This definition is consistent with the de-embedding procedure mentioned by the reviewer, in which the grating-coupler and waveguide propagation losses are removed to determine the optical power delivered to the active region(*Nanoscale Adv.* 2025,**7**,7352-7362, *Adv. Photonics Res.* 2023, **4**, 2300045). The expression is given as: $EQE = \frac{Light\ \ absorbed\ by\ PD}{Light\ inside\ WG} = \frac{I_{ph}/e}{P_{in,WG}/h\nu}$(*Opt. Express* 2020, **28**, 21279, *J. Phys. D: Appl. Phys.* 2023, **56**, 113001, *Phys. Rep.* 2024, **1081**, 1-46). The EQE reported in this work follows this widely adopted definition. The input power is calibrated to the optical power inside the waveguide before reaching the WSe_2 region (labeled as "input power" in the main text), after removing the grating coupler and propagation losses. It should be noted that this efficiency includes possible transition losses (e.g., interface scattering or mode mismatch) during light coupling from the waveguide into WSe_2, so there is no overestimation due to excessive de-embedding.

Specifically, to determine the optical power incident on WSe_2, we used an on-chip reference waveguide with the same total length as the device waveguide (without WSe_2 but subjected to the same fabrication processes, Figure R3.4). By measuring the ratio of the output power of the reference waveguide to the fiber input power, we calibrated the total loss from the fiber input to the output of the reference waveguide. Since the WSe_2 photodetector is located at the center of the device waveguide, the light passes through one grating coupler and half of the waveguide before reaching the WSe_2 region. Therefore, the loss from the fiber input to the WSe_2 region is half of the total loss of the reference waveguide, corresponding to an average efficiency of -6.7 dB. The optical power incident on the WSe_2 region is then directly obtained by subtracting this calibrated loss from the fiber input power, which serves as the input power defined in the EQE expression. This method uses the measured output of the reference waveguide as a benchmark and automatically cancels the influence of grating coupler and waveguide propagation losses, avoiding any excessive de-embedding. Accordingly, the reported responsivity and EQE should be interpreted as on-chip values rather than fiber-to-device values. This power budget approach has been widely adopted in similar works (*ACS Photonics* 2022, **9**, 859-867, *Nat. Photonics* 2015, **9**, 247-252., *Light: Science & Applications* 2020, **9**, 29). We have provided a detailed power budget in the Methods section of the revised manuscript and clearly indicated the reference point for EQE in the figure caption.

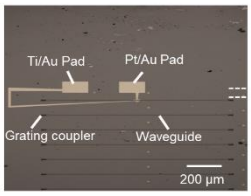


Figure R3.4(Revised Supplementary Fig.2a) Optical image of the device, showing the test structure along with reference waveguide structure.

3. The 98% EQE claim is based on the assumption that all attenuated optical power is absorbed by the WSe_2. However, the 30 nm thick flake on a 450 nm waveguide inevitably causes mode mismatch and scattering losses at the edges. Without distinguishing these non-absorption losses from actual absorption, the reported efficiency is likely misinterpreted. The authors must provide numerical or experimental evidence to quantify scattering losses and prove that they are not being misinterpreted as photocurrent-generating absorption

We thank the reviewer for raising the concern that mode mismatch and scattering losses may lead to an overestimation of EQE. We agree that optical attenuation and optical absorption are not necessarily identical. In this work, the reported EQE is calculated using the optical power inside the waveguide immediately before the WSe_2 region as the reference, following the commonly adopted definition of on-chip EQE for waveguide-integrated photodetectors. Therefore, any optical power lost through mode mismatch or scattering after entering the WSe_2 region does not contribute to the photocurrent and would reduce, rather than artificially increase, the reported EQE. Consequently, scattering losses cannot lead to an overestimation of the reported EQE. Accordingly, the reported responsivity and EQE should be interpreted as on-chip values rather than fiber-to-device values. In fact, the scenario the reviewer is concerned about is more applicable to internal quantum efficiency (IQE), which requires strict subtraction of non-absorption losses after light enters the detector. To fully address the reviewer's concern, we provide an assessment of scattering losses from the following aspects.

First, from the actual device geometry (Figure R3.5a), the WSe_2 flake covers only the top of the waveguide and bridges the trench with a gentle tilt of approximately 10°, extending to the lateral platforms without direct contact with the waveguide sidewalls. This configuration provides a gradual geometric transition at the onset and termination of the WSe_2 coverage, rather than an abrupt cross-sectional change, which minimizes abrupt perturbation of the guided mode, as confirmed by the 3D FDTD simulations presented in Figure R3.5b. Therefore, the "edge mode-mismatch scattering" that the reviewer is concerned about is significantly suppressed by the device structure.

Second, Three-dimensional finite-difference time-domain (FDTD) simulations based on the actual geometry (Figure R3.5b) show that the transmission of the *TE* mode decays by more than 99% after 12 μm of WSe_2 coverage. In the experiment, a WSe_2 length of 16 μm was chosen. Within the linear range where the photocurrent scales linearly with incident optical power, no output optical signal was detected at the device output (residual power below the detection limit of the power meter, 100 pW). This indicates that the vast majority of the total attenuation originates from the strong intrinsic absorption of WSe_2, rather than residual scattering. Similar difficulties in separating absorption and scattering have also been discussed for other waveguide-integrated TMDC photodetectors (*ACS Photonics* 2018, **5**, 1846-1852). Although directly separating absorption and scattering experimentally remains challenging in waveguide-integrated photodetectors, the upper bound of the scattering loss can be rigorously estimated from the measured responsivity. At 638 nm, the theoretical maximum responsivity for 100% IQE is 0.514 A/W, while the average responsivity corresponding to our EQE has reached 0.492 A/W. For $P_{in} = P_{real} + P_{scat}$, where P_{real} is the actually absorbed optical power and P_{scat} represents non-absorption losses such as scattering, we have $\frac{I_{\mathrm{ph}}}{P_{real}+P_{scat}} = 0.492\mathrm{A/W}$ and $\frac{I_{\mathrm{ph}}}{P_{real}} \leq 0.514\mathrm{A/W}$. Therefore, even in the worst-case scenario, scattering losses account for less than approximately 4.9% of the absorbed optical power. We have included the corresponding analysis of scattering losses in the revised manuscript. We thank the reviewer for the careful review.

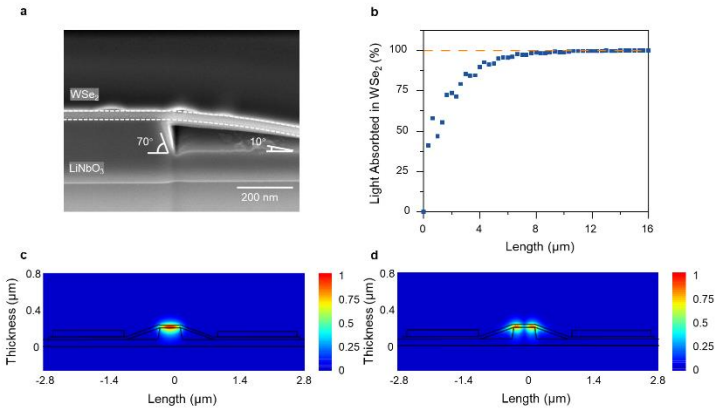


Figure R3.5(Revised Supplementary Fig.3) a, Cross-section scanning electron microscope (SEM) image of WSe_2 on TFLN waveguide. b, Simulated absorbance of WSe_2 as a function of propagation length. c,d, Simulated mode profile between WSe_2 and TFLN waveguide for TE (c) fundamental mode and (d) second-order mode.

4. The contact potential difference (CPD) variations reported for p- and n-type regions are relatively small (35 mV and 44 mV, respectively). Given the high carrier mobilities

and bandwidths claimed, one would expect a more significant Fermi level shift. Furthermore, the n-type region was characterized after Ar plasma exposure, which could introduce surface states or adsorbates in ambient conditions. How can the authors distinguish the intrinsic doping effect from surface-damage-induced potential changes in their KPFM analysis?

We thank the reviewer for the valuable comments.

First, we acknowledge that KPFM measurements are affected by various factors including tip condition, ambient humidity, and surface adsorbates. Therefore, in this work, KPFM is used only for qualitative identification of doping type and relative potential changes. Compared with the pristine region, the MoO_x-treated region shows a decrease in contact potential difference (CPD) (increase in work function), while the Ar-plasma-treated region shows an increase in CPD (decrease in work function). These observations are fully consistent with the characteristic signatures of p-type and n-type doping, respectively, and the qualitative conclusion is reliable. KPFM has been widely adopted in similar studies on surface doping of transition metal dichalcogenides (TMDCs) (*Adv. Mater. Interfaces* 2021, **8**, 2100718, *ACS Nano* 2016, **10**, 6853-6860). Therefore, the absolute CPD values should not be directly interpreted as the junction built-in potential or the magnitude of the Fermi-level shift.

We agree that the relatively small CPD contrast alone cannot account for the high mobility and GHz bandwidth operation observed in our devices. To quantitatively evaluate the Fermi level difference of the p-i-n junction, we extracted a built-in potential of 547 meV from temperature-dependent I_{ds}-V_{ds} measurements (Supplementary Fig. 5), which directly confirms that the actual Fermi-level difference across the p-i-n junction is much larger than the surface CPD contrast measured by KPFM.

Regarding the relatively small CPD changes observed for both p-type and n-type regions, the following points should be noted. For the MoO_x -treated region, the MoO_x layer (3 nm thick) was deposited on the LN substrate before WSe_2 transfer. KPFM only probes the WSe_2 surface rather than the WSe_2/ MoO_x interface. MoO_x, as a strong electron acceptor, extracts electrons from WSe_2 at the interface, but the electrostatic potential sensed by KPFM is attenuated with increasing distance from the buried interface, so the surface CPD change significantly underestimates the true Fermi level shift at the interface. As a result, KPFM mainly probes the electrostatic potential at the WSe_2 surface rather than the buried MoO_x/WSe_2 interface where charge transfer occurs.

For the Ar-plasma-treated n-type region, although KPFM measurements were performed immediately after treatment to minimize environmental exposure, it is still

difficult to completely avoid the influence of adsorbates, which can partially screen the intrinsic potential change and result in a relatively limited CPD magnitude. It is worth noting that in this work, metal electrodes were deposited immediately after doping treatment on each terminal, covering the doped regions beneath the metal, which effectively prevents interference from environmental adsorption.

On the other hand, While Ar plasma treatment may inevitably introduce additional surface defects alongside Se vacancies, distinguishing their individual contributions in KPFM is intrinsically challenging due to its sensitivity to surface electrostatic potential rather than specific defect chemistry. Nevertheless, we provide multiple independent lines of evidence beyond KPFM to confirm that the Ar plasma induced n-type doping originates from Se vacancies rather than other surface states. It has been shown that the formation energy of chalcogen vacancies is much lower than that of metal vacancies, so Ar plasma preferentially generates Se vacancies (*Nano Res.* 2025, **18**, 94907853). DFT calculations have also theoretically confirmed that the introduction of Se vacancies in pristine WSe_2 creates two degenerate defect levels near the conduction band; as the number of Se vacancies increases, more donor defect bands approach the conduction band, forming a defect-rich system (*Adv Mater* 2017, **29**,1606760). Experimentally, XPS analysis has shown that the Se/W peak area ratio decreases after plasma treatment, indicating gradual loss of Se atoms. This directly proves that plasma treatment induces n-type doping through the introduction of Se vacancies (*ACS Nano* 2016, **10**, 6853-6860).

In addition, we performed photoluminescence (PL) spectroscopy analysis on monolayer WSe_2 after Ar plasma treatment (Figure R3.6). The PL spectrum of WSe_2 consists of neutral exciton and trion emission peaks. After Ar plasma treatment, the trion peak fraction increases significantly, consistent with literature reports (*J. Phys. Chem. C* 2017, **121**, 12294-12299). This result further confirms that Ar plasma treatment controllably introduces Se vacancies while preserving the lattice structure, achieving n-type doping. Overall, the combined KPFM, PL, electrical transport, and theoretical results consistently confirm that the observed CPD variation is primarily governed by Se-vacancy-induced donor doping rather than plasma-induced other surface states or adsorbate effects.

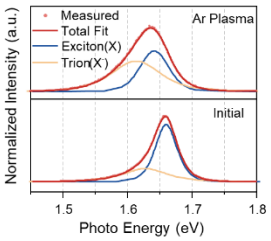


Figure R3.6 Peak fitting of monolayer-WSe_2 PL spectra before and after Ar plasma treatment.

5. The authors emphasize the role of the MoO_x buffer layer by citing a work function of 6.7 eV for MoO_3. While the results are promising, a few points require further clarification to strengthen the manuscript's physical interpretation. First, since the work function of molybdenum oxide is highly sensitive to the O/Mo ratio, it is recommended to provide XPS data to confirm the stoichiometry of the evaporated film. The reported KPFM shift of 35 mV appears relatively small for a 6.7 eV layer, raising questions about whether the actual work function is lower due to oxygen vacancies or if the 3 nm film is discontinuous. Providing high-resolution TEM or AFM data would help confirm the uniformity of this ultra-thin interface and further validate the high device performance

We thank the reviewer for the detailed questions regarding the stoichiometry of the MoO_x buffer layer and the KPFM measurement results.

As mentioned in our response to Comment 4, the MoO_x layer (3 nm thick) was deposited on the LN substrate before WSe_2 transfer, and KPFM measurements were performed on the WSe_2 surface. Therefore, the measured CPD change (35 mV) is relatively small, reflecting the effective work function change at the WSe_2 surface rather than the intrinsic work function of MoO_x.

To quantitatively analyze the composition of the MoO_x film, X-ray photoelectron spectroscopy (XPS) measurements were performed on the 3 nm MoO_x sample. As shown in Figure R3.7(a), the O 1s XPS spectrum reveals two characteristic peaks at 530.9 eV and 531.7 eV, corresponding to the Mo=O bond and the OH^- species, respectively. In Figure R3.7(b), the Mo 3d XPS spectrum shows two peaks at 233.1 eV and 236.2 eV, corresponding to the Mo^{6+} $3d_{5/2}$ and Mo^{6+} $3d_{3/2}$ binding energies, respectively. No peaks corresponding to lower valence states such as Mo^{5+} or Mo^{4+} were observed. The fitting of the Mo 3d and O 1s core levels was performed following the procedure reported in *Appl. Surf. Sci.* 2020, **519**, 1462632. Using the relative sensitivity factors ($RSF_{(O\ 1s)}$: 0.78, $RSF_{(Mo\ 3d)}$: 3.321), the O/Mo atomic ratio was calculated to be 2.76. This value falls within the characteristic range of annealed sub-stoichiometric MoO_x (2.58-2.75), which is consistent with previously reported compositions of thermally evaporated MoO_x films (*Mater. Sci. Semicond. Process.* 2018, **75**, 166-172).

We further characterized the morphology of the same 3 nm MoO_x film. The cross-sectional HR-TEM images (Figures R3.7c and R3.7d) confirm that the MoO_x film is continuously and uniformly spread on the WSe_2 surface. We thank the reviewer for the careful review.

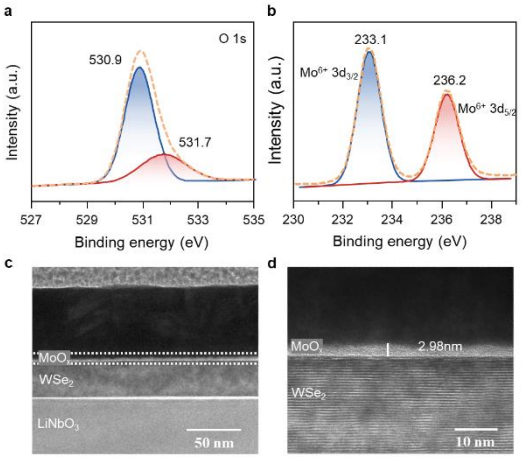


Figure R3.7(Revised Supplementary Fig.10) a,b, XPS wide scan (a)O 1s and (b) Mo 3d of MoO_x. c, Cross-sectional HR-TEM along the p-type region. d, Enlarged-view HR-TEM image of MoO_x and WSe_2 interface.

6. The waveguide has a total thickness of 300 nm with an etch depth of approximately 135 nm, creating a substantial step height of ~165 nm for the 30 nm thick WSe_2 flake to conform over. Given that this step height is more than five times the flake's thickness, significant localized strain is inevitable. How did the authors ensure uniform interfacial contact and account for the resulting strain-induced variations in the device's electronic and optical properties?

We thank the reviewer for raising the concern regarding the potential strain in the WSe_2 flake caused by the step height and its possible influence on device performance.

Regarding the uniformity of interfacial contact, as shown in Figure R3.5a, the WSe_2 flake does not conformally contact the waveguide sidewalls. Instead, it bridges across the etched trench with a gentle tilt of approximately 10° from the ridge top to the lateral platforms, without direct contact with the sidewalls. Therefore, the WSe_2 flake does not need to bend sharply to accommodate the step height, and the actual bending strain is much smaller than the reviewer estimated. This suspended bridging configuration effectively avoids sharp bending at the waveguide edges and ensures a uniform and flat interfacial contact between WSe_2 and waveguide ridge (*Adv. Photonics Res.* 2023, **4**, 2300045).

To quantitatively evaluate the strain and its impact on optoelectronic properties, Raman line-scan analysis was performed across the WSe_2 flake from the waveguide platform to the trench region (Figure R3.8, step size: 0.2 μm). The Raman intensity in the trench region is significantly lower than that on the waveguide platform, which is attributed to laser defocusing over the etched trench, reducing the collection efficiency of the Raman scattering signal. Across the entire scan range, no measurable shift of the E_{2g} peak is

observed relative to the WSe_2 on the unetched platform (variation < 0.05 cm^{-1}), indicating that any possible strain is below the detection limit of Raman spectroscopy and that the suspended structure effectively releases the intrinsic strain introduced during transfer. Referring to the work of Shen et al., where the 1% tensile strain reduces the bandgap of WSe_2 by approximately 50-75 meV (*ACS Nano* 2016, **10**, 4712-4718), and using the reported Raman strain coefficient of WSe_2 (approximately -1.5 cm^{-1}/% strain, *ACS Appl. Mater. Interfaces* 2024, **16**,49602-49611), the local strain estimated from the E_{2g} peak shift is less than 0.033%, yielding an estimated bandgap change of less than 2.5 meV. This value is far below the typical bandgap shifts resolvable reported in that work, further confirming that the strain has no observable impact on the optical absorption or detection performance of the device. The relevant data and discussion have been incorporated into the revised manuscript for the reviewer's reference.

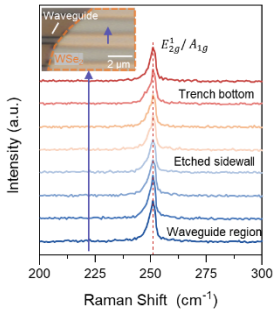


Figure R3.8(Revised Supplementary Fig.4) Raman line-scan of WSe_2 across the waveguide, etched sidewall, and trench bottom, acquired with a 488 nm laser at power of 0.15 mW (100× objective, NA=0.8). The inset is the optical image with the scan path indicated by the purple solid line.

7. The manuscript lacks a clear definition of the intrinsic (i) region's width. This dimension is critical as it determines the trade-off between junction capacitance and carrier transit time. Without specifying this width, it is difficult to evaluate the physical origin of the reported 1.12 GHz bandwidth. The authors should provide the dimensions and explain how the i-region was optimized.

We thank the reviewer for the question regarding the physical origin of the bandwidth. In this work, the i-region width of the p-i-n photodiode is 1.9 μm, which has now been explicitly indicated in the revised manuscript and annotated in the device schematic. The i-region width determines the trade-off between junction capacitance and carrier transit time: a narrower i-region reduces the transit time (beneficial for bandwidth) but increases the junction capacitance (detrimental for *RC* bandwidth); a wider i-region has the opposite effect. Based on this i-region width, we evaluated the influence of *RC* time and carrier transit time on the device bandwidth.

The *RC* time constant (τ_{RC}) is given by $\tau_{RC} = R_{total} \times C$, where R_{total} is the total series resistance and C is the junction capacitance. The junction capacitance C is calculated using the Nabet's formula (*RSC Appl. Interfaces* 2024, **1**, 728-733), a model verified for lateral devices: $C = L(N-1)\varepsilon_0(1+\varepsilon_r)K(k)/K(k')$. L denotes the channel width, N is the number of electrodes, and ε_r is the dielectric constant (WSe_2: ε_r=11.7, *Physica B* 2012, **407**, 4627-4634), $K(k)/K(k')$ serves as a planar fringe electric field correction factor. These yields $C = 2.8$ *f*F. For the R_{total}, the total series resistance R_{total}≈22 kΩ is extracted from the illuminated I_{ds}-V_{ds} curves under the operating bias using the Lambert W-modified Shockley equation (*Nat. Nanotechnol.* 2014, **9**, 262-267). Substituting these values gives an τ_{RC}= 61.6ps, corresponding to an *RC*-limited bandwidth of 2.6 GHz. Since this value is 2.3 times higher than the measured 1.12 GHz, *RC* delay is unlikely to be the bandwidth-limiting mechanism in the present device.

After excluding *RC*, the bandwidth is mainly limited by the carrier transit time. According to the standard carrier transport model, the drift velocity is $v = \mu E = \mu \frac{V}{d}$, and the transit time is $\tau = \frac{d}{v} = \frac{d^2}{\mu V}$, where μ is the carrier mobility, E is the electric field in the i-region, V is the voltage across the i-region, and d*d* is the i-region width. The carrier mobility in the channel was extracted from FETs without gate bias, giving a mobility of 128cm^2/V·s, The calculated transit time is $\tau = 141$ps, corresponding to a theoretical 3-dB bandwidth limit of 3.1 GHz. In the actual device, due to the voltage drop across the contact and series resistances, the effective voltage across the i-region is reduced, and the transit time should be longer than 141 ps, resulting in a bandwidth lower than the theoretical limit. To further identify the bandwidth-limiting factor, we measured the bandwidth as a function of bias voltage. In the 0-2 V range, the bandwidth increases linearly with voltage, indicating that the drift process is dominant. In the 2-4 V range, the bandwidth saturates, while the drift velocity remains significantly below the theoretical upper limit (*Nano. Lett* 2023, **23**, 664-6672), indicating that the drift process is no longer the limiting factor. Notably, considering that the bandwidth of the laser diode used in this work is 1.18 GHz (Figure R3.9), we conclude that the current detector bandwidth is mainly limited by the modulation capability of the light source rather than the drift or *RC* limitations of the detector itself.

Regarding the optimization of the i-region width, the primary consideration is to avoid the insertion loss introduced by the metal electrodes. If the i-region is too narrow, the metal electrodes would be too close to the optical mode region, causing additional insertion loss and affecting the coupling between WSe_2 and the optical mode. Therefore, the selected i-region width of ~2 μm represents a compromise between minimizing the

carrier transit time and avoiding optical loss induced by the nearby metal electrodes. 3D FDTD simulations based on the actual device geometry (Figure R3.5a) and the i-region width (Figure R3.5c,d) show that the metal electrodes do not overlap with the optical mode, so the current i-region width does not introduce significant metal insertion loss. This is an important prerequisite for our optimization. After addressing the light source limitation, if further improvement of the intrinsic detector bandwidth is desired, the i-region width could be appropriately reduced to decrease the carrier transit time without introducing metal loss, while still considering the trade-off with increased junction capacitance. Therefore, in future device designs, the i-region width may be further reduced after optimizing the electrode layout and eliminating the bandwidth limitation imposed by the light source. We have included the i-region width and the analysis of its influence on bandwidth in the revised manuscript.

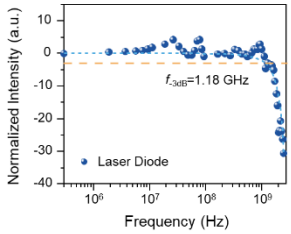


Figure R3.9(Revised Supplementary Fig.15) Broadband frequency response of the 638nm laser diode.

Minor comments:

8. The manuscript contains several typographical errors that should be carefully reviewed. (e. g. “peek shift” in page 7). Additionally, some acronyms are not defined on first mention, for example, FLP or AFM.

We thank the reviewer for the careful proofreading. We have thoroughly reviewed the entire manuscript and corrected the typographical errors.

9. While Ar plasma treatment aligns the Fermi level, the resulting Se vacancies can create trap states that lead to FLP, potentially limiting gate tunability. The authors should clarify if these traps induce persistent photoconductivity or sub-threshold degradation.

We thank the reviewer for raising the concern that Ar plasma treatment may introduce trap states and lead to Fermi level pinning (FLP), potentially causing persistent photoconductivity or sub-threshold degradation. We have carefully analyzed this issue in the context of our device characteristics, as detailed below.

First, unlike conventional silicon-based ion implantation processes, this work employs a self-aligned localized doping and contact scheme. The Ar plasma treatment generates Se vacancies (n-type doping) on the WSe_2 surface in the contact regions defined by a photoresist mask, followed immediately by metal electrode deposition. Since the metal electrodes completely cover the plasma-treated regions, the Se vacancy defects are located beneath the metal/WSe_2 interface and are shielded by the metal layer. During normal device operation, light is coupled from the waveguide into the intrinsic WSe_2 region, and photogeneration occurs only in the i-region. Therefore, the plasma-treated regions do not participate in light absorption, greatly suppressing the influence of defect-induced photogenerated carrier trapping/release processes on photodetection.

Second, while the reviewer is concerned that FLP may limit gate tunability, for the p-i-n photodiode in this work, the device operation is based on the built-in electric field between the p- and n-regions and does not rely on gate voltage control. FETs are used only for extracting auxiliary information such as mobility. In fact, the introduction of FLP in the n-type contact region, where the Fermi level is pinned near the conduction band, is precisely the desired effect for reducing the Schottky barrier at the metal/ WSe_{2-y} contact, facilitating electron injection and collection, and thus enabling Ohmic contact formation. Therefore, the localized Fermi-level pinning beneath the metal contacts is not detrimental to the operation of the present p-i-n photodiode; instead, it facilitates low-resistance electron injection by reducing the Schottky barrier. We agree that FLP beneath the contacts may reduce the gate tunability of the semiconductor underneath the metal electrodes.

Regarding the reviewer's concern about persistent photoconductivity, which typically arises from trap-induced capture of photogenerated minority carriers that causes prolonged conductivity after light illumination ceases, in this work the p-i-n photodiode operates under reverse bias. The electric field rapidly separates and sweeps photogenerated electron-hole pairs out of the i-region, with electrons drifting toward the n-type contact and holes toward the p-type contact, thereby minimizing carrier residence time in the intrinsic region and significantly reducing the probability of trap-assisted carrier capture that could lead to persistent photoconductivity. Experimentally, the high-speed photoresponse and the linear increase of photocurrent with incident optical power (Supplementary Fig. 6) is consistent with the absence of obvious trap-dominated carrier dynamics (*Adv. Sci*. 2017, **4**, 1700323). In addition, no persistent photocurrent or delayed current recovery was observed during repeated photoresponse measurements (Figure R3.10), further indicating that persistent photoconductivity is negligible under our operating conditions.

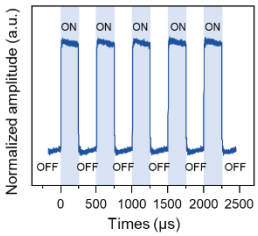


Figure R3.10 Repeatable photocurrent switching behavior of the device at V_{ds}=-4V.

10. The reported dark current values show inconsistencies between the figures. For instance, between Supplementary Figures 3a and 5a. Can you clarify this?

We thank the reviewer for noticing the discrepancy in the dark current values. Supplementary Figures 5a and 9a (Revised; originally Supplementary Fig. 3a and 5a) correspond to two different p-i-n photodiodes fabricated on the same chip using the same fabrication process and identical channel widths. Supplementary Figures 5a and 9a correspond to two different p-i-n photodiodes fabricated on the same chip using the same fabrication process. Although the devices share the same design and fabrication flow, slight device-to-device variations are expected in transferred 2D-material devices due to differences in the transferred WSe_2 flakes size, contact characteristics, and local interface conditions, resulting in variations in the absolute dark current levels. Nevertheless, the rectification ratios and the overall trends of the I_{ds}-V_{ds} curves for the two devices are highly consistent, indicating good junction quality and process reproducibility. The device shown in Supplementary Fig. 9 is the same device used for the optical communication measurements, whereas the temperature-dependent electrical measurements shown in Supplementary Fig. 5 were performed on another representative device. We have clarified the corresponding device information in the revised figure captions to avoid confusion. These device-to-device variations do not affect the conclusions regarding the junction characteristics or the performance analysis presented in this work. We thank the reviewer for the careful review.

Manuscript ID: NCOMMS-26-024174A

Title: Thin-Film Lithium Niobate Waveguide-Integrated p-i-n Tungsten Selenide Photodiodes with 10Gbit/s Data Reception

Point by Point response to the Reviewer's Comments

We sincerely appreciate the reviewers for their precious time and professional insights, which have further strengthened our manuscript. We have carefully considered all the comments and provided detailed point-by-point responses below. For easy identification, all revisions in the main text and Supplementary Information are highlighted in blue. We hope the revised manuscript meets the publication standard.

Reviewer #1 (Remarks to the Author):

The revised manuscript has improved, and the newly added cross-sectional characterization of the heterogeneous waveguide is particularly valuable. Nevertheless, several important issues remain unresolved. At this stage, I am not yet in a position to support publication, and I would like to evaluate a further revised version before making a final recommendation.

We sincerely appreciate the reviewer for the recognition of our revised manuscript and the further valuable suggestions. Below we respond to each point and have made corresponding revisions in both the main text and the Supplementary Information.

1. I remain unconvinced by the authors' response regarding the motivation for operating in the visible wavelength range. Current optical I/O technologies are predominantly developed for the O- or C-band, supported by mature light-source, waveguide, coupling, and packaging ecosystems. By comparison, operation in the visible range presents substantial challenges, particularly in terms of efficient light sources, waveguide loss, sidewall-scattering sensitivity, and system integration. The present manuscript does not yet establish a sufficiently compelling system-level advantage that would justify moving optical I/O into this wavelength regime. The authors should therefore substantially strengthen the motivation in the main text and clearly explain what unique benefit the visible band provides, which application scenario requires it, and whether this benefit outweighs the associated technological challenges. Without such clarification, the broader significance and practical relevance of the proposed direction remain questionable.

We thank the reviewer for raising this question. We agree with the reviewer's assessment that the O- and C-bands benefit from a mature industrial ecosystem for long-

haul optical communications, and we have no intention of replacing these established bands with visible light. However, we would like to clarify that we believe extending optical input/output (OIO) into the visible wavelength regime holds promise and justification for specific ultra-short-reach, high-density interconnect scenarios. Our response is structured as follows:

(1) Evolution of OIO architectures

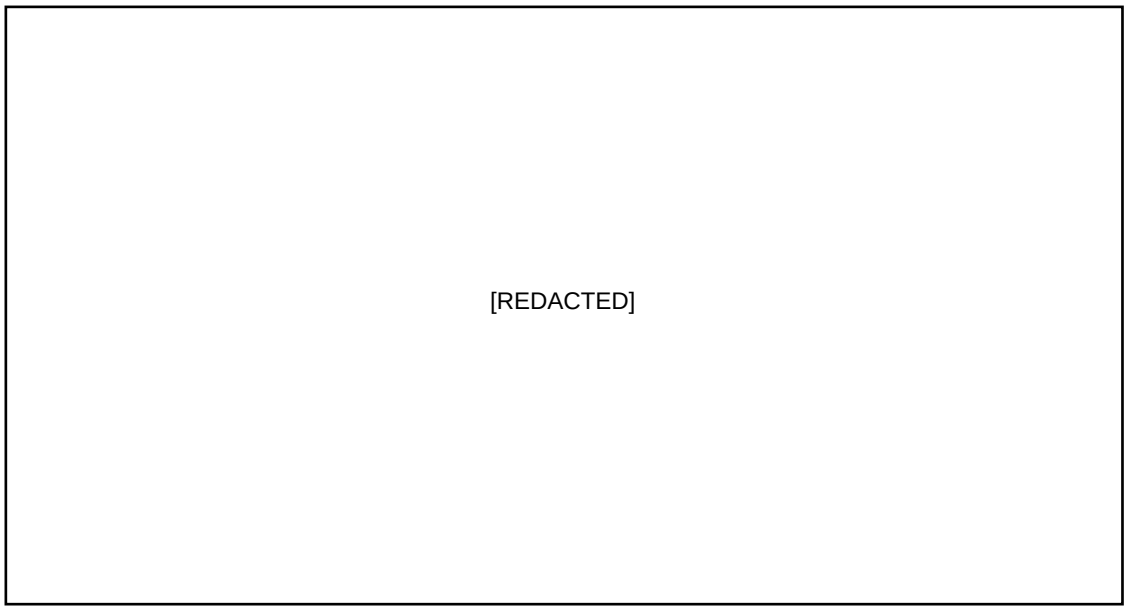


Figure R1.1 The evolution of optics and ASIC interconnects across different architectures. Four successive interconnect architectures illustrate the progressive shortening of interconnection distances between optical engine and ASIC. As optoelectronic conversion modules move increasingly close to the chip, both latency and power consumption are further reduced. (*Nat. Photonics* 2025, **19**, 740-750, *SPIE* 2025, **13990**, 267-272, *IEEE Int. Electron Devices Meeting* 2025, 1-4)

As illustrated in Figure R1.1, optoelectronic conversion modules are moving progressively closer to application-specific integrated circuit (ASIC) dies across four types of architectures: Pluggable, near-package optics (NPO), co-packaged optics (CPO), and OIO. This evolution pushes routing distances down toward the sub-millimeter regime to lower latency and power consumption, yet existing mature supply-chains only cover decimeter- to centimeter-scale applications. Sub-centimeter and millimeter-scale chip-to-chip interconnects still under active development, leaving room for alternative wavelength and material platforms. For emerging chip-scale OIO, multiple wavelength and device platforms are currently being actively explored, including C-band silicon photonics, 850-nm vertical-cavity surface-emitting laser (VCSEL)-based systems, and visible-light approaches. No single wavelength or device platform has become the universal optimal choice. The eventual winner among the

candidate solutions will be determined by their combined performance in areal interconnect density, energy per bit, and thermal stability.

Among the four candidate OIO architectures, as shown in Table R1.1, conventional silicon photonic Mach-Zehnder modulator (MZM) generally require a larger active-device footprint than resonant modulators, which can become a consideration for extremely dense chip-edge integration. Micro-ring modulator (MRM) schemes, despite their smaller footprint, are highly temperature-sensitive—during computation or inference, XPU chip temperatures can exceed 120°C, and temperature drift causes resonant wavelength shift, introducing additional requirements for thermal stabilization and wavelength control. VCSEL-based approaches also face challenges including thermal-management-driven reliability, along with integration density from thermal-electrical crosstalk and array pitch constraints, particularly in long-term high-load operation scenarios in AI data centers. In contrast, visible-light OIO offers more stable performance at high temperatures, which will be further elaborated in the following sections.

Table R1.1 Comparison of candidate OIO architectures. MZM, Mach-Zehnder modulator; MRM, micro-ring modulator; VCSEL, vertical-cavity surface-emitting laser; OIO, optical input/output.

Architecture	Key Advantages	Key Disadvantages
Silicon Photonics MZM	• Large bandwidth, high linearity • CMOS-compatible, mature process	• Large footprint (mm-scale) • Relatively high drive voltage, high power
Silicon Photonics MRM	• Small footprint (μm-scale) • Low drive voltage, low power • Supports wavelength-division multiplexing	• Temperature-sensitive, resonant wavelength drift • Requires wavelength locking and dynamic thermal tuning • Tight fabrication tolerance
VCSEL	• Surface-emitting, forms 2D arrays • Direct modulation, simple structure • Low cost, mature manufacturing	• High-temperature reliability concerns • Density limited by thermal-electrical crosstalk and array pitch
Visible-Light OIO	• High areal density (small bending radius) • High thermal tolerance, reduced cooling overhead • Van der Waals integration with platforms • Low energy-per-bit via dense low-speed parallel channels	• Immature source/waveguide/coupling ecosystem • Early-stage packaging and standardization

(2) System-level advantages of visible-light OIO

For sub-centimeter and millimeter-scale chip-to-chip fanout, wavelength compatibility with long-haul fiber communication is no longer a fundamental system constraint. In this regime, the smaller optical wavelength can enable tighter optical confinement and potentially higher areal integration density, while the low thermo-optic coefficient of TFLN provides excellent thermal stability. These characteristics make visible-light TFLN particularly attractive for dense, short-reach OIO.

In terms of integration density, the visible-light approach enables higher-density array integration. On the TFLN platform, waveguides operating at visible wavelengths can achieve a smaller bending radius compared with the infrared counterparts. This physical advantage provides the foundation for substantially improving the areal interconnect density within the same chip footprint, offering a system-level solution to the bottleneck of XPU edge fan-out density.

In terms of thermal stability, unlike silicon photonic MRM schemes that rely on resonant effects and are highly temperature-sensitive, the thermal robustness of the proposed platform primarily arises from the low thermo-optic coefficient of TFLN and the non-resonant operation of the WSe_2 p-i-n photodetector, which can provide improved tolerance to thermal fluctuations compared with resonance-dependent devices.

High-density parallelization may enable lower per-lane data rates while maintaining high aggregate bandwidth, which can potentially alleviate the electrical and thermal burden associated with ultra-high-speed single lanes. In addition, high-density, low-speed array aggregation can achieve high aggregate bandwidth while avoiding the power consumption and thermal dissipation burden associated with single-lane high-speed data rates.

(3) Visible-light challenges in efficient light sources, waveguide loss, sidewall-scattering sensitivity, and system integration

Regarding efficient light sources. The essence of "efficiency" lies in the wall-plug efficiency (WPE), which, for on-chip integrated light sources, directly determines the system's power consumption and thermal management burden. In the infrared band, commercial InP-based DFB laser diodes can achieve a chip-level WPE of up to 30% (Semiconductor Today 2025). In the visible band, commercial 638-nm AlGaInP laser diodes can achieve a WPE of 38%-45% (Oclaro launches 700mW 638nm red laser diode). The two are on the same order of magnitude. This demonstrates the feasibility of integrating this material system with photonic platforms.

Regarding waveguide propagation loss and sidewall-scattering sensitivity. For the millimeter- to sub-centimeter propagation distances considered here, such propagation loss contributes only a minor fraction of the overall optical power budget. In the infrared band (1550 nm), propagation losses typically range from 0.15-0.6dB/cm, with the lowest reported value reaching 0.027 dB/cm (*Nat. Commun* 2020, **11**, 3911, *Nature* 2018, **562**, 101-104, *Opt. Lett* 2015, **40**, 3013-3016, *Nanomaterials* 2018, **8**, 910) In the visible band (637 nm), a propagation loss of 0.06 dB/cm has been demonstrated (*Optica* 2019, **6**, 380-384), placing the two in the same order of magnitude. That work confirmed a sidewall roughness of only 0.7 nm RMS by AFM characterization and employed a shallow-etched rib waveguide structure to reduce the overlap between the optical mode and the sidewalls, experimentally validating the feasibility of low-loss visible-light waveguides. More recently, Khalatpour et al. (*Adv. Optical Mater.* 2026, **14**, e02355) further showed that through optimized waveguide design and etching processes, visible-light waveguide losses can be expected to approach the optimal levels achieved in the infrared band. For applications such as OIO with millimeter- or sub-millimeter-scale transmission distances, the waveguide loss is typically much lower than 0.1 dB.

Regarding system integration. At the system level, the proposed visible-light architecture can leverage established semiconductor fabrication and TFLN photonic components, while the remaining challenges mainly concern efficient optical coupling, visible-light source integration, and packaging. These challenges are primarily engineering challenges, rather than fundamental limitations of the proposed wavelength regime, particularly for millimeter- to sub-centimeter links where propagation loss and compatibility with long-haul fiber wavelengths are less critical.

(4) Regarding whether the advantages outweigh the technical challenges

Based on the above analysis, we believe that for millimeter- to sub-centimeter-scale chip-to-chip interconnects, the potential benefits may justify further exploration despite the additional engineering challenges. Although the issues raised by the reviewer, such as efficient light sources, waveguide loss, sidewall scattering, and system integration, are real, recent research progress has demonstrated that these challenges are being progressively addressed. Meanwhile, the advantages offered by visible light in terms of integration density and thermal stability can provide benefits that are difficult to simultaneously achieve in conventional telecom-band platforms. Therefore, when the application scenario is clearly defined as ultra-short-reach, high-density interconnects, the gains brought by visible light are worth the engineering overhead. A general rule of thumb in optical interconnect ecosystems is that the shorter the distance, the more transparent the protocol, and the less concern there is about

wavelength compatibility and other system-level factors. We believe that as AI clusters continue to scale and the bandwidth density requirements of XPUs keep increasing, visible-light wavelengths will gain increasing acceptance and attention in ultra-short-reach optical interconnects. We therefore position this work not as a replacement for mature telecom-band optical interconnects, but as an enabling step toward an alternative wavelength platform for emerging ultra-short-reach, high-density OIO.

Based on the above discussion, we have made the following additions and revisions in the Introduction of the main text:

"Integrated photonics is a fast-developing technology that acts as a key enabler for addressing high-speed, low-power chip-scale optical interconnection challenges [1, 2]. As the gap between computing performance and interconnect bandwidth continues to widen, optical interconnect technology is undergoing a profound evolution, with the central trend being the progressive movement of photonic conversion modules closer to the chip to reduce latency and power consumption [3]. This trend is driving the transition from pluggable optical modules toward chip-scale optical input/output (OIO) architectures (Supplementary Figure 1) [4]. For emerging chip-scale OIO, multiple device platforms are currently being actively explored, including silicon photonics, thin-film lithium (TFLN) photonics and vertical-cavity surface-emitting laser (VCSEL)-based systems [1, 5, 6]. Among these, the TFLN platform offers strong light confinement with large waveguide refractive index contrast, providing significant advantages for high-bandwidth, high-density and low-power OIO applications.

While most TFLN photonic devices have been developed for telecom-band optical communications, the broad transparency window of lithium niobate (LN, approximately 400-5000nm) also enables operation across the visible spectrum. Recent advances in visible-light TFLN modulators have demonstrated the potential of LN for visible integrated photonic circuits [7-9]. Benefiting from the smaller bending radius at visible wavelengths and the low thermo-optic coefficient of TFLN, visible-light TFLN OIO can offer advantages in areal integration density and thermal stability compared with silicon-based Mach-Zehnder/micro-ring modulator and VCSEL-based systems (Supplementary Table 1). These features make visible wavelengths particularly attractive for sub-centimeter and millimeter-scale chip-to-chip fanout interconnects, as depicted in Figure 1a, targeting emerging computing applications such as large language models (LLMs), high-performance computing (HPC), and digital twins (DTs)."

2. Figure 1a remains confusing and requires substantial revision. First, the label "Fiber Edge-Coupling" is inaccurate; "fiber-to-chip edge coupling" or simply "edge coupling" would be more appropriate. Second, the signal-flow direction and the ordering of the

components are unclear. If the intended signal direction is from the fiber to the chip, it is not evident why the signal first passes through the demultiplexer, particularly because the input is illustrated as broadband. The functions of several other elements are also not identified. For example, the components electrically connected to the XPU through wire bonds should be explicitly labelled. Although this is intended as a conceptual illustration, it should still present a physically consistent architecture with unambiguous signal directions, component functions, and interconnections. In its current form, the figure raises more questions than it resolves.

We sincerely thank the reviewer for pointing out these shortcomings in Figure 1a. We fully agree with the reviewer's assessment—even for a conceptual architecture illustration, the signal flow should be physically self-consistent, and the functions of the components as well as their interconnections should be clearly indicated. In response to the reviewer's comments, we have made the following revisions, as shown in Figure R1.2:

(1) The imprecise terminology has been corrected. "Fiber Edge-Coupling" has been changed to the more concise and accurate "Edge coupling"

(2) The optical signal generation, transmission, and modulation process has been clarified. The optical fiber provides a single-wavelength continuous-wave (CW) pump light to the chip, which itself carries no data information. After being coupled into the photonic chip via edge coupling, the pump light first enters an on-chip micro-ring resonator, where a broadband optical frequency comb is generated through the Kerr nonlinear process. The generated multi-wavelength optical frequency comb is then fed into a demultiplexer (DEMUX), separated into multiple independent wavelength channels, and routed as parallel optical carriers to multiple electro-optic modulators. The visible light in each channel passes through an electro-optic modulator (labeled "E/O Modulation, controlled by XPU"), which is driven by XPU1 via on-chip CMOS Electronics. The data-encoded optical signals are received by the on-chip photodetector (this work), and the converted electrical signals are transmitted to XPU2 via CMOS Electronics.

To further clarify this signal flow, the optical field states at different positions in the revised Figure 1a have been annotated. "Broadband CW light (carrier only)" is labeled along the path from the micro-ring output to the DEMUX, indicating the pure carrier before data encoding, "Modulated optical signal (data encoded)" is labeled after the modulator. These annotations clearly illustrate the physical process whereby "the external fiber provides only the pump light, and the data is encoded on-chip by XPU via CMOS-driven electro-optic modulators." Finally, "HPC/LLM/DT" in the upper right corner is marked with a dashed box labeled "Target Applications" to clarify that

these represent target application scenarios rather than functional components in the optical signal chain. The revised Figure 1a has been updated in the revised manuscript.

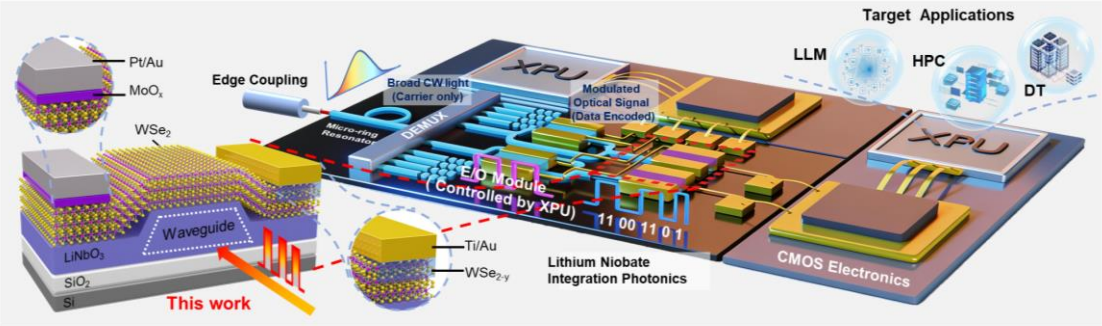


Figure R1.2(Revised Figure 1a) Architecture of a chip-to-chip optical interconnect.

3. There appears to be an inconsistency between the simulation geometry in Extended Data Fig. 1b and the experimentally characterized waveguide cross-section. The simulated waveguide still appears to have vertical sidewalls, whereas the newly provided characterization indicates a sidewall angle of approximately 70°. The authors should clarify how the sidewall angle is defined and whether the experimentally measured geometry was incorporated into the simulation. If not, the simulations should be repeated using the realistic waveguide cross-section, and the authors should discuss whether the nonvertical sidewalls affect the modal properties, confinement, coupling behavior, or other key performance metrics reported in the manuscript.

We thank the reviewer for pointing out this inconsistency. We agree that the simulated geometry should accurately reflect the experimentally characterized waveguide cross-section. In the revised manuscript, we have incorporated the experimentally measured sidewall angle of ~70° into the relevant simulations. Here, the sidewall angle is defined as the angle between the etched sidewall and the horizontal substrate surface. For the grating coupler, the grating teeth are therefore modeled with a trapezoidal cross-section rather than ideal vertical sidewalls.

The revised simulation results show that at a wavelength of 638 nm, the grating coupler enables efficient coupling of light from the optical fiber into the waveguide (Figure R1.3a). Mode analysis indicates that the structure still supports TE_{00} fundamental mode transmission, satisfying the single-mode condition with good optical confinement (Figure R1.3b). In addition, the optimized grating coupler has a pitch of approximately 440 nm and a duty cycle of approximately 0.58 (Figure R1.3c-e), corresponding to a top tooth width of 185 nm and a top trench width of 255 nm. SEM imaging was performed on the fabricated device (Figure R1.3f-h), and the measured top tooth width is 191 nm, which is in good agreement with the simulation results.

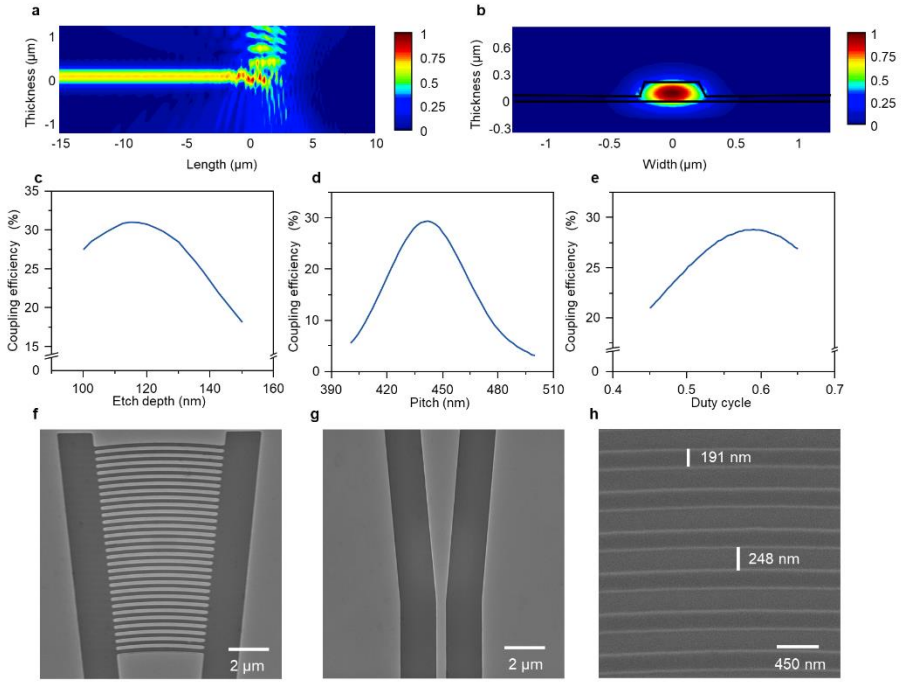


Figure R1.3(Revised Supplementary Figure 2) **a,b,** Calculated electric field profiles of the *TE* mode (**a**) in and (**b**) along the TFLN waveguide. **c,d,e,** Simulated coupling efficiency of the TFLN grating coupler at 638 nm wavelength: (**c**) Etch depth dependence, (**d**) Pitch variation, and (**e**) Duty cycle optimization. **f,** SEM image of the TFLN waveguide with grating coupler. **g,** SEM image of waveguide tapper and guided wave region, highlighting the smooth transition and uniform waveguide geometry. **h,** High-magnification SEM image of the grating coupler.

We further considered the effect of the nonvertical sidewalls on the grating coupling behavior. Compared with the idealized vertical-sidewall geometry, the ~70° sidewalls can reduce the coupling efficiency. However, the coupling efficiency was experimentally calibrated using an on-chip reference waveguide, yielding a measured value of -6.7 dB. Because the grating coupler is spatially separated from the WSe_2-covered detector region, its coupling loss does not directly determine the intrinsic detector response or bandwidth. The calibrated coupling loss was taken into account when evaluating the optical power delivered to the detector.

Based on the experimentally measured sidewall angle of ~70° and the actual WSe_2-waveguide configuration, simulations of optical absorption and mode distribution in the waveguide with WSe_2 were performed. The results show that WSe_2 achieves saturated absorption for lengths exceeding 12 μm, and the optical mode is predominantly confined within the WSe_2 layer. This strong optical mode overlap combined with near-complete optical absorption provides the physical basis for the high EQE observed

experimentally (see Supplementary Figure 4 and Supplementary Note 1 for details). The grating coupler simulations added in this revision do not alter the above optical absorption analysis or the conclusions regarding device performance. The corresponding simulation results for the grating coupler have been updated in the Supplementary Information.

Reviewer #2 (Remarks to the Author):

I thank Weiijia Tang et al. for the revisions. I'll go through them by each point:

-1) The readability of the paper is much improved with the changes made to the manuscript. -2) The new discussion on the specific detectivity is good.

We sincerely thank the reviewer for the careful point-by-point review and valuable comments on our manuscript.

-3) The repeated measurements, inclusion of error bars and averaging to evaluate the EQE is a great improvement. The value of 95.6% is already more realistic than the value of 98%, considering that the absorption is not 100%. However, the new Figure 1e still shows EQE above 100% without error bars on the data point that go below 100%. One issue could be deviations between the calibration grating and channel and the actual device leading to further uncertainty. Or could the authors bias the device higher to show that their EQE flattens out?

We thank the reviewer for acknowledging the introduction of repeated measurements, error bars, and averaging in the EQE evaluation.

It should be noted that each data point in the original Figure 1d and 1e (corresponding to Figure R2.1a and 2.1c) already includes error bars representing standard deviations. However, due to the relatively large scale of the vertical axes, the error bars for some data points are not clearly visible. To more clearly demonstrate this to the reviewer, Figure R2.1b and d have been added as enlarged views of the responsivity in the low-power region (Figure R2.1a) and the EQE in the high-bias region (Figure R2.1c). In these magnified plots, the error bars for each data point are clearly visible, allowing a more straightforward assessment of the measurement uncertainty near 100% EQE.

In addition, following the reviewer's suggestion to apply a higher bias voltage to observe whether the EQE flattens out, additional EQE measurements were performed at an extended bias of -5 V. The results have been included in the revised Figure R2.1c and 2.1d. The measured EQE shows no further increase from -4 V to -5 V, further confirming that the carrier collection efficiency of the device is already near saturation at this bias. Therefore, the 95.6% EQE value does not arise from an underestimation due to insufficient carrier collection, but rather represents a value close to the actual upper limit of the external quantum efficiency of the device.

We agree with the reviewer's assessment that coupling differences between the calibration grating and the actual device channel may exist. The uncertainty in the EQE measurement shown in Figure 1e may arise not only from statistical fluctuations in repeated measurements, but also from differences in coupling efficiency between the

calibration grating and the device grating, as well as the power measurement accuracy of the optical power meter. These factors may contribute to the systematic uncertainty of the absolute EQE measurement. Therefore, in the revised manuscript, the single measurement result is no longer treated as a precise absolute value; instead, repeated measurements and error bars are used to reflect the measurement uncertainty. Considering the repeated measurement results, the saturation behavior of EQE at -5 V, and the fact that the device itself does not exhibit 100% optical absorption, the 95.6% EQE is considered a more reliable and physically reasonable result.

We have incorporated the new EQE data points (extending the range to -5 V) into the revised Figure 1e, and updated the main text accordingly. These modifications are believed to adequately address the reviewer's concerns.

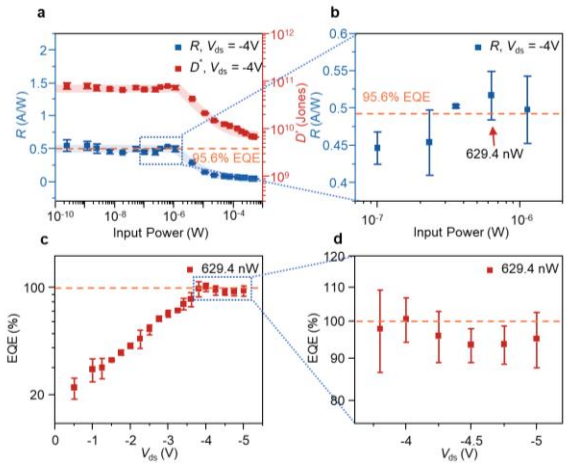


Figure R2.1 a, *R* and D^* as a function of incident light power with reverse bias voltage of -4 V. The orange dash line corresponds to *R* at 95.6% EQE. **b**, Enlarged view of *R* in the incident power range from 52.4nW-1.1μW. **c**, EQE versus V_{ds} measured at a fixed incident power of 629.4 nW. The orange dashed line marks 100% EQE. **d**, Enlarged view of EQE data in bias range from -4V to -5V.

-4) On the photocurrent: 11.8 uA at an optical input power of 193.5 uW means the device is operated at the maximum that is reported in Figure 1c and is the saturated output of device. This corresponds to a responsivity of ~61 mA/W which means the device is operated at a point where it has diminished efficiency. However, it seems this was necessary to perform the data measurements. This seems to be a major drawback of the device that only ~10 uA of photocurrent can be produced and the maximum achievable photocurrent should be discussed in the manuscript.

On the SNR: I'm well aware of SNR and NPDR. The limiting factor when doing data transmission with such low RF powers can be the scope. E.g., a 6 GHz Keysight

oscilloscope has on the highest sensitivity an input noise floor of ~ -61 dBm (or ~4 uA rms). This noise floor is typically higher if you go to higher bandwidth oscilloscopes. Just as a rough estimation, with ~10 uA the ideal RF power is P = I^2 * 50 Ohm = 5 nW = -53 dBm, so not even an order of magnitude above the noise floor.

The papers the authors cite further point towards my question regarding an amplifier:

"For context, similar photocurrent levels have been reported to achieve Gbit/s data rates: 30 μA (corresponding to NPDR ≈ 85 mW^{-1}) for 3 Gbit/s (Nat. Photonics 2015, 9, 247-252) …"

-> "… of the measurement instruments (including the current amplifier and bias tee)"

" …4 μA (NPDR ≈ 600 mW-1) for 1 Gbit/s rate (ACS Photonics 2018, 5, 1846-1852)"

-> "The generated electrical signal was read out with a microwave probe, amplified with a 50 Ω RF power amplifier, and finally detected by a digital storage oscilloscope."

Depending on the measurement equipment, I see that that an amplifier free operation is possible to achieve data transmission. Nevertheless, I would disagree with the statement that "the μA level photocurrent is sufficient for the real time oscilloscope (RTO) to sample the signal". Or could the authors show data comm where the device is not operated in a regime where the responsivity is diminished? This would be at around ~1 μA output current?

Again, I would ask the authors to include a complete experimental description of the data measurements, i.e., include the operation condition on the device, especially the optical input power. Additionally, as mentioned above, the authors should discuss the maximum photocurrent output which is a major drawback of the device.

We thank the reviewer for raising these questions. The reviewer's main concerns focus on three aspects: (1) whether the maximum photocurrent of ~11.8 μA represents a major drawback of the device, given the significant reduction in responsivity at this operating point (2) the completeness of the experimental measurement conditions (3) whether the signal-to-noise ratio (SNR) is sufficient to support 10 Gbit/s transmission. We address each of these points below.

(1) Discussion on the maximum photocurrent: responsivity roll-off and saturation mechanisms

First, it should be clarified that 11.8 μA is not the maximum photocurrent demonstrated by the device. In the revised manuscript, photocurrent data at higher optical powers have been added. As shown in Figure R2.2, when the incident optical power is increased to 536 μW, the output photocurrent reaches 24 μA.

In the main text, a peak responsivity of ~492 mA/W is reported (Figure 1d), while the responsivity at the operating point under an optical power of 500 μW has already decreased. This behavior corresponds to a responsivity roll-off rather than a strict photocurrent saturation—it is the collection efficiency that decreases, while the photocurrent itself continues to increase. The observed responsivity roll-off does not by itself imply a corresponding degradation of the small-signal bandwidth, because responsivity and bandwidth characterize different aspects of the device response. The bandwidth is primarily governed by carrier transit time and the RC time constant under the small-signal operating condition. Of course, the responsivity roll-off indeed remains an important limitation for high-power operation.

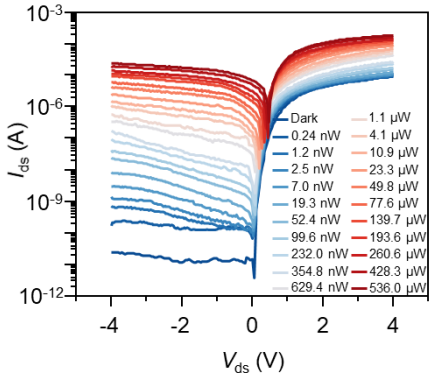


Figure R2.2(Revised Figure 1c) I_{ds}-V_{ds} curves of the WSe_2 p-i-n junction diode under dark condition and varying laser powers at a wavelength of 638 nm.

The responsivity roll-off could be attributed to the non-uniform optical field distribution within the waveguide structure. In the waveguide-integrated configuration, the optical field decays exponentially along the propagation direction, with the majority of the optical power absorbed at the front end of the waveguide, resulting in a much higher photocarrier density in the front-end region than in the rear-end region. This non-uniform distribution affects the device performance in two aspects. On one hand, the locally enhanced optical field causes the front-end region to reach saturation earlier; as the incident optical power increases, the saturated region gradually expands, reducing the effective area that continues to contribute to the photocurrent and slowing down the photocurrent rise rate (*IEEE Photonics J.* 2022, **14**, 6817304). On the other hand, the non-uniformly distributed photocarriers induce a space-charge effect, where the high-density photocarriers reduce the effective electric field intensity in the depletion region, degrading carrier transport efficiency (*ACS Photonics* 2022, **9**, 1764-1774). Both mechanisms may contribute to the responsivity roll-off with increasing optical power.

The analysis is further supported by the test data from the same device under free-space vertical illumination (Supplementary Figure 10): Importantly, under the same -4 V bias,

the same device produces >30 μA photocurrent under free-space illumination while maintaining a responsivity of ~0.1 A/W, indicating that the intrinsic carrier generation and collection capability of the WSe_2 p-i-n junction is not fundamentally limited to the ~10 μA level，The responsivity roll-off observed in the waveguide configuration is attributed to the waveguide structure. Future waveguide design optimization, such as tapered waveguides or distributed absorption structures, could homogenize the optical field along propagation, mitigate front-end absorption concentration, and improve the photocurrent of waveguide-integrated photodetectors at high optical powers. Corresponding discussions have also been added to the main text and Supplementary Information.

(2) Description of the experimental configuration and measurement conditions for optical communication tests

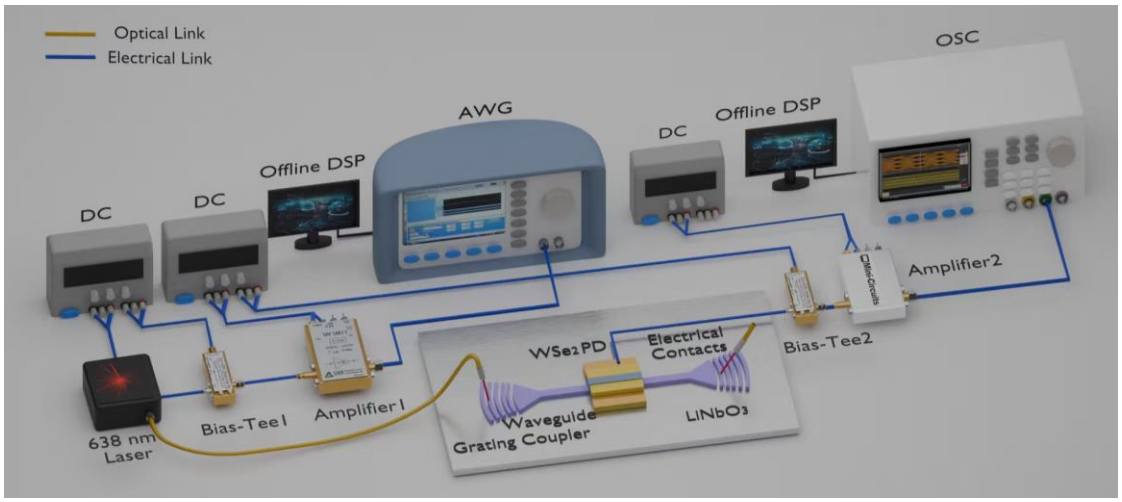


Figure R2.3(Revised Figure 4d) Schematic of the optical communication measurement system with DSP flow.

We thank the reviewer for pointing out this issue. Upon verification, we acknowledge that the measurement conditions were not clearly described in our previous response, where we confused the configurations for small-signal frequency response measurements with those for optical communication tests. We apologize for this confusion.

In practice, the two types of measurements employed different configurations. Bandwidth characterization only requires the detector response to be resolved above the measurement-system noise floor. Therefore, the small-signal frequency response was directly measured using a vector network analyzer, whose noise floor can reach -110 dBm, allowing the detector bandwidth to be obtained without an amplifier (Figure 4a and 4b). In contrast, optical communication requires sufficient SNR to satisfy a target BER under the complete receiver noise budget. For optical communication

measurements, a broadband amplifier with 40 dB gain (Mini-Circuits ZKL-1R5+) was placed in front of the oscilloscope (Tektronix MSO73304DX), as shown in Figure R2.3. Under this condition, the detector was reverse-biased at -4 V, and the incident optical power on the WSe_2 region was above 400 μW. The eye diagrams presented in the main text were measured under these conditions. The complete experimental configuration descriptions above have been included in the Supplementary Information of the revised manuscript.

(3) Quantitative SNR analysis and determination of the minimum photocurrent for optical communication tests.

Based on the above measurement configuration, we further performed a quantitative analysis of the SNR in the receiver chain to confirm the feasibility of optical communication under these conditions.

The effective noise bandwidth of the detection chain is primarily determined by the detector response within the relevant frequency range. For a conservative estimate, we approximate the effective noise bandwidth by the measured 3-dB bandwidth of 1.12 GHz, yielding an equivalent noise bandwidth of $B_n = (\pi/2) \times 1.12\ \text{GHz} = 1.76\ \text{GHz}$.

1) Total noise spectral density evaluation

The total noise consists of three components: the noise spectral density of the broadband amplifier, the noise spectral density of the oscilloscope, and the noise spectral density of the photodetector and optical link. To quantify the contributions of noise sources at different positions at a unified reference point, all noise sources in the receiver chain are referred to the amplifier input. These three types of noise are quantitatively analyzed below.

(i) Broadband amplifier noise

The broadband amplifier (Mini-Circuits ZKL-1R5+) has a noise figure of 3 dB, corresponding to a linear noise factor $F = 10^{NF/10} = 2.0$. The noise figure accounts for the amplifier's excess noise. Taking the thermal noise of the 50 Ω input impedance as the baseline, the equivalent input-referred noise spectral density is: $S_{I,Amp} = F \cdot \frac{4kT}{50} = 6.4 \times 10^{-22}\ \text{A}^2/\text{Hz}$.

(ii) Oscilloscope noise referred to the amplifier input

According to the Tektronix MSO73304DX specifications (50 mV/div, bandwidth filter on), the input-referred vertical noise is 0.58% of full scale, corresponding to an root-mean-square (RMS) noise of $V_{n,Scope} \approx 1.03 mV_{RMS}$. Referred to the amplifier input through the 40 dB gain, the equivalent noise spectral density is:

$$S_{I,Scope} = \frac{(V_{n,Scope} / (10^{40/20} \times 50\Omega))^2}{B_n} = 2.4 \times 10^{-23}\,\mathrm{A^2/Hz}.$$

(iii) Measured noise from the optical link and photodetector

The noise spectral density of the optical link and photodetector was measured directly using a lock-in amplifier (Zurich Instruments MFLI) under the conditions of V_{ds}=-4V and an incident power ~470μW. over a frequency range from 10 Hz to 50 kHz. Above 10 kHz, the noise spectrum is flat, exhibiting white noise characteristics (Supplementary Figure 18). The measured flat high-frequency noise floor was therefore used as a conservative approximation for the noise spectral density over the detector bandwidth, giving: $S_{I,PD+Opt} = 1.0 \times 10^{-22}\,\mathrm{A^2/Hz}$.

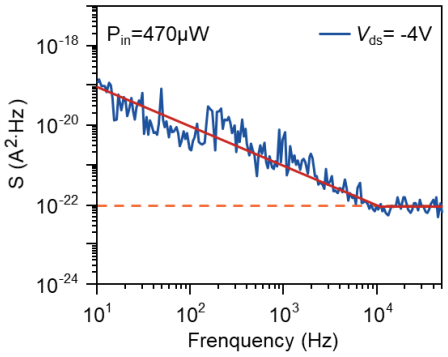


Figure R2.4 The noise power density as a function of frequency at V_{ds} = -4V under incident power of ~470 μW.

Thus, combining all noise contributions:

$$S_{I,total} = S_{I,Amp} + S_{I,Scope} + S_{I,PD+Opt} = 7.64 \times 10^{-22}\,\mathrm{A^2/Hz}.$$

2) Determination of the minimum photocurrent

Based on the above noise parameters, a quadratic equation model is used to extract the minimum DC photocurrent required for the PAM4 signal to reach the O-FEC threshold (BER=2×10^{-2}). The model takes into account the contributions of shot noise, thermal noise, amplifier noise, and oscilloscope noise: $A \cdot I_{DC}{}^2 - B \cdot I_{DC} - D = 0$.

Where $A = m^2$, $B = 2q \cdot SNR \cdot B_n$, $D = S_{I,total} \cdot SNR \cdot B_n$. The modulation depth is m=0.7, and the minimum SNR required for PAM4 at 2×10^{-2} is 11 dB (linear value 12.59). Substituting these values yields: $I_{DC,\min} = \frac{B + \sqrt{B^2 + 4AD}}{2A} = 5.9\ \mu\mathrm{A}$.

Thus, the calculated minimum photocurrent required to satisfy the O-FEC SNR criterion is approximately 5.9 μA under the measurement conditions. The photocurrent used in our 10-Gbit/s experiment was substantially higher than this value, providing sufficient SNR margin for the demonstrated data reception.

These additional measurements and quantitative noise analysis clarify both the high-power photocurrent capability of the device and the experimental conditions underlying the 10-Gbit/s data reception. We have incorporated these clarifications into the revised manuscript and Supplementary Information.

-5) I thank the authors for the inclusion of this discussion; this is an interesting addition. -6) The new discussion on the motivation of short-reach applications in the introduction is done well. -7) The clarification on the scope of the SOTA table helps.

We sincerely thank the reviewer for the supportive remarks on our revisions. These encouragements mean a great deal to us.

Reviewer #3 (Remarks to the Author):

The authors have provided a thorough and careful response, which I believe has significantly strengthened the manuscript.

We sincerely thank the reviewer for the positive feedback on our revised manuscript, and we are grateful to all reviewers for their precious time and professional insights.

Manuscript ID: NCOMMS-26-024174B

Title: Thin-Film Lithium Niobate Waveguide-Integrated p-i-n Tungsten Selenide Photodiodes with 10 Gbit/s Data Reception

Point by Point response to the Reviewer's Comments

We sincerely appreciate the reviewers for their continued attention to our manuscript and their professional insights. We are also grateful for the positive assessment of our work. We have carefully addressed each remaining comment and made corresponding revisions to further improve the clarity and accuracy of the manuscript. For easy identification, all revisions in the main text and Supplementary Information are highlighted in red, and our detailed point-by-point responses are provided below.

Reviewer #1 (Remarks to the Author):

I thank the reviewer for their efforts in revising the manuscript.

A minor issue that should be addressed is in Introduction, "For emerging chip-scale OIO, multiple device platforms are currently being actively explored, including silicon photonics, thin-film lithium (TFLN) photonics and vertical-cavity surface-emitting laser (VCSEL)-based systems [1, 5, 6]", the VCSEL-based system should not be listed along with silicon photonics and TFLN platforms, since they are different materials systems. The author may say "III-V materials systems."

I can recommend its publication after correcting the abovementioned sentence.

We thank the reviewer for this helpful suggestion. We agree that VCSEL-based systems should not be listed as a category equivalent to silicon photonics and thin-film lithium niobate (TFLN) photonics. We have revised the relevant text accordingly, replacing "VCSEL-based systems" with "III-V semiconductor photonics" to ensure a more consistent classification. The revision is as follows:

"For emerging chip-scale OIO, multiple device platforms are currently being actively explored, including silicon photonics, thin-film lithium (TFLN) photonics and III-V materials systems**"**

Reviewer #2 (Remarks to the Author):

Reviewer #2 (Remarks to the Author):

I thank the Author's for their detailed answers. The manuscript is now in a status where everything seems to be present such that a reader can come to their own conclusion. I recommend the manuscript for publishing. Yet, I still want to add two comments where I'll leave the decision with the editor:

1) “Upon verification, we acknowledge that the measurement conditions were not clearly described in our previous response, where we confused the configurations for small-signal frequency response measurements with those for optical communication tests. We apologize for this confusion.”

Such a statement does not spark a lot of confidence in the presented results, making me question what else is wrongly described and presented. Omitting a 40dB amplification is not a small thing, as the results would otherwise not have been achieved. And answering my remark by insisting it is possible and citing other works that contradict the author's own statement adds to the point where confidence in the author’s understanding and description of results is further lowered.

We thank the reviewer for pointing out this important issue. We acknowledge that our previous description of the amplifier usage was inaccurate, particularly in failing to clearly distinguish the receiver configurations between the small-signal frequency response test and the optical communication test, which caused unnecessary confusion. We sincerely apologize for this.

After re-verifying the experimental configurations and building upon the previous revisions, we have further refined the measurement configuration descriptions in the Methods section. This clearly distinguishes the two tests and ensures that the reported results strictly correspond to their respective measurement configurations.

1) The frequency response (bandwidth) measurement

As explicitly stated in the revised Methods, this test was performed directly using a vector network analyzer (VNA). The revision is as follows:

“For optical-to-electrical frequency response measurements, the output signal from the device was connected directly to a vector network analyzer. The 3-dB bandwidth was extracted under various reverse bias voltages.”

2) The optical communication (data transmission) measurement

This measurement utilized a completely independent receiver chain, with its system configuration illustrated in Figure 4d. The specific receiver configuration, including the

40 dB broadband amplifier and oscilloscope, is described in the Methods. The exact wording is as follows:

"For data transmission experiments, the photodetector was reverse-biased at -4 V. The incident optical power on the WSe_2 region was set above 400 μW. The receiver chain consisted of a broadband amplifier (Mini-Circuits ZKL-1R5+) and a Tektronix MSO73304DX oscilloscope."

Furthermore, the parameters and gain calculation of this amplifier for the optical communication receiver are explicitly provided in Supplementary Note 8. The exact wording is as follows:

"The broadband amplifier (Mini-Circuits ZKL-1R5+) has a noise figure of 3 dB... Referred to the amplifier input through the 40 dB gain, the equivalent noise spectral density is..."

We recognize that accurate descriptions of experimental configurations are fundamental to methodological reporting and essential for readers to independently evaluate the results. Therefore, we have corrected the relevant descriptions and verified the test configurations in both the manuscript and the Supplementary Information to improve transparency and reproducibility. We thank the reviewer again for this rigorous review.

2) While a careful reader might notice and draw their own conclusions, I still find statements like the one in the abstract potentially misleading: "The device features a record-high measured bandwidth of 1.12 GHz and at least 10 Gbit/s data rate in visible light. Simultaneously, the integration of TFLN waveguide overcomes the traditional speed-responsivity trade-off, achieving a responsivity of ~492 $mA \cdot W^{-1}$ ..."

10 Gbit/s data rate was achieved, but not simultaneously with a responsivity of 0.5 A/W. The responsivity was clearly diminished by operating the device at high optical input powers. In the conclusion, the same statement is made using "alongside" instead of "simultaneously". And yes, the device itself achieves these parameters (all on in device), but not with the same operation conditions. While re-reading this I also noticed that the abstract states at least 10 Gbit/s, while the conclusion states up to 10 Gbit/s.

We thank the reviewer for pointing out this issue. We completely agree with your comments and have made the following revisions:

1) The revision of statements such as "Simultaneously"

We agree with the reviewer's point. Although the test conditions for each figure of merit are specified separately in the main text, the wording in the original abstract was indeed likely to mislead readers into thinking that the high responsivity and the 10

Gbit/s data rate were achieved “simultaneously” under the same conditions. Therefore, we have revised the relevant statements in the abstract and conclusion, removing wording that implies the simultaneous achievement of both performances, to more clearly and accurately distinguish the performance results obtained under different test conditions.

2) The unified statement of the data rate

We acknowledge the previous inconsistency between the abstract and the conclusion. As stated in Supplementary Note 5, the physical bandwidth of the laser used in the test system (1.18 GHz) limits the upper limit of the system’s frequency response. This limitation also restricts the maximum data rate supported by the system in the data transmission tests. With the current system, 10 Gbit/s is the highest error-free rate that can be verified. Without experimental data supporting higher rates, using “at least” is not rigorous. Therefore, we have uniformly revised the relevant statements throughout the manuscript to “up to 10 Gbit/s”.

The specific revisions are as follows:

Abstract

“The device features a measured bandwidth of 1.12 GHz and a data rate up to 10 Gbit/s in visible light. Separately, the device achieves a responsivity of ~492 mA·W^{-1} and a near-unity external quantum efficiency of 95.6% at 638 nm.”

Discussion

“Notably, the TFLN waveguide integration mitigates the conventional responsivity-speed trade-off, achieving a competitive *R* of ~492 mA/W with a near unity EQE of 95.6% at 638nm. The device also exhibits a measured O-E bandwidth of 1.12 GHz and a data transmission up to 10 Gbit/s within 15.3% O-FEC threshold of 2 ×10^{-2}.”

We sincerely thank all the reviewers and the editor again for their valuable time and efforts during the review process.

Manuscript ID: NCOMMS-26-024174B

Title: Thin-Film Lithium Niobate Waveguide-Integrated p-i-n Tungsten Selenide Photodiodes with 10 Gbit/s Data Reception

Point by Point response to the Reviewer's Comments

We sincerely appreciate the reviewers for their continued attention to our manuscript and their professional insights. We are also grateful for the positive assessment of our work. We have carefully addressed each remaining comment and made corresponding revisions to further improve the clarity and accuracy of the manuscript. For easy identification, all revisions in the main text and Supplementary Information are highlighted in red, and our detailed point-by-point responses are provided below.

Reviewer #1 (Remarks to the Author):

I thank the reviewer for their efforts in revising the manuscript.

A minor issue that should be addressed is in Introduction, "For emerging chip-scale OIO, multiple device platforms are currently being actively explored, including silicon photonics, thin-film lithium (TFLN) photonics and vertical-cavity surface-emitting laser (VCSEL)-based systems [1, 5, 6]", the VCSEL-based system should not be listed along with silicon photonics and TFLN platforms, since they are different materials systems. The author may say "III-V materials systems."

I can recommend its publication after correcting the abovementioned sentence.

We thank the reviewer for this helpful suggestion. We agree that VCSEL-based systems should not be listed as a category equivalent to silicon photonics and thin-film lithium niobate (TFLN) photonics. We have revised the relevant text accordingly, replacing "VCSEL-based systems" with "III-V semiconductor photonics" to ensure a more consistent classification. The revision is as follows:

"For emerging chip-scale OIO, multiple device platforms are currently being actively explored, including silicon photonics, thin-film lithium (TFLN) photonics and III-V materials systems**"**

Reviewer #2 (Remarks to the Author):

Reviewer #2 (Remarks to the Author):

I thank the Author's for their detailed answers. The manuscript is now in a status where everything seems to be present such that a reader can come to their own conclusion. I recommend the manuscript for publishing. Yet, I still want to add two comments where I'll leave the decision with the editor:

1) “Upon verification, we acknowledge that the measurement conditions were not clearly described in our previous response, where we confused the configurations for small-signal frequency response measurements with those for optical communication tests. We apologize for this confusion.”

Such a statement does not spark a lot of confidence in the presented results, making me question what else is wrongly described and presented. Omitting a 40dB amplification is not a small thing, as the results would otherwise not have been achieved. And answering my remark by insisting it is possible and citing other works that contradict the author's own statement adds to the point where confidence in the author’s understanding and description of results is further lowered.

We thank the reviewer for pointing out this important issue. We acknowledge that our previous description of the amplifier usage was inaccurate, particularly in failing to clearly distinguish the receiver configurations between the small-signal frequency response test and the optical communication test, which caused unnecessary confusion. We sincerely apologize for this.

After re-verifying the experimental configurations and building upon the previous revisions, we have further refined the measurement configuration descriptions in the Methods section. This clearly distinguishes the two tests and ensures that the reported results strictly correspond to their respective measurement configurations.

1) The frequency response (bandwidth) measurement

As explicitly stated in the revised Methods, this test was performed directly using a vector network analyzer (VNA). The revision is as follows:

“For optical-to-electrical frequency response measurements, the output signal from the device was connected directly to a vector network analyzer. The 3-dB bandwidth was extracted under various reverse bias voltages.”

2) The optical communication (data transmission) measurement

This measurement utilized a completely independent receiver chain, with its system configuration illustrated in Figure 4d. The specific receiver configuration, including the

40 dB broadband amplifier and oscilloscope, is described in the Methods. The exact wording is as follows:

"For data transmission experiments, the photodetector was reverse-biased at -4 V. The incident optical power on the WSe_2 region was set above 400 μW. The receiver chain consisted of a broadband amplifier (Mini-Circuits ZKL-1R5+) and a Tektronix MSO73304DX oscilloscope."

Furthermore, the parameters and gain calculation of this amplifier for the optical communication receiver are explicitly provided in Supplementary Note 8. The exact wording is as follows:

"The broadband amplifier (Mini-Circuits ZKL-1R5+) has a noise figure of 3 dB... Referred to the amplifier input through the 40 dB gain, the equivalent noise spectral density is..."

We recognize that accurate descriptions of experimental configurations are fundamental to methodological reporting and essential for readers to independently evaluate the results. Therefore, we have corrected the relevant descriptions and verified the test configurations in both the manuscript and the Supplementary Information to improve transparency and reproducibility. We thank the reviewer again for this rigorous review.

2) While a careful reader might notice and draw their own conclusions, I still find statements like the one in the abstract potentially misleading: "The device features a record-high measured bandwidth of 1.12 GHz and at least 10 Gbit/s data rate in visible light. Simultaneously, the integration of TFLN waveguide overcomes the traditional speed-responsivity trade-off, achieving a responsivity of ~492 $mA \cdot W^{-1}$ ..."

10 Gbit/s data rate was achieved, but not simultaneously with a responsivity of 0.5 A/W. The responsivity was clearly diminished by operating the device at high optical input powers. In the conclusion, the same statement is made using "alongside" instead of "simultaneously". And yes, the device itself achieves these parameters (all on in device), but not with the same operation conditions. While re-reading this I also noticed that the abstract states at least 10 Gbit/s, while the conclusion states up to 10 Gbit/s.

We thank the reviewer for pointing out this issue. We completely agree with your comments and have made the following revisions:

1) The revision of statements such as "Simultaneously"

We agree with the reviewer's point. Although the test conditions for each figure of merit are specified separately in the main text, the wording in the original abstract was indeed likely to mislead readers into thinking that the high responsivity and the 10

Gbit/s data rate were achieved “simultaneously” under the same conditions. Therefore, we have revised the relevant statements in the abstract and conclusion, removing wording that implies the simultaneous achievement of both performances, to more clearly and accurately distinguish the performance results obtained under different test conditions.

2) The unified statement of the data rate

We acknowledge the previous inconsistency between the abstract and the conclusion. As stated in Supplementary Note 5, the physical bandwidth of the laser used in the test system (1.18 GHz) limits the upper limit of the system’s frequency response. This limitation also restricts the maximum data rate supported by the system in the data transmission tests. With the current system, 10 Gbit/s is the highest error-free rate that can be verified. Without experimental data supporting higher rates, using “at least” is not rigorous. Therefore, we have uniformly revised the relevant statements throughout the manuscript to “up to 10 Gbit/s”.

The specific revisions are as follows:

Abstract

“The device features a measured bandwidth of 1.12 GHz and a data rate up to 10 Gbit/s in visible light. Separately, the device achieves a responsivity of ~492 mA·W^{-1} and a near-unity external quantum efficiency of 95.6% at 638 nm.”

Discussion

“Notably, the TFLN waveguide integration mitigates the conventional responsivity-speed trade-off, achieving a competitive *R* of ~492 mA/W with a near unity EQE of 95.6% at 638nm. The device also exhibits a measured O-E bandwidth of 1.12 GHz and a data transmission rate up to 10 Gbit/s within 15.3% O-FEC threshold of 2 ×10^{-2}.”

We sincerely thank all the reviewers and the editor again for their valuable time and efforts during the review process.