Opto-Electronic Science

Supplementary information

February 2025, Vol. 4, No. 2

DOI: 10.29026/oes.2025.240020 CSTR: 32246.14.oes.2025.240020

Direct detection with an optimal transfer function: toward the electrical spectral efficiency of coherent homodyne detection

Xingfeng Li[1], Jingchi Li[1], Xiong Ni[1], Hudi Liu[1], Qunbi Zhuge[1], Haoshuo Chen[2], William Shieh[3] and Yikai Su[1]*

[1]State Key Lab of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; [2]Nokia Bell Labs, 600 Mountain Ave, Murray Hill, NJ 07974, USA; [3]School of Engineering, Westlake University, Hangzhou 310030, China.

*Correspondence: YK Su, E-mail: yikaisu@sjtu.edu.cn

This file includes:

Supplementary information for this paper is available at https://doi.org/10.29026/oes.2025.240020

Section 1: Simulation setup

The simulation setup and digital signal processing (DSP) stacks are provided in Fig. S1. In the transmitter side, the binary data are first mapped into a total of 1638400 16-QAM symbols for two independent left sideband (LSB) and right sideband (RSB) signals. Random binary strings are constructed using the Mersenne Twister algorithm[S1]. Then, two root raised cosine (RRC) filters are used to produce the corresponding Nyquist-shaped LSB and RSB signals at the roll-off factors of 0.01. The sampling rate is set to 2*B*, where *B* is the Baud rate. We combine the LSB and RSB signals to generate a dual-SSB 16-QAM signal with a guard band. A carrier is added at the central frequency of the signal, according to the required carrier-to-signal power ratio (CSPR). After loading additive white Gaussian noise (AWGN) based on the optical signal-to-noise ratio (OSNR), our proposed scheme is used to detect a complex-valued double-sideband (CV-DSB) signal. To avoid the spectral aliasing induced by the square-law detection of the photodiode (PD), the optical signal is resampled to 4*B*. After the two PDs, the resulting electrical signals are resampled to 2*B*.

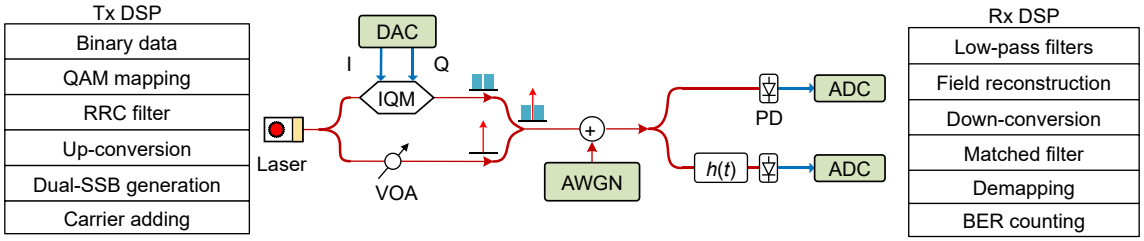

Fig. S1 | Simulation setup and DSP flow charts. DAC: digital-to-analog converter; IQM: in-phase and quadrature modulator; VOA: variable optical attenuator; ADC: analog-to-digital converter; BER: bit error ratio.

In the receiver-side DSP, two rectangular low-pass filters are used to remove the out-of-band SSBI components and undesired components within the low-frequency guard band. For field reconstruction, an iterative algorithm is used for signal-to-signal beat interference (SSBI) cancellation. Alternatively, a convolutional neural network (CNN) can be used to seamlessly achieve signal reconstruction and SSBI mitigation in a unified step. We use the same CNN structure employed in our previous work[S2]. It is also shown in Fig. S2. In the deep CNN, 80% of the data samples are used for training, and the remaining 20% samples are used for testing. The kernel sizes are set to 19 and 1 for the convolutional layers and shortcut connections, respectively. To reduce the computational complexity and increase the depth of the CNN, 1×1 convolutions are used for all six convolutional layers in the three residual blocks[S2]. After field recovery, the CV-DSB signal is decomposed via down-conversions and matched filters. Finally, the BER is calculated to characterize the system performance.

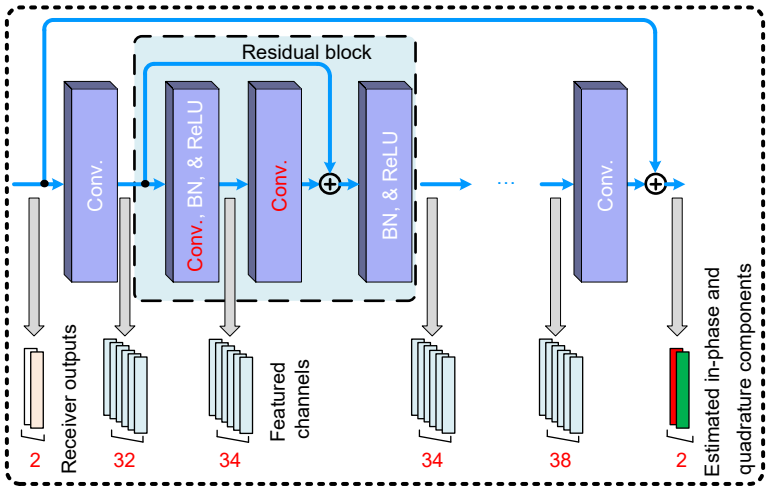

Fig. S2 | Schematic of the deep CNN structure. Conv.: convolutional layer; BN: batch normalization; ReLU: rectified linear unit.

Section 2: Theoretical ESE limit

2.1 Theoretical derivation

In this section, we will derive the theoretical electrical spectral efficiency (ESE) limit by invoking Shannon's formula[S3]:

$$R = \alpha B \cdot \log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{B}\right), \tag{S1}$$

where R is the net data rate, α is an interface rate scaling factor, B is the optical signal bandwidth, β_1 is a signal power scaling factor that is related to polarization-division multiplexing, β_2 is a noise enhancement factor, and B_{ref} is 12.5 GHz.

In an amplifier noise-limited direct detection (DD) system, if the OSNR and the power of amplified spontaneous emission noise are fixed, the sum of the signal and carrier powers is also fixed[S4]. We assume that SSBI can be eliminated by using an optimal receiver structure with three phase shifters. In this scenario, the CSPR required by the system is minimal. After the PD, the desired carrier-signal beating term and the electrical SNR of the electrical signal can be maximized when the signal power is equal to the carrier power[S4,S5]. Thus, the lowest optimum CSPR for our proposed scheme is 0 dB.

Because only a near-zero guard band is required at the signal's central frequency, the lower limit of the receiver electrical bandwidth is $B/2$. The ESE is expressed as

$$ESE = \frac{R}{B/2} = \frac{2R}{B}. \tag{S2}$$

Substituting Eq. (S2) into Eq. (S1), we obtain the theoretical ESE limit of the proposed scheme as follows:

$$ESE = 2\alpha \cdot \log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{2R}ESE\right), \tag{S3}$$

where α, β_1, and β_2 are provided in Table S1. Eq. (S3) is an implicit function. A bisection method is demanded to solve for the ESE.

For coherent detection, the CSPR is $-\infty$ dB. Since the coherent receiver mixes a local oscillator with a randomly polarization-rotated signal, it inherently requires a polarization-diversity setup. The ESE of the homodyne coherent is expressed as:

$$ESE = \frac{R/2}{B/2} = \frac{R}{B}. \tag{S4}$$

Plugging Eq. (S4) into Eq. (S1), we obtain the theoretical ESE limit of the homodyne coherent as follows:

$$ESE = \alpha \cdot \log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{R}ESE\right). \tag{S5}$$

For a Kramers-Kronig (KK) receiver, the ESE can be written as

$$ESE = \frac{R}{B}. \tag{S6}$$

Plugging Eq. (S6) into Eq. (S1), we obtain the theoretical ESE limit of the KK receiver as follows:

$$ESE = \alpha \cdot \log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{R}ESE\right). \tag{S7}$$

A Stokes-vector receiver (SVR) uses two orthogonal polarizations to transmit information-bearing signal and carrier. The theoretically lowest CSPR is 0 dB, which maximizes the desired carrier-signal beating term and SNR of the electrical signal[S4,S5]. The theoretical ESE limit of the SVR is also expressed as shown in Eq. (S5), but the parameters α, β_1, and β_2 are different.

Carrier-assisted differential detection (CADD) receiver requires guard bands at every periodic frequency null of the transfer function[S6]. We use GB to denote the total guard bands. The GB is related to the transfer function of the CADD receiver and is not a very small value. It is difficult to obtain the minimum value of GB. The ESE is expressed as:

$$ESE = \frac{R}{(B+GB)/2} = \frac{2R}{B+GB}\ . \tag{S8}$$

Substituting Eq. (S8) into Eq. (S1), we obtain the theoretical ESE limit of the CADD receiver as follows:

$$ESE = \alpha \cdot \frac{2R - GB \cdot ESE}{R} \log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR} \cdot \frac{B_{\mathrm{ref}} \cdot ESE}{2R - GB \cdot ESE}\right)\ . \tag{S9}$$

The theoretical ESE limits for various receiver schemes are summarized in Table S1.

Table S1 | The theoretical ESE limits for various receiver schemes.

Receiver schemes	α	β_1	β_2	Theoretical *ESE* limits
KK	1	2	2	$ESE = \alpha\log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{R}ESE\right)$
SVR	1	2	2	$ESE = \alpha\log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{R}ESE\right)$
CADD	1	2	1	$ESE = \alpha \cdot \frac{2R - GB \cdot ESE}{R}\log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR} \cdot \frac{B_{\mathrm{ref}} \cdot ESE}{2R - GB \cdot ESE}\right)$
Homodyne coherent	2	1	1	$ESE = \alpha\log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{R}ESE\right)$
This work	1	2	1	$ESE = 2\alpha\log_2\left(1 + \frac{\beta_1}{\beta_2}\frac{OSNR}{1+CSPR}\frac{B_{\mathrm{ref}}}{2R}ESE\right)$

2.2 ESE and cost metrics at a 200-Gb/s net rate and 30-dB OSNR

Table S2 shows the ESE and cost metrics between the coherent detection and various direct detections at a 200-Gb/s net rate and 30-dB OSNR. This table is reproduced from ref.[S3,S6]. For KK and CADD receivers, the theoretical minimum CSPR and guard band are challenging to derive. We only provide the simulated values. Finally, the ESE limits are normalized to the homodyne-coherent case. The theoretical ESE limit of the proposed scheme is 88.22% of the homodyne coherent. Our proposed scheme is the closest to the homodyne coherent.

Table S2 | ESE and cost metrics of the 200-Gb/s net interface rate per polarization per wavelength detection system with a 30-dB OSNR.

Receiver scheme	CSPR (dB)	Guard band (GHz)	*ESE* (b/s/Hz)	Normalized *ESE*	# of PDs	# of ADCs	Requirement of LO
KK[S7]	6 (Simulated)	0 (Theoretical)	6.33	30.43%	1	1	No
SVR[S8]	0 (Theoretical)	0 (Theoretical)	7.94	38.17%	6	3	No
CADD[S6]	8 (Simulated)	2.5 (Simulated)	12.7	61.06%	5	3	No
Homodyne coherent [S3]	−∞ (Theoretical)	0 (Theoretical)	20.8	100%	4	2	Yes
This work	3 (Simulated)	~1×10^{-9} (Simulated)	16.95	81.49%	2	2	No
This work	0 (Theoretical)	~0 (Theoretical)	18.35	**88.22%**	2	2	**No**

Section 3: Experimental setup

The experimental setup and photographs of key devices are presented in Fig. S3. In the DSO, we utilize Channel 2B and 3B with a 3-dB bandwidth of 59 GHz and a sampling rate of 160 GSa/s.

The WS is a key device for constructing the optimal transfer function. We use MATLAB to generate the desired amplitude and phase responses. The MATLAB codes are as follows:

```
clear
close all
clc
Degr = 90; % Degree
GuardBand = 10; % GHz
Resolution = 0.0001; % THz
f = 193.35 : Resolution : 193.45; % THz
```

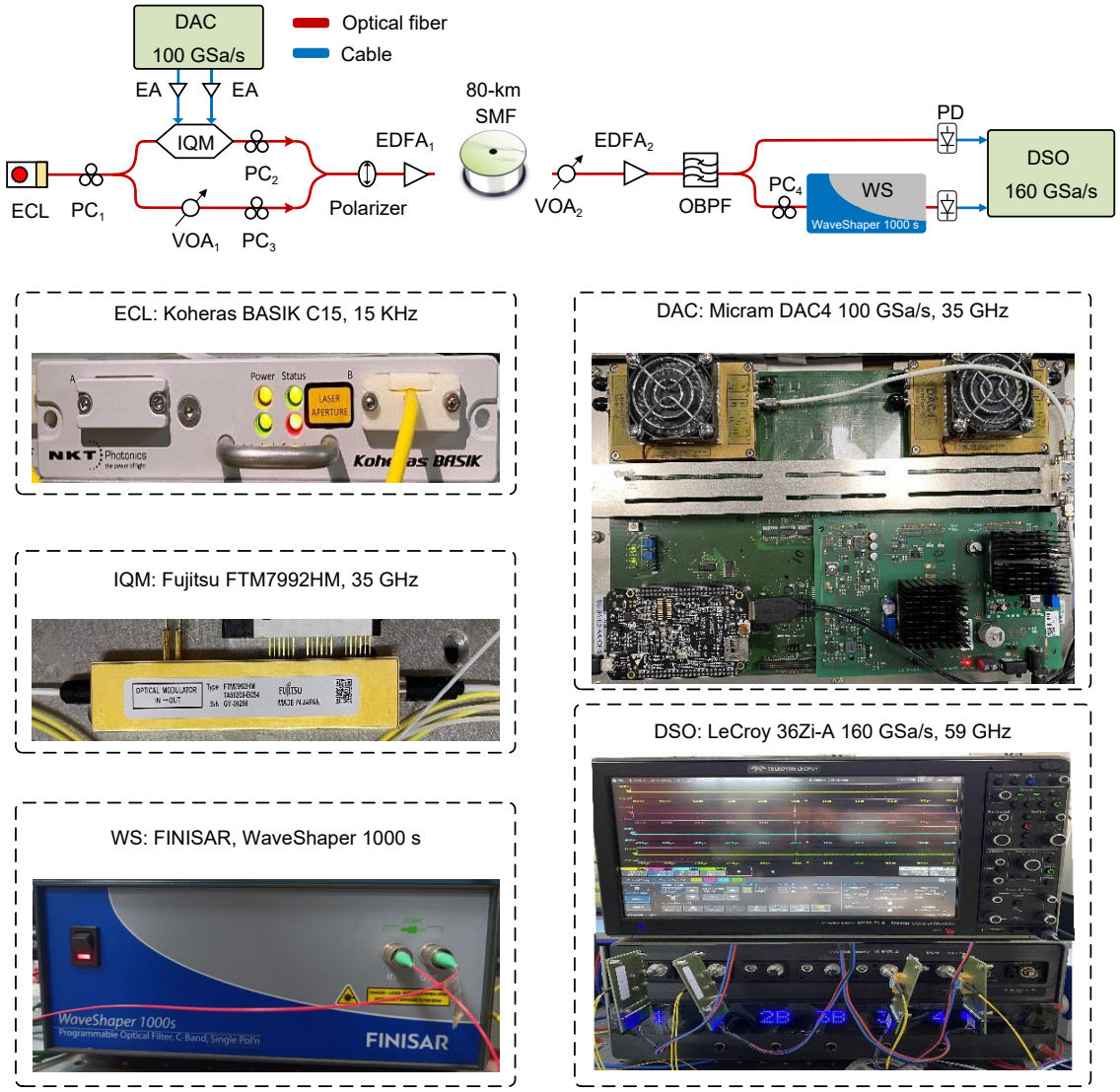

Fig. S3 | Experimental setup and photographs of the key devices. EA: electrical amplifier; ECL: external cavity laser; PC: polarization controller; EDFA: erbium-doped fiber amplifier; OBPF: optical bandpass filter; WS: WaveShaper; DSO: digital storage oscilloscope.

```
Central_number = (GuardBand.*1e-3)/Resolution + 1; % Central points
Edge_number = round(((max(f) - min(f))./Resolution + 1 - Central_number)./2); % Edge points
f = reshape(f,[],1);
Amp = zeros (length(f), 1); % dB
Pha = [linspace(Degr.*pi/180, Degr.*pi/180, Edge_number), linspace(0.*pi/180, 0.*pi/180, Central_number),...
linspace(Degr.*pi/180, Degr.*pi/180, Edge_number)];
Pha = reshape(Pha,[],1);
Port = linspace(1, 1, length(f));
Port = reshape(Port,[],1);
output = [f, Amp, Pha, Port];
%% Plot figure
figure;
plot(f, Amp, 'LineWidth', 2);
xlabel('Frequency (THz)'); ylabel('Attn (dB)');
set(gca,'XTick', [min(f):0.05:max(f)]);
```

```
set(gca,'YTick', -1:0.5:1);
set(gca, 'LineWidth', 0.5, 'XColor', [0 0 0]/255, 'YColor', [0 0 0]/255,...
'GridColor', [0 0 0]/255, 'GridLineStyle', '-', 'MinorGridColor', [0 0 0]/255, 'MinorGridLineStyle', ':');
set(gca, 'FontName', 'TimesNewRoman', 'FontSize', 20, 'FontWeight', 'normal');
set(gcf,'Position',[100 100 1000 600]);
figure;
plot(f, Pha, 'LineWidth', 2);
xlabel('Frequency (THz)'); ylabel('Phase (Rad)');
set(gca,'XTick', [min(f):0.05:max(f)]);
set(gca,'YTick', 0:0.4:1.6);
set(gca, 'LineWidth', 0.5, 'XColor', [0 0 0]/255, 'YColor', [0 0 0]/255,...
'GridColor', [0 0 0]/255, 'GridLineStyle', '-', 'MinorGridColor', [0 0 0]/255, 'MinorGridLineStyle', ':');
set(gca, 'FontName', 'TimesNewRoman', 'FontSize', 20, 'FontWeight', 'normal');
set(gcf,'Position',[100 100 1000 600]);
%% Save data
fileID=fopen(['Phase_shifter_',num2str(min(f)),'THz_to_',num2str(max(f)),'THz_',...
num2str(GuardBand),'GHz_GB_',num2str(Degr),'degree.wsp'],'w');
for i = 1: length(f)
fprintf(fileID,'%.4f %.3f %.3f %1.0f\n',f(i), Amp(i), Pha(i), Port(i));
end
```

Figure S4(a) and S4(b) illustrate the amplitude and phase responses of the optimal transfer function, respectively. The produced data are then loaded onto the WS. Figure S4(c) shows a preview of the optimal transfer function after data loading.

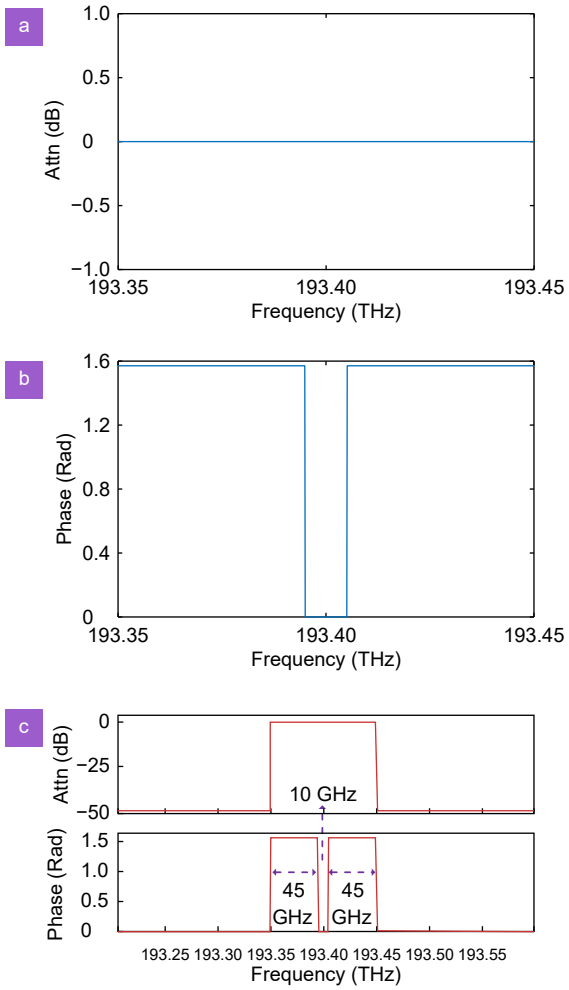

Fig. S4 | (**a**) Amplitude response of the optimal transfer function. (**b**) Phase response of the optimal transfer function. (**c**) Preview of the optimal transfer function after data loading.

References

S1. Chuang CY, Liu LC, Wei CC et al. Study of training patterns for employing deep neural networks in optical communication systems. In *European Conference on Optical Communication* 1–3 (IEEE, 2018); http://doi.org/10.1109/ECOC.2018.8535400.

S2. Li XF, Li JC, An SH et al. Deep-learning-enabled direct detection with reduced computational complexity and high electrical-spectral-efficiency. *J Lightwave Technol* **41**, 5495–5502 (2023).

S3. Chen X, Chandrasekhar S, Winzer PJ. Self-coherent systems for short reach transmission. In *European Conference on Optical Communication* 1–3 (IEEE, 2018); http://doi.org/10.1109/ECOC.2018.8535234.

S4. Li A, Peng WR, Cui Y et al. Single-λ 112Gbit/s 80-km transmission of PAM4 signal with optical signal-to-signal beat noise cancellation. In *Optical Fiber Communication Conference* (Optica Publishing Group, 2018); http://doi.org/10.1364/OFC.2018.Tu2C.5.

S5. Che D, Chen X, He JY et al. 102.4-Gb/s single-polarization direct-detection reception using signal carrier interleaved optical OFDM. In *Optical Fiber Communication Conference* (Optica Publishing Group, 2014); http://doi.org/10.1364/OFC.2014.Tu3G.7.

S6. Shieh W, Sun C, Ji HL. Carrier-assisted differential detection. *Light Sci Appl* **9**, 18 (2020).

S7. Mecozzi A, Antonelli C, Shtaif M. Kramers–Kronig coherent receiver. *Optica* **3**, 1220–1227 (2016).

S8. Che D, Sun C, Shieh W. Optical field recovery in stokes space. *J Lightwace Technol* **37**, 451–460 (2019).