

# ナノフォトニックコンピューティングの性能限界

川上 哲志<sup>1,a)</sup> 浅井 里奈<sup>2</sup> 小野 貴継<sup>1</sup> 本田 宏明<sup>1</sup> 井上 弘士<sup>1</sup> 北 翔太<sup>3,4</sup> 納富 雅也<sup>3,4</sup>

**概要:** ナノフォトニクス技術の発展に伴い、ナノフォトニック・デバイスは低遅延、小型、低消費エネルギー化が進められている。これまで、ナノフォトニック・デバイスの演算への応用事例は少なく、その潜在的な性能は未知な領域である。著者らは、ナノフォトニック・デバイスの演算応用に着目し、CMOSを凌駕する高性能・低消費電力な計算機システムの実現を狙う。本稿では、アナログ光ベクトル-行列演算器を対象にナノフォトニックコンピューティングの性能限界を解析する。具体的には、ノイズを考慮した光演算機の電力・性能限界をモデル化した。さらに、既存のCMOSニューラルネットワーク用アクセラレータと性能比較することで、光コンピューティングの優位性を明らかにするとともに、その潜在的な可能性を示す。

## 1. はじめに

情報処理通信技術の発展に伴い、組込みシステムからデータセンタ、スーパーコンピュータなど様々な階層で大量のデータが処理されるようになり、あらゆる社会システムの高性能化・効率化へ寄与している。特に、実世界から大量の情報を獲得し、高度な計算（分析・解析）に基づいて、再び実世界に情報をフィードバックするサイバーフィジカルシステム（CPS）は、医療・交通・金融といった様々な応用分野が期待され、今後もその応用範囲は拡大すると予想される。CPSの基盤となる計算機システムにおいては、大量のデータを高性能・低電力に処理することが重要な課題となる。

計算機システムの高性能化、低消費電力化は、CMOS微細加工技術に基づくプロセッサの演算能力の向上と光配線による大容量伝送技術によって牽引されてきた。マイクロプロセッサにおいては、CMOSトランジスタを極限まで微細化し高集積化することによって、回路内の寄生抵抗・容量を削減し、性能向上と低消費エネルギー化を両立してきた。しかしながら、これらの根幹技術である微細化が限界を迎えつつあり、さらなる劇的な改善は期待できない[9]。

一方で、近年のナノフォトニクス技術の発展により、フォトニック結晶に代表される光の伝搬を制御可能な物質の大規模集積化、小型化、低消費エネルギー化が可能になりつつある[11]。ナノフォトニック・デバイスは、かつては困難であった光の波長程度の極微小領域への光の閉じ込めを可能にすることが報告されており[17]、今後のさらなる発展が期待されている。これにより、ナノフォトニクス技術は、すでに実用化しているインターネット通信のみならず、CPU-メモリ間やコア間といった比較的短距離な通信にもその適応範囲を広げ、CMOS技術に対する優位性を示している[3]。しかしながら、ナノフォトニクス技術の演算処理応用はこれまで殆どされておらず、現状では未知なる領域である。

そこで、ナノフォトニクス技術に基づいた光演算処理を含む光コンピューティング手法を検討し、CMOS技術を凌駕する高性能・低消費電力な計算機システムの実現を目指す。本稿では、ナノフォトニックコンピューティングの優位性ならびに将来展望を明らかにすべく、その電力効率の限界について解析を行った。光演算器としては、様々な情報処理の基本となるベクトル-行列演算器に着目し、ナノフォトニック・デバイスの性質に基づくアナログ方式での演算を実現可能なMZI-VMM (Mach-Zehnder Interferometer Vector by Matrix Multiplier) に着目した。光波の情報密度を最大限に活かすことのできるアナログ処理方式は、依然としてCMOS回路に対し集積度の劣るナノフォトニック・デバイスの欠点を補い、高い演算処理性能を達成できる有効な処理方式であると考えられる。その一方で、ある一定精度以上の演算精度を担保することを考えるとノイズは極めて重要な性能律速要因となる。したがって、ノイズを考慮し

<sup>1</sup> 九州大学 大学院システム情報科学研究院  
Faculty of Information Science and Electrical Engineering, Kyushu University

<sup>2</sup> 九州大学 大学院システム情報科学府  
Graduate School of Information Science and Electrical Engineering, Kyushu University

<sup>3</sup> NTT ナノフォトニクスセンタ  
NTT Nanophotonics Center

<sup>4</sup> NTT 物性科学基礎研究所  
NTT Basic Research Laboratories

a) satoshi.kawakami(at)cpc.ait.kyushu-u.ac.jp

た MZI-VMM の電力効率のモデルを構築した。構築したモデルに基づき検証した結果、現行の光デバイスを用いた際の MZI-VMM は高い電力効率を達成可能であることが明らかになった。また、今後の光デバイスの発展を想定した際の光コンピューティングの電力効率限界について解析した結果、現行の光デバイスを用いた場合と比較して最大で 152 倍の電力効率改善が見込めることが分かった。さらに、MZI-VMM の電力効率の改善に大きく影響を与える光デバイスパラメータを明らかにした。これは光演算器を前提とした際の光デバイスのパラメータ改善の指針を示す重要な知見である。

## 2. 光コンピューティングの問題点と今後

現在、光通信はインターネット社会を支える極めて重要な要素技術として広く普及している。また、広域通信のみならず、計算ノード間接続やチップ内コア間接続（いわゆる Network-on-Chip）といったコンピュータ・システム内部の比較的狭域な通信路においても光の適用が進みつつある [2], [18], [20]。その一方、光の特性をデータ処理に利用する「光コンピューティング」の研究開発も盛んに行われた次期があったが、90 年代後半以降は衰退の一途を辿っているのが実状である。そこで本節では、これまでの光コンピューティング研究の変遷を鑑み、問題点を整理し、本研究が目指す「ナノフォトニックコンピューティング」の方向性を議論する。

### 2.1 問題点

これまでに光コンピューティングに関する多くの研究開発が行われてきたが、CMOS による電子式コンピュータを凌駕する性能を達成するには至っていない。その最大の原因は、現代の電子式コンピュータと同じ高性能化手法を指向するアーキテクチャ的アプローチにあると考える。空間系光アナログ/デジタルコンピューティングでは、処理機構を物理的に空間並列配置した SIMD 処理を高性能化の拠り所としている。一般に、SIMD 処理における実効性能は、「演算器の数（演算並列度）」とそれに見合った「入出力バンド幅」に依存する。空間伝搬光を活用した並列処理（スケールアウト型電子並列処理）では、自由空間データ通信により十分な入出力バンド幅を提供できる一方、演算素子のスケールアップには光学装置の大規模化や増加が必要となる。これは、小型化が求められる近年の情報処理システムにおいては、非常に厳しい要求となる。これに対し、電子式コンピュータにおいては、ムーアの法則に従って搭載する演算器数を順調に増加してきた。ここで、チップサイズ一定で搭載する演算器数を増大するアプローチをスケールアップ型電子並列処理と呼ぶ。また、メモリチャネル数の増加やメモリバス動作周波数の向上、階層メモリ構造の採用など様々な工夫により高い入出力バンド幅を実現して

いる。その結果、スケールアウト型光並列処理とスケールアップ型電子並列処理の性能差は「演算並列度」に大きく依存する形となり、半導体集積度の向上を拠り所とする後者が圧倒的優位性を獲得したのである。これに加え、光アナログ処理から光デジタル処理への転換も負の要因である。集積度の観点では半導体に対し圧倒的な差が存在するにも関わらず、機能レベルを電子式コンピュータと同程度にまで低下させたため、光本来が有する利点を失う結果を招いた。その一方で、光の特性上、DRAM や SRAM といった半導体メモリのような高い制御性を持つ記憶素子の実現や、複雑な順序回路の実装は依然として難しく、光デジタル処理方式導入の狙いの 1 つであった制御容易性の向上には十分貢献できていない。以上を整理すると、演算精度や可制御性の向上を目的として光デジタルコンピューティングを導入した結果、電子式コンピュータと比較して演算粒度は同レベルとなり、かつ、実現可能な並列度に圧倒的な差が生じたことが、光コンピュータの可能性を制限している最大の問題である。

### 2.2 ナノフォトニックコンピューティングの方向性

光の利点を最大限に活用すると同時に欠点を隠蔽し、電子式コンピュータを凌駕する高い性能や電力効率を達成するには、デバイス/アーキテクチャレベルでのコデザインが必要不可欠である。これまでの光コンピューティングの研究を鑑みると計算機アーキテクチャの観点では、以下に関する検討が重要となる。

- **光アナログ処理の導入**：ナノフォトニクス導入により集積化光コンピューティングが可能となるが、依然として電子式トランジスタの微細化に対しては 3~4 桁の開きがあり、これはそのままスループットの差として顕在化する。このギャップを埋めるためには次元の異なる最適化が必要であり、その有効な手段として光アナログ処理による高機能演算の実現が挙げられる。
- **不完全計算モデルの導入**：光アナログ処理では演算精度の低下が最大の問題となる。そこで、このような欠点を許容するコンピューティング・モデル（例えば、アプロキシメイト・コンピューティング）の導入が必要不可欠となる。アプロキシメイト・コンピューティングは、演算精度を犠牲にして演算処理量を削減することで性能向上や消費電力削減を達成する処理形態である。FFT、JPEG エンコーディング、k-means クラスタリングや sobel フィルタによるエッジ検出といった演算をアプロキシメイト・コンピューティングに適応する例が報告されている [8]。
- **極めて高い光入出力バンド幅の活用**：過去の光コンピュータと同様、光による高い入出力バンド幅（低レイテンシ化も可能）を有効活用すべきである。特に、光通信路に直結した光速処理機構を搭載することで、

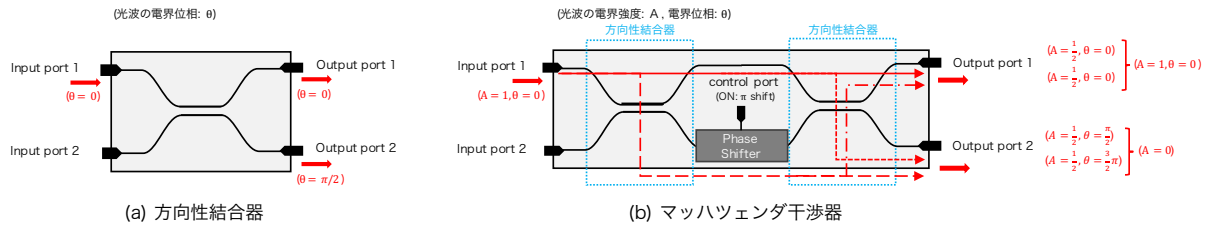


図 1: 基本光素子

光電変換のオーバーヘッドを隠蔽しつつ、極めて低いレイテンシでの情報処理が可能となる。すなわち、光通信中の情報処理 (In-Optical-Network Computing) の実現である。

- **光演算における多重化技術の適用**: 光通信技術で用いられる波長多重技術 (DWDM: Dense Wavelength Division Multiplexing) をナノフォトニック・デバイスによる演算処理へ適用することで、同時刻に同一デバイスで MIMD もしくは SIMD と同等の機能が実現可能となる。これは、異なる波長の光信号はお互いに干渉しないという性質を有効に活用しているため、従来の電気回路では考えられなかったことである。

これらの条件を満たす計算機システムとして、著者らはナノフォトニック・ニューラルアクセラレータを提案している [21]。一方で、デバイスの観点ではどのような特性・性能を改善することが計算機システムとしての性能向上に効果的か明らかではない。一般的に、デバイスの各性能は材質や構造によってトレードオフの関係にある。たとえば、位相変調器の素子長を小さくすると変調速度が遅くなる傾向にある。そこで、本稿では各デバイスパラメータの改善がどの程度システムレベルでの性能向上に寄与するかを解析する。これにより、デバイス/アーキテクチャレベルでのコデザインが加速し、より高性能・低消費電力な計算機システムの構築が可能になる。

### 3. 光ベクトル-行列演算器

光コンピューティングにおけるベクトル-行列演算器 (VMM: Vector-by-Matrix Multiplier) は、様々な方式で提案されている。Stanford VMM では、空間伝搬光が行列要素と対応付いた変調器を通過することでベクトル-行列積を実現する [5]。また、WDM-VMM では波長多重と波長毎に対応したリング共振器により各行列要素の積を実現した上で、受光器にて全波長の光強度を合算観測することでベクトル-行列積を実現している [16]。本稿では、信号媒体としてコヒーレント光を用いることにより行列要素として負の数も扱うことが可能な MZI-VMM [7], [13] に着目し、性能解析を行う。本節では、まず MZI-VMM の構成要素となる基本的な光素子について紹介し、次に MZI-VMM の動作原理について説明する。

### 3.1 基本素子

#### 方向性結合器

方向性結合器 (DC: Directional Coupler) とは、1つの光の伝達経路 (以下、導波路) からの光信号を2つの導波路に分岐したり、あるいは2つの導波路からの光信号を1つの導波路に結合する機能を有するデバイスである。図 1 (a) のように2つの導波路を十分に近い距離で平行に並べると、光波は2つの導波路間を移動する。導波路型の DC においては、光波が導波路間を移動する際に位相が  $\pi/2$  シフトするという特徴を持つ。

#### 位相シフタ

位相シフタ (PS: Phase Shifter) は、制御ポートからの信号によって入力光波の位相を任意に変化させるという特徴を有する光デバイスである。

#### マッハツェンダ干渉器

マッハツェンダ干渉器 (MZI: Mach-Zehnder Interferometer) の概略を図 1 (b) に示す。導波路型 MZI の回路は、図 1 (a) の DC と PS から構成される。MZI は、PS よって生じた位相差に従って出力光の強度を変化させるという特徴を有する。たとえば、2つの DC が入射光強度を等分配させ、PS は制御信号により位相を  $\pi$  シフトする (ON 時) もしくは位相シフトしない (OFF 時) という2値制御が可能な構成を考える。この際、入力ポート1から入力された光波は、PS が ON 時に出力ポート1から出力され、逆に OFF の時には出力ポート2から出力される経路切り替えスイッチとしての機能を果たす。これは、図 1 (b) に示すように、2つの DC によって分割された光波が出力ポートにて強め合う (もしくは、弱め合う) ことで実現されている。

### 3.2 MZI-VMM の動作原理

MZI-VMM は、2つのユニタリ変換回路とアッテネータ (もしくはアンプ) によって構成される。これは、任意の行列を特異値分解により2つのユニタリ行列と1つの対角行列に分解することに対応する。すなわち、任意の  $M \times N$  行列  $A$  を式 (1) の通り分解できる。

$$A = U \Sigma V \quad (1)$$

ここで、 $U$  は  $M \times M$  ユニタリ行列、 $V$  は  $N \times N$  ユニタリ行列を表す。また、 $\Sigma$  は  $M \times N$  行列であり非対角要素は0、かつ、対角要素は非負で降順の特異値 ( $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ )

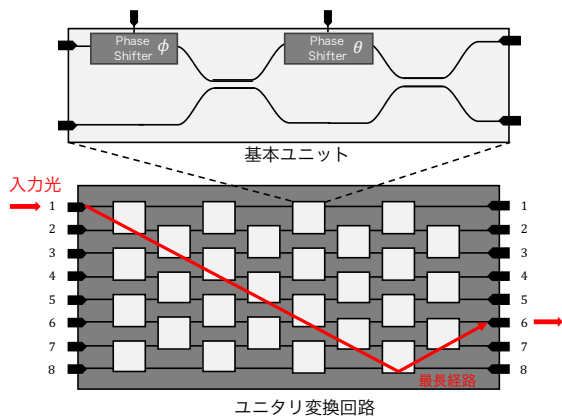


図2: MZIによるユニタリ変換回路の構成図.

を持つ行列である.  $r$ は  $A$ のランクに等しい. 所望のベクトル-行列演算を  $y = Ax$ とすると, 特異値分解により  $y = U\Lambda Vx$ と表すことができる. したがって, 各入力ポートの光波をベクトルの要素に対応付けし, ユニタリ変換回路, 光アンプ, ユニタリ変換回路の順で通過させることで任意のベクトル行列演算が可能となる.

光デバイスによるユニタリ変換回路の実現方法とその原理は, 1994年に Reck らによって提案されている [13]. 本稿では, Reck らの回路と比較し, 実装面積が小さく, 入力光波の通過する最長経路が短く, 回路忠実度が高い Clements らのユニタリ変換回路実装方式に着目する. この回路の詳細な動作原理とユニタリ変換回路となる数学的証明については, 文献 [7]を参照されたい. 図2に,  $8 \times 8$ ユニタリ変換回路の構成図を示す. 図1(b)に示す MZIの入力ポートに PSを追加したものを基本ユニットとし, その組合せで実現されている. 基本ユニットは,  $2 \times 2$ のユニタリ変換を実現しており, このユニットの直列接続は  $2 \times 2$ ユニタリ行列の積を意味する. すなわち, 基本ユニットの組合せで表現された回路はユニタリ変換回路となる. また, 基本ユニットで表現されるユニタリ行列を左(右)からかけることにより, 特定の2行(列)の任意の要素を0にすることができる. この要素消去によって, 任意の行列は対角化することができる. 対角化されたユニタリ行列は単位行列であるため, 任意のユニタリ行列はこの基本ユニットの組合せで実現できる.

#### 4. 光演算器の性能モデル

著者らは, MZI-VMMの性能モデルについて文献 [21]で提案している. この性能モデルにおいては, MZI-VMM回路遅延(演算器の最長経路), PSの動作周波数, 受光器(PD: Photo Detector)の動作周波数によって演算器性能が律速される. しかしながら, アナログ演算器において重要なノイズの影響を考慮出来ていない. ある一定精度以上の演算精度を担保することを考えると, ノイズは極めて重要な性能律速要因となる. ノイズの影響を性能モデルに組み込むこと

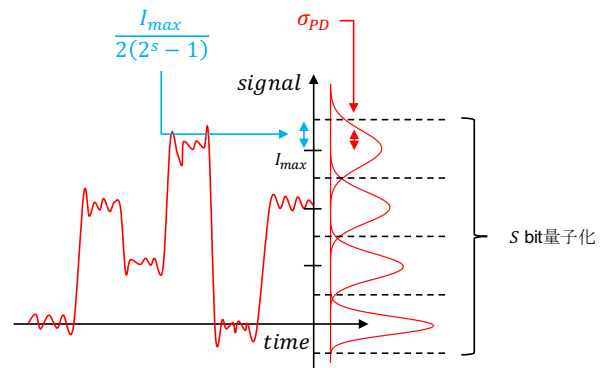


図3: 受光器の信号電流と揺らぎ

により, ある演算精度を保障した上での性能上限の見積りが可能となる. 本節では, まず光源と受光器のみによって構成される系, すなわち, 光通信系でのノイズの影響をモデル化した後に, MZI-VMMを含む光演算器としての性能モデルについて説明する.

#### 4.1 光通信系におけるノイズの影響

光源(LS: Light Source)とPDを導波路で接続した系における受信光強度の揺らぎ(ノイズの分散)は以下の3つに大別できる.

- (1) 信号光波のショットノイズ ( $\sigma_s^2$ ): 光子の出現確率の不確定性に起因する.
- (2) PDの暗電流によるショットノイズ ( $\sigma_d^2$ ): 信号光波が無の場合でもPDに流れる暗電流の電子のゆらぎ.
- (3) PDの電気回路の熱雑音 ( $\sigma_{th}^2$ ): 抵抗体内の電子の不規則な熱振動によって生じる.

この系におけるPDで観測される信号の揺らぎ  $\sigma_{PD}^2$ は式(2)で表される.

$$\sigma_{PD}^2 = \sigma_s^2 + \sigma_d^2 + \sigma_{th}^2 \quad (2)$$

$$\sigma_s^2 = 2q^2 \eta \left( \frac{P_{PD}}{h\nu} \right) \frac{f_{PD}}{2} M^{2+x} \quad (3)$$

$$\sigma_d^2 = 2qI_d \frac{f_{PD}}{2} \quad (4)$$

$$\sigma_{th}^2 = \frac{4kT}{R_L} \frac{f_{PD}}{2} \quad (5)$$

ここで,  $q$ は電子素量,  $\eta$ は量子効率,  $P_{PD}$ はPDの最大消費電力,  $h$ はプランク定数,  $\nu$ は信号光波の周波数,  $f_{PD}$ はPDの動作周波数,  $M$ はADP(Avalanche Photo Diode)の電流増倍係数,  $x$ はAPDの過剰雑音指数,  $I_d$ はPDの暗電流,  $k$ はボルツマン係数,  $T$ はPDの絶対温度,  $R_L$ はPDの電気回路の負荷抵抗である. 信号光波のショットノイズの分散は, 式(3)で示す通り信号強度に比例する.

本稿では最悪条件での揺らぎに対する演算精度を保障すべく, 最大信号強度(最大消費電力)におけるノイズをモデル化する. PDにおいて最大信号を観測した際の電流を  $I_{max}$ とし  $s$  bitで量子化する場合を想定すると, 観測された電流値  $i$ が  $I_{max} - I_{max}/2(2^s - 1) < i \leq I_{max} + I_{max}/2(2^s - 1)$ を

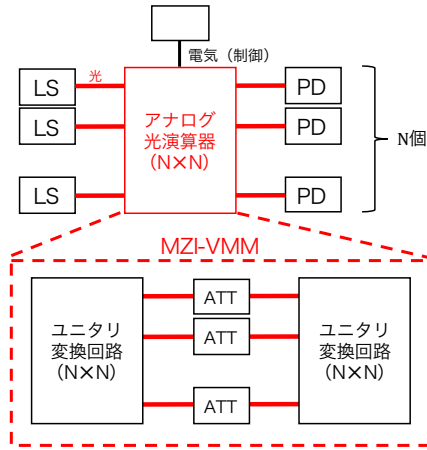


図 4: MZI-VMM によるアナログ光演算器

満たした場合に  $I_{max}$  相当の信号として認識される (図 3)。つまり,  $I_{max}/2(2^s - 1)$  が  $\sigma_{PD}$  よりも十分に大きい場合は, 誤った値として識別される確率が小さくなる。この考え方を一般化し, 式 (2)-(5) にて整理すると式 (6) で表すことができる。

$$\frac{q\eta(\frac{P_{PD}}{h\nu})M}{2(2^s - 1)} \geq \alpha \sqrt{2q^2\eta(\frac{P_{PD}}{h\nu})\frac{f_{PD}}{2}M^{2+s} + 2qI_d\frac{f_{PD}}{2} + \frac{4kT}{R_L}\frac{f_{PD}}{2}} \quad (6)$$

ここで, 左辺の  $\alpha$  以外の項は PD で検出される電流揺らぎの標準偏差を示しており, 右辺は最大信号を s bit で量子化するための量子化幅を表している。  $\alpha$  は 0 より大きい値を持つ誤差保障係数であり, s bit 量子化を実現するために担保すべき「信号揺らぎの標準偏差に関するマージン」である。

## 4.2 MZI-VMM の性能モデル

本節では, 図 4 で示す MZI-VMM を有する光演算器のノイズを考慮した性能モデルを説明する。なお, 3.2 節で述べたように MZI-VMM では任意の行列-ベクトル演算が可能であるが, 簡単のために  $N \times N$  の任意の行列のみを対象とする。この系においては, 4.1 節で説明した (1) (2) (3) のノイズのみならず, PS の制御ノイズに起因する MZI-VMM の出力信号の揺らぎ  $\sigma_{VMM}$  が加わる。  $\sigma_{VMM}$  が最大となる条件を明らかにすべく, まず図 2 に示す基本ユニットの出力信号に最大のノイズが現れる条件を考える。基本ユニットの入力光の電界振幅を  $E_{11}, E_{12}$ , 出力光の電界振幅を  $E_{01}, E_{02}$  とすると, 入出力光波の関係は伝達行列を用いて式 (7) で示せる。

$$\begin{pmatrix} E_{01} \\ E_{02} \end{pmatrix} = e^{i(\frac{\theta+\pi}{2})} \begin{pmatrix} e^{i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \\ e^{i\phi} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} E_{11} \\ E_{12} \end{pmatrix} \quad (7)$$

ここで, 伝達行列はユニタリ行列であるため, 入力光の電界振幅のノルムの和 (入力光の強度の和) は出力光の電界振幅のノルムの和に等しい。すなわち, 基本ユニットにおいて (理想的には) エネルギーは保存される。これは, 片方の出力ポートから観測される信号に最大の揺らぎが観測

される場合に逆側の出力ポートからも最大ノイズが観測されることを意味する。したがって, 片方の出力ポートからの信号値にのみ着目し, 最悪条件を考察する。

PD によって検出される信号値は電界強度  $|E_{01}|^2$  に比例するため, 電界強度の傾きが最大 (もしくは最小) となる場合が最大ノイズが検出される条件となる。基本ユニットを構成する 2 つの PS のノイズがガウス分布  $N(0, \sigma_{PS}^2)$  にしたがうとすると, 基本ユニットの揺らぎは式 (8) で表される。

$$\begin{aligned} & \left\{ \frac{\partial |E_{01}|^2}{\partial \theta} \right\}^2 \sigma_{PS}^2 + \left\{ \frac{\partial |E_{01}|^2}{\partial \phi} \right\}^2 \sigma_{PS}^2 = \\ & \sigma_{PS}^2 \left\{ \frac{1}{4} (E_{12}^2 - E_{11}^2)^2 + E_{11}^2 E_{12}^2 (\sin^2 \phi \sin^2 \theta + \cos^2 \phi \cos^2 (\theta + \theta_0)) \right\} \\ & \sin \theta_0 = \frac{E_{12}^2 - E_{11}^2}{\sqrt{(E_{12} - E_{11})^2 + 4E_{11}^2 E_{12}^2 \cos^2 \phi}} \\ & \cos \theta_0 = \frac{2E_{11} E_{12} \cos \phi}{\sqrt{(E_{12} - E_{11})^2 + 4E_{11}^2 E_{12}^2 \cos^2 \phi}} \end{aligned} \quad (8)$$

ここで, 基本ユニットに対する 2 つの入力パターンにおけるノイズが最大となる条件を考える。(A)  $E_{E1} = E_{E2} = E_{max}$  ( $E_{max}$  は各 LS の最大振幅) の場合に式 (8) が最大となる条件は,  $\phi = k\pi$  ( $k = 0, 1, 2, \dots$ ) かつ  $\theta = l\pi$  ( $l = 0, 1, 2, \dots$ ), または,  $\phi = (2k - 1)\pi/2$  ( $k = 1, 2, 3, \dots$ ) かつ  $\theta = (2l - 1)\pi/2$  ( $l = 1, 2, 3, \dots$ ) である。この際,  $E_{max}^4 \sigma_{PS}^2$  のノイズが検出されることが分かる。また, (B)  $E_{E1} = 0$ , または  $E_{E2} = 0$  の場合は,  $\phi, \theta$  の値に依らず  $\frac{E_{11}^4}{4} \sigma_{PS}^2$ , または  $\frac{E_{12}^4}{4} \sigma_{PS}^2$  のノイズが観測される。

次に,  $N \times N$  の MZI-VMM におけるノイズの最悪条件を考える。4.1 節で述べた通り, 信号光波のショットノイズの分散は信号強度に比例するため, 全ての LS のエネルギーがある 1 つの出力ポートに集約される場合にノイズが最大となる。すなわち, 最悪条件である出力ポートの信号強度は式 (9) の通りである。

$$P_{LSall} T_{MZI}^{2N} \quad (9)$$

ここで,  $P_{LSall}$  は図 4 の系における全ての LS の最大消費電力の和であり,  $T_{MZI}$  は基本ユニット 1 段を通過した際の透過率 (= 1 - 損失率) である。図 2 に示すユニタリ変換回路において最長経路を通過した際の基本ユニット段数は  $N$  であり, また図 4 (3.2 節) で示す通り MZI-VMM では 2 つのユニタリ変換回路を用いるため, 最大で  $2N$  段の基本ユニットを通過する。ここで,  $N \times N$  の MZI-VMM において, 最長経路上に存在する基本ユニットが上述した最悪条件を満たし続けた場合を「揺らぎの最悪ケース」とする。図 5 に  $8 \times 8$  ユニタリ変換回路が最悪条件を満たす際の各段の信号強度を示す。1, 2, 4 段目の基本ユニットは上述の (A) パターンに相当し, その他の基本ユニットは (B) の場合になることがわかる。  $N$  を 4 以上の 2 の冪数に限定し, MZI-VMM の最大の揺らぎを一般化すると式 (10) で示される。

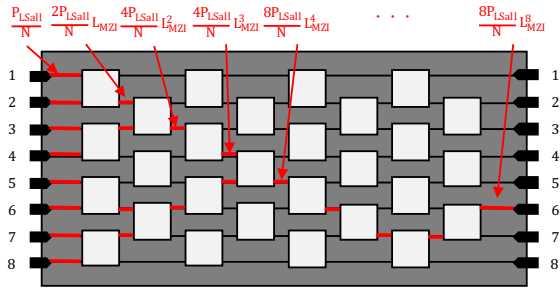


図 5: ユニタリ変換回路における各段の信号強度

$$\sigma_{VMM}^2 = \left(\frac{P_{LSall}}{N}\right)^2 \sigma_{PS}^2 + \frac{1}{4} \left(\frac{P_{LSall}}{N}\right)^2 \sigma_{PS}^2 \sum_{k=1}^{\log_2 N} 4^k T_{MZI}^{2k} \frac{(1 - T_{MZI}^{2k})}{1 - T_{MZI}^{2k}} \quad (10)$$

$$+ \frac{3}{4} \left(\frac{P_{LSall}}{N}\right)^2 \sigma_{PS}^2 \sum_{k=1}^{(\log_2 N)-1} 4^k T_{MZI}^{2^{k+1}-2} + \frac{1}{4} P_{LSall}^2 \frac{T_{MZI}^{2N}(1 - T_{MZI}^{2N})}{1 - T_{MZI}^{2N}} \sigma_{PS}^2$$

以上より、式 (6) を拡張し式 (9) と式 (10) を用いることで、MZI-VMM の性能モデルは式 (11) で表せる。

$$\frac{q\eta \left(\frac{P_{LSall} T_{MZI}^{2N}}{h\nu}\right) M}{2(2^s - 1)} \geq \alpha \sqrt{2q^2 \eta \left(\frac{P_{LSall} T_{MZI}^{2N}}{h\nu}\right) \frac{f_{PD}}{2} M^{2+x} + 2qI_d \frac{f_{PD}}{2} + \frac{4kT}{R_L} \frac{f_{PD}}{2} + \sigma_{VMM}^2} \quad (11)$$

## 5. モデルに基づく電力効率限界推定

4 節で説明した性能モデルに基づき MZI-VMM 光演算器の電力効率を見積もる。式 (11) の  $\alpha$  を特定の値に設定することで MZI-VMM の演算誤差の上限を規定し、電力効率について式を整理することで電力効率の上限を見積もることが可能になる。MZI-VMM は PS の制御信号を固定してベクトル-行列演算を行う実行モデルと PS 制御信号が可変な実行モデルによってスループットの性能モデルが異なる。そこで本節ではそれぞれの実行モデルを対象に解析する。なお、スループットの性能モデルの詳細は文献 [21] を参照されたい。

### 5.1 MZI-VMM のスループットモデル

まず、PS の制御信号を固定した際のスループットについて考える。PS の制御信号を固定して MZI-VMM を駆動させることは、行列要素を固定したベクトル-行列演算を行うことと等しい。本実行モデルにおいてスループット  $throughput$  [OPs/s] は、式 (12) で示される。

$$throughput = N \times N \times \min(f_{LS}, f_{PD}) \quad (12)$$

ここで、 $f_{LS}$  は光源の周波数であり、スループットの単位である 1OP (Operation) は 1 積和演算を意味する。LS と PD の動作周波数は等しいとすると、 $throughput = N^2 f_{PD}$  である。この演算器系において、LS で生じた光波のエネルギーは MZI やアッテネータを経て PD で全て失われる (変換効率 0) とすると消費電力は  $P_{LSall}$  となる。PS の消費電力は、電気光学効果の微小な移相器を用いることで相対的に無視できるレベルにすることができる [15]。したがって、

表 1: 性能モデルの固定パラメータ一覧

Symbol	Description	Value
$q$ [C]	電気素量	$1.6021766208 \times 10^{-19}$
$h$ [Js]	プランク定数	$6.62607004 \times 10^{-34}$
$c$ [m/s]	光速	299792458
$\lambda$ [ $\mu$ m]	光の波長	1.55
$\eta$ [-]	量子効率	0.7
$k$ [J/K]	ボルツマン定数	$1.38064852 \times 10^{-23}$
$T$ [K]	絶対温度	300
$R_L$ [ $\Omega$ ]	負荷抵抗	50
$I_d$ [nA]	暗電流	1
$M$ [-]	電流増倍係数	1
$x$ [-]	過剰雑音指数	1
$\sigma_{PS}^2$ [-]	PS の揺らぎ	$10^{-12}$
$\alpha$ [-]	誤差保障係数	3
$s$ [bit]	ビット数	8

表 2: 性能モデルの可変パラメータ

Symbol	Description	Current	Advanced
$L_{MZI}$ [ps]*	基本ユニットのレイテンシ	1	0.01
$f_{PS}$ [GHz]*	PS の周波数	12.5	40
$P_{LSall}$ [dBm]	全 LS の最大電力	17	30
$f_{PD}$ [GHz]	PD の周波数	40	100
$N$ [-]	MZI-VMM の規模	6	1024
$T_{MZI}$ [-]	基本ユニットの透過率	0.9	0.99

\*行列要素可変 MZI-VMM のみで使用

式 (11) により電力効率限界は式 (13) で表すことができる。

$$\frac{N^2 f_{PD}}{P_{LSall}} \leq \frac{2N^2}{P_{LSall}} \left\{ \frac{\left( q\eta \left( \frac{P_{LSall} T_{MZI}^{2N}}{h\nu} \right) M \right)^2}{4(2^s - 1)^2 \alpha^2} - \sigma_{VMM}^2 \right\} \times \frac{1}{2q^2 \eta \left( \frac{P_{LSall} T_{MZI}^{2N}}{h\nu} \right) M^{2+x} + 2qI_d + \frac{4kT}{R_L}} \quad (13)$$

次に、PS の制御信号が可変な実行モデルについて考える。PS の制御信号を可変して MZI-VMM を駆動させることは、任意のベクトルと任意の行列要素の積を行うことに等しい。PS の制御信号を変えながらベクトル-行列演算を実行する場合、スループット  $throughput$  [OPs/s] は、式 (14) で示される。

$$throughput = N \times N \times \min\left(f_{LS}, \frac{1}{(L_{MZI} + 1/f_{PD})}, f_{PS}\right) \quad (14)$$

PS の制御信号が固定の場合と同様に、LS と PD の動作周波数が等しいとすると、スループットは「PS の周波数が律速する場合」と「MZI-VMM レイテンシと PD 遅延時間の和の逆数が律速する場合」に分けられる。前者の場合、電力効率限界は式 (15) で表せる。

$$\frac{N^2 f_{PS}}{P_{LSall}} \leq \frac{f_{PS}}{f_{PD}} \frac{2N^2}{P_{LSall}} \left\{ \frac{\left( q\eta \left( \frac{P_{LSall} T_{MZI}^{2N}}{h\nu} \right) M \right)^2}{4(2^s - 1)^2 \alpha^2} - \sigma_{VMM}^2 \right\} \times \frac{1}{2q^2 \eta \left( \frac{P_{LSall} T_{MZI}^{2N}}{h\nu} \right) M^{2+x} + 2qI_d + \frac{4kT}{R_L}} \quad (15)$$

後者の場合、電力効率限界は式 (16) で表すことができる。

表 3: 行列要素固定 MZI VMM の電力効率上限

Configuration No.	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Take advanced parameters ?	$P_{LSall}$		○				○	○	○			○	○	○		○
	$f_{PD}$			○			○			○	○	○	○		○	○
	$N$				○			○		○		○	○		○	○
	$T_{MZI}$					○			○		○	○		○	○	○
電力効率 [TOPs/W]	22	22	23	25	67	23	26	67	25	70	2043	27	72	2700	2430	2820
電力効率比 * Baseline	1x	1x	1x	1.1x	3x	1x	1.2x	3x	1.1x	3.2x	111x	1.2x	3.3x	123x	111x	128x

表 4: 行列要素可変 MZI VMM の電力効率上限

Configuration No.	1*	2	3	4	5	6	7	8	9	10	11	12	13	14
Take advanced parameters ?	$L_{MZI}$		○		○		○		○		○		○	○
	$f_{PS}$			○	○			○	○			○	○	○
	$P_{LSall}$									○	○	○	○	○
	$f_{PD}$									○	○	○	○	○
	$N$					○	○	○	○	○	○	○	○	○
	$T_{MZI}$					○	○	○	○	○	○	○		○
電力効率 [TOPs/W]	18	18	18	22	1200	2220	1200	2340	1200	2430	1200	28	70	2740
電力効率比 * Baseline	1x	1x	1x	1.2x	67x	123x	67x	130x	67x	135x	67x	1.6x	3.9x	152x

$$\frac{N^2 \frac{1}{2NL_{MZI} + 1/f_{PD}}}{P_{LSall}} \leq \frac{2N^2}{P_{LSall}} \left\{ \frac{(q\eta \frac{P_{LSall} T_{MZI}^{2N}}{hc/\lambda} M)^2}{4(2^x - 1)^2 a^2} - \sigma_{VMM}^2 \right\} \times \frac{1}{1} \quad (16)$$

$$\frac{1}{2q^2 \eta \left( \frac{P_{LSall} T_{MZI}^{2N}}{hc/\lambda} M \right)^{2+x} + 2qI_d + \frac{4kT}{R_L} + 4NL_{MZI}} \left\{ \frac{(q\eta \frac{P_{LSall} T_{MZI}^{2N}}{hc/\lambda} M)^2}{4(2^x - 1)^2 a^2} - \sigma_{VMM}^2 \right\}$$

## 5.2 パラメータ

表 1 に式 (13), および式 (15), 式 (16) で用いる固定パラメータの一覧を示す. 本稿の評価において受光器は PIN-PD の使用を想定するため,  $M = 1$  かつ  $x = 1$  とし, 8bit で量子化を行うとする. また,  $\alpha = 3$  は演算誤差率 0.3% に相当する. PIN-PD とは, PN 接合間に真性半導体を挟み込んだ PD であり, 光通信において最もよく使われる受光器である. PS は電気によって制御されているため, 300K の環境温度かつ, 50Ω の負荷抵抗を有する電気回路の熱雑音と同じオーダーの雑音として  $\sigma_{PS}^2 = 10^{-12}$  とした. 表 2 は式 (13) の可変パラメータの一覧を示す. Current 列には現在のデバイス技術によって実現可能なデバイスパラメータを示し, Advanced 列にはナノフォトニクス技術の進展により各パラメータが改善されたと仮定した際の値を示す. 表 2 に示すデバイスパラメータは取りうる値の上限を表す. なお, Current なパラメータは, 実在する光デバイスや現在のナノフォトニクス技術によって実現可能な値に基づいている [1], [4], [12], [19], [21].

## 5.3 MZI-VMM の電力効率限界

まず, PS 制御信号が固定の場合における電力効率を求める. 可変パラメータを様々な組合せで改善させた際の最適な電力効率上限を表 3 に示す. Advanced パラメータを選択した場合を○で標記しており, Configuration No.1 は全て Current パラメータである. No.2-5 を見ると各パラメータの個別の改善では, あまり電力効率向上に寄与しないこと

が分かる. 一方で, No.11, 14, 15, 16 では劇的な電力効率向上を達成している. これらに共通している点は, VMM の規模と基本ユニットの透過率が改善していることである. また, No.12, 13 にて VMM の規模もしくは基本ユニットの透過率のどちらかが改善されない場合も電力効率が改善されていないことが分かる. したがって, ナノフォトニクス技術により VMM の規模と基本ユニットの透過率の改善が共に重要であることが分かる.

次に, PS 制御信号が可変な場合の電力効率を求める. 固定パラメータ並びに可変パラメータは表 1 と表 2 に各々示す. また, 可変パラメータの種々組合せにおける最適な電力効率を表 4 に示す. No.2-4 より  $2 \times 2$ MZI のレイテンシもしくは PS の周波数の改善では, 劇的な電力効率向上改善は見込めない. 行列要素固定 MZI - VMM の結果と同様に, VMM の規模と透過率を同時に改善することは電力効率向上に大きく寄与する (No.5-11, 14). さらに, No.6, 8, 10 より VMM の規模と透過率が改善された状況においては, 基本ユニットのレイテンシ (素子長) を小さくすることによる改善効果も確認できる.

## 5.4 考察

本節では, VMM の規模と透過率が電力性能向上に重要となる根拠を考察する. 図 6 は, No.12 の設定における N に対する電力効率のグラフを示す. ただし, 凡例にて基本ユニットの透過率を変化させている. ピークの電力効率を達成するためには, 適切な VMM 規模 (N) を選択する必要があることが分かる. MZI-VMM においては, N に応じてスループットは  $N^2$  増加し消費電力は N で増加するため, 基本的には N を増加させると高い電力効率を達成できる. 一方で, 提案している性能モデルではある精度の演算誤差を

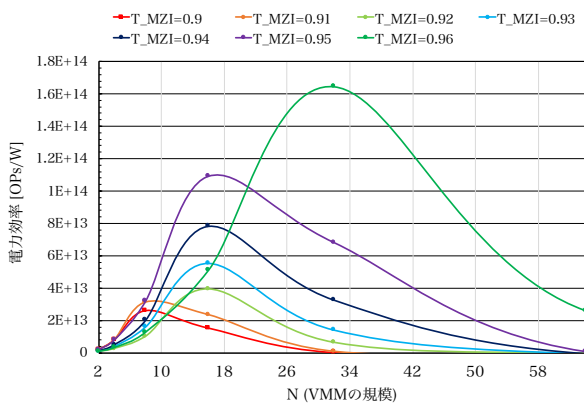


図 6: 各透過率における最適な VMM 規模の関係

保障するよう考慮しているため、PD に到達した際の光強度を一定以上に保つ必要がある。LS から照射された光波は  $N$  に対して指数的に減衰するため、LS 強度も指数的に増加させる必要があり、これが高消費電力の要因となる。この 2 つの要因により  $N$  に対してピークの電力効率が現れる。つまり、透過率を改善することで電力効率のピークをより  $N$  の大きい領域へと移行することができる。透過率  $T_{MZI}$  が如何に改善されても  $N$  が改善できない状況 (No.13) においては、ピークの  $N$  を設定することができないために高い電力効率を達成できない。また、 $N$  が如何に改善されても透過率  $T_{MZI}$  が改善できない状況 (No.12) においては、ピークが  $N$  の小さい値となるので高い電力効率を達成できない。以上により、VMM の規模  $N$  と基本ユニットの透過率  $T_{MZI}$  を同時に改善することが最重要課題となる。

最新の CMOS ニューラルネットアクセラレータである TPU [10], ISAAC [14], DaDianNao [6] の電力効率とを比較する。表 4 の No.1 の構成と比較した結果、TPU の電力効率比は 0.07 倍、ISAAC は 0.02 倍、DaDianNao は 0.02 倍であった。何れのアクセラレータも演算器のみならずメモリ、I/O 等の周辺機器を含んだ消費電力で電力効率を算出しており、また ISAAC, DaDianNao では 16bit 積和演算をベースとしているため、公平な性能比較とはならないことに注意されたい。その一方で、光演算器は現行のデバイスパラメータでも 10 倍以上の電力性能を達成できる。すなわち、今後のナノフォトニクス技術の進歩と適切なアーキテクチャの設計によって、ナノフォトニックコンピューティングが CMOS を凌駕する性能を達成できると考えられる。

## 6. おわりに

本稿では、ナノフォトニックコンピューティングの優位性ならびに将来展望を明らかにすべく、その電力効率の限界について解析を行った。光演算器としては、MZI-VMM を対象とし、その性能をモデル化した。アナログ演算器においては、ある一定精度以上の演算精度を担保することを考えるとノイズは極めて重要な性能律速要因となる。本モ

デルにおいては、最大ノイズが観測される条件を前提とし、性能限界についての予測を行った。その結果、ナノフォトニックコンピューティングは現行の光デバイス技術においても高い電力効率を達成可能であることが分かった。さらに、MZI-VMM の電力効率の改善に大きく影響を与える光デバイスパラメータを明らかにした。

今後、MZI-VMM の精度と詳細な消費電力を実機で計測し、本モデルの妥当性について検討する。また、メモリシステムや CPU インタフェースを含むシステムアーキテクチャの考案、プログラミングモデルの検討、ならびに、試作チップやシミュレータを用いた詳細評価を行い、ナノフォトニックコンピュータの有効性を実証する予定である。

**謝辞** 本研究の一部は、科学技術振興機構の戦略的創造研究推進事業「新たな光機能や光物性の発見・利活用を基軸とする次世代フォトニクス」の助成により行われた。

## 参考文献

- [1] Akiyama, S., Baba, T., Imai, M., Akagawa, T., Takahashi, M., Hirayama, N., Takahashi, H., Noguchi, Y., Okayama, H., Horikawa, T. et al.: 12.5-Gb/s operation with 0.29-V $\cdot$ cm V $\pi$  L using silicon Mach-Zehnder modulator based-on forward-biased pin diode, *Optics express*, Vol. 20, No. 3, pp. 2911–2923 (2012).
- [2] Batten, C., Joshi, A., Orcutt, J., Khilo, A., Moss, B., Holzwarth, C. W., Popovic, M. A., Li, H., Smith, H. I., Hoyt, J. L., Kartner, F. X., Ram, R. J., Stojanovic, V. and Asanovic, K.: Building Many-Core Processor-to-DRAM Networks with Monolithic CMOS Silicon Photonics, *IEEE Micro*, Vol. 29, No. 4, pp. 8–21 (2009).
- [3] Beamer, S., Sun, C., Kwon, Y.-J., Joshi, A., Batten, C., Stojanovic, V. and Asanovic, K.: Re-architecting DRAM Memory Systems with Monolithically Integrated Silicon Photonics, *Proceedings of the 37th Annual International Symposium on Computer Architecture*, ISCA '10, New York, NY, USA, ACM, pp. 129–140 (2010).
- [4] Carolan, J., Harrold, C., Sparrow, C., Martin-Lopez, E., Russell, N. J., Silverstone, J. W., Shadbolt, P. J., Matsuda, N., Oguma, M., Itoh, M. et al.: Universal linear optics, *Science*, Vol. 349, No. 6249, pp. 711–716 (2015).
- [5] Cartwright, S.: New optical matrix-vector multiplier, *Applied optics*, Vol. 23, No. 11, pp. 1683–1684 (1984).
- [6] Chen, Y., Luo, T., Liu, S., Zhang, S., He, L., Wang, J., Li, L., Chen, T., Xu, Z., Sun, N. et al.: Dadiannao: A machine-learning supercomputer, *Proceedings of the 47th Annual IEEE/ACM International Symposium on Microarchitecture*, IEEE Computer Society, pp. 609–622 (2014).
- [7] Clements, W. R., Humphreys, P. C., Metcalf, B. J., Kolthammer, W. S. and Walmsley, I. A.: An Optimal Design for Universal Multiport Interferometers, *arXiv preprint arXiv:1603.08788* (2016).
- [8] Esmailzadeh, H., Sampson, A., Ceze, L. and Burger, D.: Neural Acceleration for General-Purpose Approximate Programs, *Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO-45, Washington, DC, USA, IEEE Computer Society, pp. 449–460 (2012).
- [9] Haron, N. Z. and Hamdioui, S.: Why is CMOS scaling com-



- ing to an END?, *Design and Test Workshop, 2008. IDT 2008. 3rd International*, IEEE, pp. 98–103 (2008).
- [10] Jouppi, N. P., Young, C., Patil, N., Patterson, D., Agrawal, G., Bajwa, R., Bates, S., Bhatia, S., Boden, N., Borchers, A. et al.: In-datacenter performance analysis of a tensor processing unit, *arXiv preprint arXiv:1704.04760* (2017).
- [11] Notomi, M., Nozaki, K., Shinya, A., Matsuo, S. and Kuramochi, E.: Toward fJ/bit optical communication in a chip, *Optics Communications*, Vol. 314, pp. 3–17 (2014).
- [12] Optilab, LLC: 40 GHz Linear InGaAs PIN Photodetector PD-40, [http://www.optilab.com/devices/category/photo\\_diode/40\\_ghz\\_linear\\_ingaas\\_pin\\_photodetector/j](http://www.optilab.com/devices/category/photo_diode/40_ghz_linear_ingaas_pin_photodetector/j).
- [13] Reck, M., Zeilinger, A., Bernstein, H. J. and Bertani, P.: Experimental realization of any discrete unitary operator, *Phys. Rev. Lett.*, Vol. 73, pp. 58–61 (1994).
- [14] Shafiee, A., Nag, A., Muralimanohar, N., Balasubramonian, R., Strachan, J. P., Hu, M., Williams, R. S. and Srikumar, V.: ISAAC: A Convolutional Neural Network Accelerator with In-Situ Analog Arithmetic in Crossbars, *Proc. ISCA* (2016).
- [15] Shen, Y., Harris, N. C., Skirlo, S., Prabhu, M., Baehr-Jones, T., Hochberg, M., Sun, X., Zhao, S., Larochelle, H., Englund, D. et al.: Deep learning with coherent nanophotonic circuits, *Nature Photonics* (2017).
- [16] Tait, A. N., Zhou, E., de Lima, T. F., Wu, A. X., Nahmias, M. A., Shastri, B. J. and Prucnal, P. R.: Neuromorphic Silicon Photonics, *arXiv preprint arXiv:1611.02272* (2016).
- [17] Takahashi, Y., Inui, Y., Chihara, M., Asano, T., Terawaki, R. and Noda, S.: A micrometre-scale Raman silicon laser with a microwatt threshold, *Nature*, Vol. 498, No. 7455, pp. 470–474 (2013).
- [18] Taubenblatt, M.: Optical interconnects for high performance computing, *Optical Fiber Communication Conference*, Optical Society of America, p. OThH3 (2011).
- [19] Thorlabs, Inc.: High-Speed Photodetectors, [https://www.thorlabs.co.jp/NewGroupPage9.cfm?ObjectGroup\\_ID=1295](https://www.thorlabs.co.jp/NewGroupPage9.cfm?ObjectGroup_ID=1295).
- [20] Vantrease, D., Schreiber, R., Monchiero, M., McLaren, M., Jouppi, N. P., Fiorentino, M., Davis, A., Binkert, N., Beausoleil, R. G. and Ahn, J. H.: Corona: System implications of emerging nanophotonic technology, *ACM SIGARCH Computer Architecture News*, Vol. 36, No. 3, IEEE Computer Society, pp. 153–164 (2008).
- [21] 川上哲志, 磯部聖, 浅井里奈, 小野貴継, 本田宏明, 井上弘土, 納富雅也: ナノフォトニック・ニューラルアクセラレーション構想, 情報処理学会研究報告 (デザインガイア), Vol. 2016-ARC (2016).