SCIENCE CHINA
Information Sciences

• Supplementary File •

Supplemental material: InGaN-based high-speed mini laser diode surpasses PAM-4 visible light links by over 30 Gbps

Junhui Hu[1], Zhenqian Gu[1], Haolin Jia[1], Zhen Yang[1], Zengxin Li[1], Jiabin Wu[1], Leihao Sun[1], Aolong Sun[1], Ouhan Huang[1], Changsheng Xia[2], Boon S. Ooi[3], Jianyang Shi[1], Ziwei Li[1], Junwen Zhang[1], Shaohua Yu[4], Nan Chi[1] & Chao Shen[1*]

[1]*Key Laboratory for Information Science of Electromagnetic Waves (MoE), College of Future Information Technology, Fudan University, Shanghai, 200438, China*
[2]*GMPT Company Ltd., Shanghai, 201210, China*
[3]*Photonics Laboratory, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955, Saudi Arabia*
[4]*Peng Cheng Laboratory, Shenzhen, Guangdong, 518055, China*

Appendix A Experimental setup of the mini-LD based high-speed VLC system

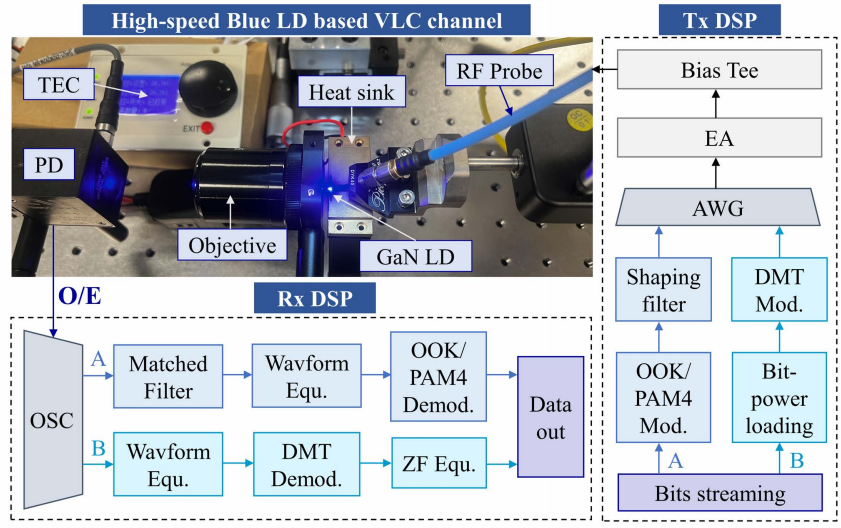

Figure A1 Experimental setup of high-speed blue LD based VLC system, including digital signal processing (DSP) at transmit (Tx) and receive (Rx) ends.

To verify the communication performance of the mini-LD, a chip-level high-speed VLC system is established in Figure A1. For digital signal processing (DSP) at transmit (Tx) end, a random bit data stream is generated on the computer, and the binary bit stream is then modulated into OOK, PAM4, or DMT formats. Here, we analyze the data rates of the system under these three modulation formats. OOK and PAM4 are baseband modulation signals, employing a factor of 2 for up-sampling with the copy-and-write method. A roll-off filter is applied using an RC-type Nyquist filter, with a roll-off factor of 0.05 to ensure the signal bandwidth does not exceed the device bandwidth. DMT is a multicarrier modulation scheme that assigns bit loading to each subcarrier based on the measured SNR using a 4-QAM signal. Additionally, the Levin-Campello (LC) bit-power loading algorithm is employed to maximize spectral efficiency [1]. The up-sampling factor is also set as 2, with 1024 subcarriers, each containing 200 data points. A cyclic prefix with a length of 32 is used to mitigate inter-symbol interference (ISI). Furthermore, to avoid the stop-band at low frequency caused by electrical amplification/attenuation components in the link, a zero-padding sequence with a length of 8 is added at the beginning of the sequence. Then, the modulated signal is electrically amplified (ADM1-0026PA) and connected into a bias-Tee (BTN2-0018) with the DC

* Corresponding author (email: chaoshen@fudan.edu.cn)

signal. The RF+DC signal is input into an arbitrary waveform generator (AWG, Keysight M8195A) for digital-to-analog conversion. The output analog signal is directly applied to the blue LD through an RF probe for electro-optical modulation. The mini-LD is placed on a heat sink, and its temperature is maintained at room temperature by utilizing a thermoelectric cooler (TEC). The modulated optical signal is focused by a 10× objective lens and received by a high-speed PD. After optical-to-electrical conversion, the signal is displayed on a high-speed oscilloscope (DSA-X 92004A). The received signal is captured on the computer for Rx DSP. The OOK and PAM4 signals undergo matched filtering to eliminate out-of-band noise, followed by compensation of the received signal using a Volterra filter. After demodulation, the recovered signal is obtained. For DMT format, we use bidirectional gated recurrent unit (BiGRU) post-equalization to mitigate nonlinear effects [2]. After DMT demodulation, zero-forcing (ZF) equalization is applied to obtain the recovered signal. Subsequently, the BER of the transmitted data stream can be calculated.

Appendix B Principles involved in modulation and demodulation

In practical communication systems, signal-to-noise ratio (SNR) of the channel is often estimated using the error vector magnitude (EVM) [3]. Due to the impact of noise and device nonlinearities during transmission, there is inevitably a deviation between the received signal and the ideal signal. To quantify this deviation, EVM is defined as the magnitude of the error vector. In calculations, EVM is typically expressed as a normalized percentage or in decibels (dB). To reflect the overall performance, the average EVM across all subcarriers is computed. The corresponding formula is

$$\overline{EVM} = \sqrt{\frac{\frac{1}{N}\sum_{n=1}^{N}[S_R(n) - S_O(n)]^2}{\frac{1}{N}\sum_{n=1}^{N}[S_R(n)]^2}} \tag{B1}$$

Where average EVM represents the average normalized error magnitude, S_R denotes the received symbols, S_O denotes the ideal symbols, and N is the total number of received symbols. Here, it is assumed that the only source of signal distortion is noise. Therefore, a larger EVM indicates greater noise and larger symbol error, while a smaller EVM implies lower noise and more accurate signal recovery. When N is sufficiently large, it is evident from the formula that the numerator approximates the average noise power N_0, and the denominator corresponds to the average signal power P_0. Thus, EVM and SNR satisfy the following relationship.

$$SNR = 10\log_{10}\left(\frac{1}{\overline{EVM}^2}\right) \tag{B2}$$

After estimating the channel SNR through EVM by the QPSK pre-transmission, it is necessary to allocate the number of bits based on the relationship between SNR and QAM order. However, although conventional bit-loading techniques can achieve relatively high spectral efficiency, they often leave residual power unused across subcarriers and thus cannot strictly approach the Shannon limit. To address this, our work adopts a bit-power adaptive loading technique (LC algorithm) to minimize power redundancy.

The LC algorithm primarily addresses two optimization problems:

1. Bit Rate Maximization Problem (BRMP): determining the maximum achievable bit rate under a fixed total power constraint.

2. Minimum Margin Power Problem (MMP): determining the minimum total power required to transmit a fixed bit rate.

Evidently, within the scope of this study, we focus on the BRMP to achieve the highest possible data rate. The formulation of the BRMP is described as follows.

$$\begin{aligned} &\max_{b\in\mathbb{Z}_{b_m}^N} \quad \sum_{n=1}^{N} b_n \\ &\text{s.t.} \quad \sum_{n=1}^{N} p_n \leqslant P_{\text{total}} \end{aligned} \tag{B3}$$

In the above formulation, b_n represents the number of bits allocated to the n-th subcarrier, and p_n denotes the corresponding power allocation. P_{total} is the total available transmission power, and N is the number of subcarriers. To improve computational efficiency, the allocated bit number is typically constrained to be an integer not exceeding a predefined maximum value b_m. In practical communication systems, the bit number rarely exceeds 8 (corresponding to 256-QAM modulation). Therefore, in this work, we set $b_m = 8$. Based on the calculated SNR_i for each subcarrier, a preliminary bit loading is performed using the FBL scheme [4]. If only a few subcarriers (typically fewer than five when $N = 256$) are assigned the maximum bit level, the bit order of those subcarriers is reduced by one, and the excess power is transferred to P_a for reallocation. After this peak-trimming step, the current bit allocation is denoted by k_i. The LC algorithm then computes the corresponding power allocation P_a as [5].

$$P_a = \sum_{i=1}^{N}\left[SNR_i - SNR_i^{req}(k_i)\right] \tag{B4}$$

Each subcarrier's power ratio is adjusted to $SNR_i^{req}(k_i)/SNR_i$ under the assumption of a linear system response. The remaining power budget is then allocated to the subcarrier with the highest priority need. The optimal subcarrier for

adding one additional bit is the one whose projection margin has the smallest absolute value P_m, thereby minimizing the reduction in P_a. For the i-th subcarrier, the projection margin P_{mi} is defined as the difference between the estimated SNR and the SNR required for the currently assigned bit level, plus the additional SNR needed to raise the bit level by B. Its magnitude is given by

$$|P_{mi}| = SNR_i^{req}(k_i + B) - SNR_i \tag{B5}$$

In this procedure, B is fixed at 1 for each allocation step. Once the allocation meets the efficiency criterion defined in [6], the optimal power and bit distribution is achieved. The resulting spectral efficiency (SE) can be expressed as

$$SE = \frac{1}{N}\sum_{i=1}^{N} k_i \quad [\text{bits} \cdot \text{s}^{-1} \cdot \text{Hz}^{-1}] \tag{B6}$$

To compensate for the nonlinear distortions introduced by the device and the transmission channel, we employ an adaptive recursive-least-squares (RLS) Volterra filter [7]. By updating the filter coefficients recursively through the recursive-least-squares algorithm, this approach simultaneously mitigates linear impairments as well as second- and third-order nonlinear effects. In discrete-time baseband processing, the third-order Volterra series represents the system output $y[n]$ as the sum of the linear, quadratic, and cubic convolutions (or cross-convolutions) of the input $\boldsymbol{x}[n]$.

$$\begin{aligned}
&y[n] = \boldsymbol{W}^T \boldsymbol{x}[n],\ \boldsymbol{x}[n] = \big[\, \boldsymbol{x}_{lin}^T[n],\ \boldsymbol{x}_{quad}^T[n],\ \boldsymbol{x}_{cubic}^T[n] \,\big]^T, \\
&\boldsymbol{W} = \big[\, \boldsymbol{W}_{lin}^T[n],\ \boldsymbol{W}_{quad}^T[n],\ \boldsymbol{W}_{cubic}^T[n] \,\big]^T, \\
&\boldsymbol{x}_{lin}[n] = \big[\, x[n],\ x[n-1],\ \ldots,\ x[n-L_1+1] \,\big]^T, \\
&\boldsymbol{x}_{quad}[n] = \mathrm{vec}\big(x[n-k_1]\, x[n-k_2]\big)_{0 \leqslant k_2 \leqslant k_1 < L_2}, \\
&\boldsymbol{x}_{cubic}[n] = \mathrm{vec}\big(x[n-k_1]\, x[n-k_2]\, x[n-k_3]\big)_{0 \leqslant k_3 \leqslant k_2 \leqslant k_1 < L_3}.
\end{aligned} \tag{B7}$$

Here, the matrix $\boldsymbol{W}$ represents the third-order Volterra convolution kernel, which can simultaneously compensate for the channel's linear distortion, second-order nonlinearities (AM–AM/AM–PM), and third-order nonlinearities (e.g., intermodulation and memory effects). The parameters L1, L2 and L3 denote the numbers of taps in the first-, second-, and third-order kernels, respectively. The kernel $\boldsymbol{W}$ is updated by a RLS iteration as follows,

$$\begin{aligned}
&\boldsymbol{W}_i = \boldsymbol{W}_{i-1} + \boldsymbol{G}_i e_i^*,\ e_i[n] = y[n] - \boldsymbol{W}_i^T \boldsymbol{x}[n], \\
&\boldsymbol{G}_i = \frac{\boldsymbol{P}_{i-1}\boldsymbol{x}}{\gamma + \boldsymbol{x}^T \boldsymbol{P}_{i-1}\boldsymbol{x}},\ \boldsymbol{P}_i = \frac{1}{\gamma}\Big(\boldsymbol{P}_{i-1} - \boldsymbol{G}_i\, \boldsymbol{x}^T \boldsymbol{P}_{i-1}\Big).
\end{aligned} \tag{B8}$$

Here, e_i denotes the residual at the i-th iteration, γ is the forgetting factor, and the matrix $\boldsymbol{P}$ is initialized as an identity matrix scaled to the filter length. The center tap of the linear kernel is set to 1, whereas all second- and third-order taps are initialized to a small value δ. Both the received and transmitted signals are mean-square normalized to prevent numerical overflow and to keep the adaptation step stable. The first N_{train} known symbols are employed as a training sequence. The Volterra kernel $\boldsymbol{W}$ is updated iteratively, and the error is recorded to monitor convergence. Once training is complete, $\boldsymbol{W}$ is frozen and applied to the subsequent unknown data stream, thereby performing blind nonlinear compensation. The parameters used for the third-order RLS-Volterra filter in this work are summarized in Table B1.

Table B1 The parameters used for the third-order RLS-Volterra filter.

Parameters	Value
Training length N_{train}	7168
Numbers of taps in the first-order kernels L_1	73
Numbers of taps in the second-order kernels L_2	13
Numbers of taps in the third-order kernels L_3	9
Forgetting factor γ	0.9999
Peak-tap initialization coefficient δ	0.00001

After symbol-level post-equalization is applied to compensate for nonlinear impairments, the received waveform must be demodulated and its BER evaluated for the chosen modulation format. Taking M-order pulse-amplitude modulation (PAM) as an example (the procedure for QAM is analogous), the workflow is as follows.

1. Constellation construction: the ideal PAM constellation is generated as an equally spaced zero-mean set of M real amplitudes.

$$C = \{-M+1,\ -M+3,\ \cdots,\ M-1\} \tag{B9}$$

For fair comparison the constellation is power-normalized.

$$\overline{C} = \frac{C}{\sqrt{\frac{1}{M}\sum_{c \in C} |c|^2}} \tag{B10}$$

2. Symbol decision (nearest-neighbor mapping): for each equalized sample $r[n]$, compute its Euclidean distance for every ideal point c_k.

$$d_k[n] = |r[n] - c_k| \tag{B11}$$

Select the index of the closest point,

$$\hat{k}[n] = \arg\min_k d_k[n] \tag{B12}$$

and assign $s[n] = c_{\hat{k}[n]}$. This implements maximum likelihood (nearest-neighbor) symbol detection. The following figure illustrates a decision example for 16.5 and 17.5 Gbaud PAM4 signals transmission.

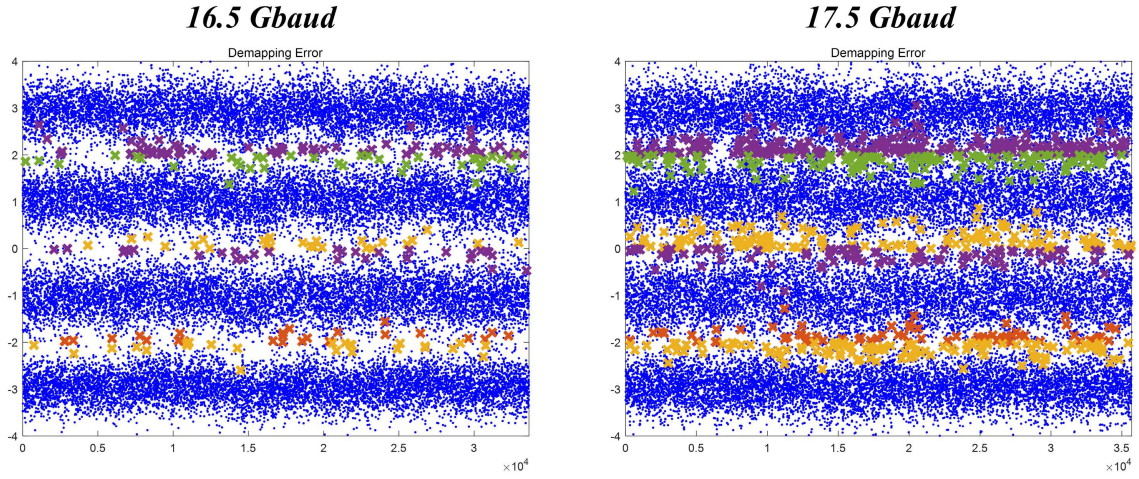

Figure B1 Decision-error diagram for 16.5/17.5 Gbaud PAM-4 signals. Blue dot is the right symbol, and the '×' represents the error symbol.

3. Gray decoding: each symbol index $k[n]$ is converted to a binary word via the M-PAM Gray map, yielding the received bit stream $\boldsymbol{b}_{rx}$. The transmitted training/payload symbols are processed in the same way, producing the reference bit stream $\boldsymbol{b}_{tx}$.

4. BER evaluation: the initial training samples in the bit stream are discarded, and the bit error count is computed solely over the valid payload portion.

$$\mathrm{BER} = \frac{\mathrm{num}(\boldsymbol{b}_{\mathrm{tx}} \neq \boldsymbol{b}_{\mathrm{rx}})}{\mathrm{num}(\boldsymbol{b}_{\mathrm{tx}})} \tag{B13}$$

For the received signals shown in Figure B1, the calculated BERs are 3.47×10^{-3} at 16.5 Gbaud and 1.13×10^{-2} at 17.5 Gbaud.

References

1 Hu J, Hu F, Jia J, et al. 46.4 Gbps visible light communication system utilizing a compact tricolor laser transmitter. Opt Express, 2022, 30: 4365–4373
2 Niu W, Xu Z, Liu Y, et al. Key technologies for high-speed Si-substrate LED based visible light communication. J Lightwave Technol, 2023, 41: 3316–3331
3 Shafik R A, Rahman M S, Islam A R. On the extended relationships among EVM, BER and SNR as performance metrics. In: Proceedings of IEEE International Conference on Electrical and Computer Engineering, Dhaka, 2006. 408–411
4 Zou P, Zhao Y, Hu F, et al. Underwater visible light communication at 3.24 Gb/s using novel two-dimensional bit allocation. Opt Express, 2020, 28: 11319–11338
5 Hu F, Chen S, Li G, et al. Si-substrate LEDs with multiple superlattice interlayers for beyond 24 Gbps visible light communication. Photonics Res, 2021, 9: 1581–1591
6 Campello J. Practical bit loading for DMT. In: Proceedings of IEEE International Conference on Communications, Vancouver, 1999. 801–805
7 Li G, Li Z, Ha Y, et al. Performance assessments of joint linear and nonlinear pre-equalization schemes in next generation IM/DD PON. J Lightwave Technol, 2022, 40: 5478–5489