

集積ナノフォトニクスに基づく近似並列乗算器を用いた 低レイテンシ光ニューラルネットワーク

塩見 準^{1,a)} 石原 亨² 小野寺 秀俊¹ 新家 昭彦^{3,4} 納富 雅也^{3,4}

概要: 集積ナノフォトニクスの実現により、超高速動作する光集積回路の設計が可能になった。本稿では、光スイッチング素子を活用したニューラルネットワークを述べる。まず、本稿では近似並列乗算と累算を低いレイテンシで行う光近似積和乗算器を提案する。超低レイテンシ動作の鍵は、回路のクリティカルパス上に存在する光電変換器 (OptoElectric converter, 以後 OE 変換器と呼ぶ) を最小化することである。ニューラルネットワークでは正確な演算は必ずしも必要とされず、演算回路に近似を導入することが可能である。提案演算器は近似誤差を持つ一方、任意のビット幅に対しクリティカルパス上に存在する OE 変換器が高々 1 個にとどまり、従来の CMOS 演算器と比較して 1 桁以上高速になることを示す。最後に提案乗算器を用いたニューラルネットワークの構成例を述べる。

1. 序論

今日の情報通信技術の飛躍的な発展は、CMOS 集積回路の進化なしでは語ることができない。CMOS 集積回路の微細化により情報通信機器の高性能化・低電力化が実現されてきた。しかし、集積回路微細化による性能改善は限界に達しつつあることが懸念されている。例えば文献 [1] では、微細化により配線抵抗・容量が増大し、その結果 CMOS 論理ゲートの伝搬遅延は 10 ps 程度に飽和することが指摘されている。

他方、大容量情報通信の需要増大を背景に、光ファイバ通信を代表とする光通信技術が急速な発展を遂げている。データセンタ間などの長距離情報通信だけでなく、データセンタ内通信やサーバラック間通信など、近距離通信に対しても光通信技術が積極的に用いられるようになってきている。近年、集積ナノフォトニクス技術の進展により、さらに近距離なオンチップ光通信が可能になった。CMOS 集積回路と異なり、光集積回路は配線の抵抗や容量に依存せず情報通信を行える。例えば、10 μm 程度の素子長で、100 fs の伝搬遅延特性を持つ光スイッチング素子の開発が現在進められており、情報通信機器の性能改善を後押しする技術として注目されている [2]。近年では、ナノフォトニクス技術を用いてチップ内通信速度を加速させる研究などが

活発に行われている [3]。

文献 [3] などの光集積回路の多くの研究では、通信に焦点を当てて研究が行われている一方、本研究は光スイッチング素子を用いた演算回路設計技術に焦点を当てる。特に、光ネットワーク上で、光信号のまま効率的に通信パケットの分配や経路最適化を行うことを目的として、ニューラルネットワークの実装に適した演算器設計技術に注目する。文献 [4-7] などの研究では、光スイッチング素子を用いた論理回路設計手法が提案されている。しかし、ニューラルネットワークの性能律速要因である並列乗算器や累算器のような複雑な演算器の設計手法は十分に研究されていない。本稿ではまず、光ニューラルネットワークに適した近似並列乗算器と累算器の提案を行う。当該演算器は、電気のデジタル信号を入力にアナログの光信号を光速で生成する。次に、提案演算器を用いたニューラルネットワークの設計事例を述べる。波長分割多重技術を活用し、効率的に並列乗算器と累算器の使用個数を削減する技術を提案する。

本稿の構成を以下に示す。まず第 2 章で関連研究と本研究の成果を述べる。第 3 章で提案演算器の構成を述べて、CMOS を用いた従来回路との比較を行う。第 4 章でニューラルネットワークの設計事例を述べる。第 5 章で結論を述べる。

2. 関連研究

光集積回路中で用いられるスイッチング素子を図 1 に示す。図 1 (a), (b) はそれぞれ 1×1 , 2×2 光スイッチング素

¹ 京都大学大学院情報学研究科
² 名古屋大学大学院情報学研究科
³ NTT ナノフォトニクスセンタ
⁴ NTT 物性科学基礎研究所
a) shiomi-jun@i.kyoto-u.ac.jp

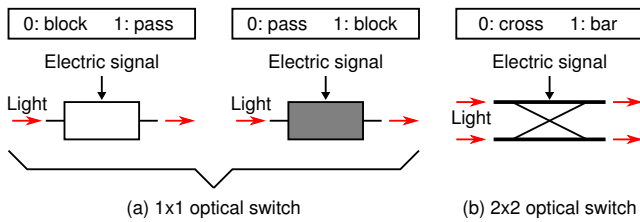


図 1 光の伝搬経路を切り替えるスイッチング素子.

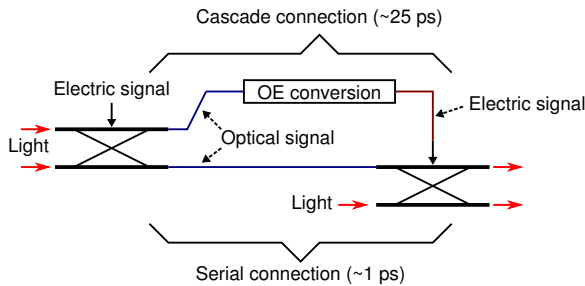


図 2 光スイッチング素子の接続方法.

子である。両方式ともに、入力デジタル電気信号に応じて光の伝搬経路を切り替える。1×1 光スイッチング素子の場合、入力電気信号に基づき光信号を伝搬させるか止めることができ、2×2 の場合 2 つの光信号を直進させるか交差させるか選択することができる。光集積回路の低レイテンシ動作を享受するためには光スイッチング素子の接続方法を理解することが重要である。図 2 に、光スイッチング素子の接続方法を示す。シリアル接続の場合、光信号が直進することで論理演算できるため低レイテンシな演算を実現できる。他方、カスケード接続を行う場合、光電変換 (OptoElectric 変換: OE 変換) が必要であり、大きな動作遅延が必要になる。本稿で使用するナノフォトニックデバイスの場合、OE 変換におよそ 25 ps 必要である一方、シリアル接続で必要な伝搬遅延は光スイッチング素子 1 個あたり 1 ps である。

光スイッチング素子のシリアル接続を最大限活用することで、低レイテンシ動作する光論理回路を構成できる。光論理回路設計に関する初期の研究は Hardy らによる [4]。Hardy らは 2×2 光スイッチング素子を用いた Directed Logic (DL) を提案している。パストランジスタロジックのように、DL では光スイッチング素子をシリアル接続し、任意の論理関数をシリアル接続のみで実現できる。その他、任意の論理関数を光スイッチング素子で実現する方式として、二分決定グラフを実装する方法 [5] などが提案されている。しかし、いずれの方式においても、並列乗算器のような複雑な回路を実装すると回路規模が爆発的に増大する。また、光スイッチング素子を用いた低レイテンシ並列加算器の設計手法が活発に研究されているが [8,9]、並列乗算器に対する実装方式は十分に研究されていない。本稿では特に、ニューラルネットワークに適した並列乗算器の実装方式に関して議論する。

光スイッチング素子を用いた低レイテンシニューラルネットワーク構成手法は [10,11] など議論されているが、例えば [10] はニューロン数の 2 乗のオーダーで光スイッチングデバイスが必要で、スケーラビリティに乏しい。また、[11] ではアナログ演算が用いられているため、事前の学習が困難である。本稿では半デジタル的にニューラルネットワークを構成する手法を提案し、事前学習可能なニューラルネットワークの実現を目指しつつ、波長分割多重技術を用いてニューロン数の 1 乗のオーダーの光スイッチング素子で実装可能なニューラルネットワーク実現例を示す。

本稿の成果を以下に示す。

- 低レイテンシに動作する近似積和演算器を提案する。提案演算器は計算誤差を有するものの、クリティカルパス上に OE 変換器が高々 1 個しか存在せず、光速度で積和演算を行える。CMOS 演算回路との比較の結果、1 桁以上高速に動作することを示す。
- 全結合層で構成されたニューラルネットワークを実現するために必要な並列乗算器の個数は、ニューロン数の 2 乗のオーダーに比例する。本稿では、波長多重技術を活用し、1 個の並列乗算器で複数の並列乗算を実現できる回路構成を示す。その結果、並列乗算器の個数がニューロン数の 1 乗算になることを示す。

3. 低レイテンシ近似並列乗算器・累算器

3.1 低レイテンシ・省素子演算器の必要性

最も基本的なニューラルワークの構造を図 3 に示す。当該ニューラルネットワークはすべて全結合層で構成されている。現在の層から次の層へ演算結果を伝搬させるためには、乗算および加算を複数回実行する必要がある。例えばレイヤ L , $L+1$ のニューロン数がそれぞれ N_L , N_{L+1} の場合、ニューラルネットワーク構成するために、 $N_L N_{L+1}$ 回の乗算と、各乗算結果を足し合わせる必要がある。すなわち以下に示すベクトル行列積を実行する必要がある。

$$\begin{aligned} X_1^{L+1} &= W_{1,1}^L X_1 + \dots + W_{3,1}^L X_{N_L} \\ X_2^{L+1} &= W_{1,2}^L X_1 + \dots + W_{3,2}^L X_{N_L} \\ &\vdots \\ X_{N_{L+1}}^{L+1} &= W_{1,N_{L+1}}^L X_1 + \dots + W_{3,N_{L+1}}^L X_{N_L} \end{aligned}$$

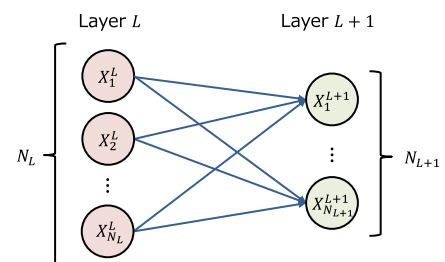


図 3 ニューラルネットワークの基本構造の一例.

したがって、大規模なニューラルネットワークを設計する場合、(i) 低レイテンシ動作し、かつ (ii) 素子数が少ない、積和演算器を設計することが重要である。本稿で述べる積和演算器では、任意のビット幅に対してクリティカルパス中に OE 変換器が 1 個のみ存在し、低レイテンシ動作を実現する。また、4 章にて波長分割多重技術を用いることで、1 個の並列乗算器で複数の乗算を同時に行えることを示し、回路面積の効率的な削減を実現する。

3.2 コンセプト

2 つの固定小数点数 (以後 W, X とする) の積を考える。 W の整数部, 小数部のビット幅をそれぞれ n_W, m_W とし, X の整数部, 小数部のビット幅をそれぞれ n_X, m_X とする。一般的に, 配列乗算器や Wallace 木乗算器等の並列乗算器を実現するためには, 部分積を足し合わせるために全加算器 (Full Adder: FA) を実装する必要がある。文献 [9] が指摘するように, FA 1 個につき OE 変換器が 1 個必要であり, クリティカルパス上にあらわれる OE 変換器の数はビット幅の増大とともに増大する。したがって, 単純に既存の並列乗算器を光スイッチング素子で実装すると低レイテンシ動作を損なう問題が存在する。本稿ではクリティカルパス上の OE 変換器の個数を削減するため, 以下に示す近似 (対数量子化) に基づき並列乗算器の設計に取り組む:

$$W \times X \simeq W \times 2^{\hat{x}} \quad (1)$$

ここで \hat{x} は $\log_2(X)$ を整数に打ち切った値で, これを対数量子化と呼ぶ。対数量子化によりビット幅は $\lceil \log_2(n_X) \rceil$ まで削減される。ここで, $\lceil \cdot \rceil$ は天井関数である。乗数 X が 2 のべき乗値に丸め込まれるため, 近似乗算による演算誤差は最大 50% になる。他方, 乗算をビットシフトに置き換えることができ回路規模を大幅に削減できる。このコンセプトに基づき, ニューラルネットワークに対数量子化を取り込んだ研究が [12,13] で行われている。文献 [12,13] では, 32 ビット浮動小数点を用いてニューラルネットワークを実装した場合と比較して 5 ビットの対数量子化を用いた場合でも推論精度の劣化は数ポイントであることが述べられている。本稿では, (1) で示した対数量子化のアプローチを光スイッチング素子を用いて実装する。

3.3 近似並列乗算器と累算器の概要

近似並列乗算器と累算器を組み合わせた近似積和演算器の概要を図 4 に示す。対数量子化された光デジタル信号 \hat{x}_i が OE 変換され, デジタル電気信号に変換される。並列乗算器には重み係数 W_i に比例した電界強度 (振幅) を持つアナログ光が入力される。デジタル信号 \hat{x}_i の値に基づきアナログ光 W_i の電界強度, 伝搬経路が切り替わり, 光速で乗算を実現する。コンバイナベースの累算器 [14] を経て乗算結果が加算され, 積和演算が実現される。コン

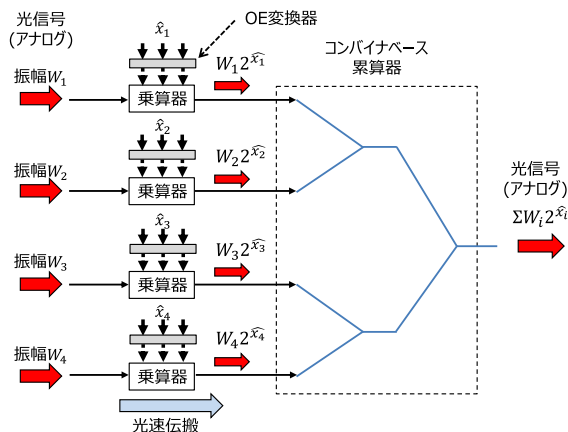


図 4 近似積和演算器の構成。

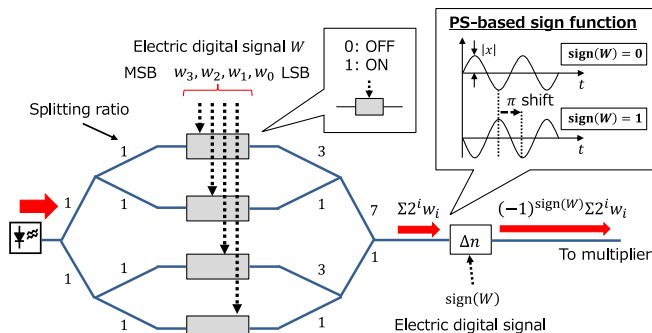


図 5 光 DA コンバータ。

バイナ 1 段あたりの伝搬遅延は 100 fs 以下である。

図 4 左部分に描いたアナログ光 W_i を生成するため, 図 5 に示す DA コンバータを用いる。光源をスプリッタで分岐し, 1×1 光スイッチング素子 [15] に入力されるデジタル電気信号 w_i により信号の通過/遮断が制御される。コンバイナベースの累算器 [14] の分岐比を調整し, W_i に比例する電界強度を持つ光信号を生成する。 W の符号ビットに応じて位相シフタ (Phase Shifter: PS) のシフト量が変化する。符号ビットにより位相が π ずれ, 結果としてマイナス演算を実現する。位相が π ずれた光信号を足し合わせると互いに信号を弱め合うため, 図 4 の累算器部分で減算を実現する。次節では近似並列乗算器に関して述べる。

3.4 近似並列乗算器

近似光並列乗算器の構成例を図 6 に示す。 2×2 スwitchング素子がシリアル接続されており, 定常光が接続されて

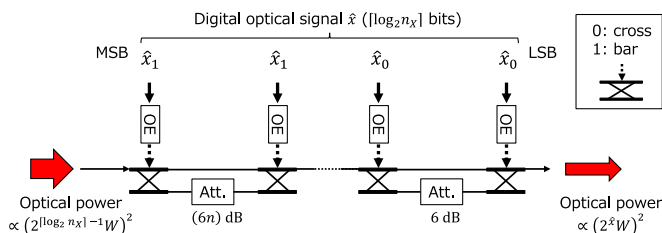


図 6 近似並列乗算器の構成。

いる。対数量子化されたデジタル信号 (\hat{x}) のビット幅は $\lceil \log_2(n_X) \rceil$ である。2×2 スイッチング素子の電気信号端子に入力され、 \hat{x} の値に応じて定常光の伝搬経路が変化する。例えば \hat{x} の各ビットがすべて 1 の場合、定常光は減衰せずに光速で伝搬する。 \hat{x} のビットが 0 の場合、光信号は減衰器を伝搬し、光信号の電界強度が減衰する。例えば光信号のパワーが 6 dB 減衰する場合、光信号の電界強度が半分に減衰する。具体的には、図 6 において、アナログ入力光のパワーを $(2^{\lceil \log_2 n_X \rceil - 1} W)^2$ に比例する値とすると、対数量子化された \hat{x} の値に応じて出力光電力が $(2^{\hat{x}} W)^2$ まで減衰するため、出力光の電界強度を観測することで乗算を行える。また、シリアル接続のみの構成のため、クリティカルパス上に存在する OE 変換器は任意のビット幅に対して 1 個である。

3.5 性能評価

本章で述べた近似並列乗算器と CMOS 回路の性能を数値計算上で比較する。本稿では、光スイッチング素子としてナノフォトニックデバイス [2, 15] を使用することを想定する。光信号がスイッチングデバイスを伝搬する遅延 τ_{sw} を 1 ps とする。また、[15] に基づき、フォトディテクタベースの OE 変換器の変換遅延 τ_{oe} を 25 ps とする。並列乗算器のクリティカルパスは、図 6 の対数量子化された入力部分の MSB 側から OE 変換を通して出力部分へ至るパスである。図 5 に示すデジタル入力部分のスイッチング時間を τ_{oe} とすると、クリティカルパス遅延 D_{opt} を以下でモデル化できる。

$$D_{opt} = \tau_{oe} + 2\tau_{sw} \log_2 n_X. \quad (2)$$

他方、提案内容と同等の機能を有する近似並列乗算器を CMOS 回路で構成した場合の遅延を検討する。最も簡単な実装方式はバレルシフタを用いる方式である。バレルシフタのクリティカルパスは、 $\log_2 n_X$ 段の 2 入力マルチプレクサ (MUX2) より構成される。文献 [1] に基づき論理ゲート 1 個分の伝搬遅延 τ_{gate} を 10 ps と仮定する。伝搬遅延 D_{ele} は以下のとおりとなる。

$$D_{ele} = \tau_{gate} \log_2 n_X. \quad (3)$$

D_{opt} , D_{ele} の比較結果を図 7 に示す。なお、簡単のため $n_X = m_X = n_W = m_W$ としている。“Optical”が提案乗算回路のクリティカルパス遅延、“Electrical”が従来の CMOS ベースのバレルシフタのクリティカルパス遅延である。乗数 X のビット幅 n_X 光集積回路の増加にともない、“Electrical”のクリティカルパス遅延は増加する一方、“Optical”のクリティカルパス遅延はほぼ一定である。これは、任意のビット幅 n_X に対して任意のパス上に OE 変換器が 1 個以下のみ存在するためである。この結果、 n_X が 16, 32 の時、“Optical”が “Electrical” よりそれぞれ 17%,

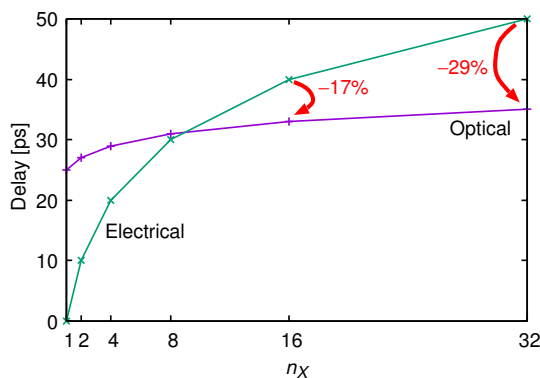


図 7 近似乗算器のクリティカルパス遅延とビット幅の依存関係 ($n_X = m_X = n_W = m_W$ のとき)。

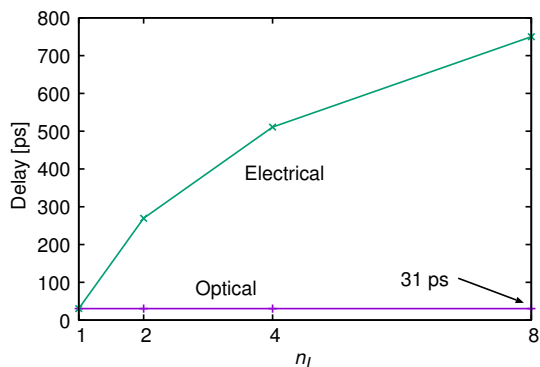


図 8 近似積和演算器のクリティカルパス遅延とニューロン数の依存関係 ($n_X = m_X = n_W = m_W = 8$ のとき)。

29% 程度高速に動作する。

“Optical”の高速性は、累算器を考慮することでさらに加速される。“Electrical”の累算器として、複数のリップルキャリアアダーを 2 分木構造で接続した構成を考える。FA 1 個あたりの伝搬遅延を τ_{gate} とし、各リップルキャリアアダーのクリティカルパス遅延を $(n_W + m_W + n_X) \tau_{gate}$ でモデル化する。クリティカルパスは $\log_2 N_L$ 個のリップルキャリアアダーを通るため、近似積和演算器全体のクリティカルパス遅延は (3) に示した D_{ele} に $(n_W + m_W + n_X) \tau_{gate} \log_2 N_L$ を足し合わせた値である。ただし、 N_L はレイヤ L のニューロン数である。同様に、“Optical”で構成する近似積和演算器全体のクリティカルパス遅延を考える。[14] に基づきコンバイナの伝搬遅延 τ_{com} を 100 fs とする。図 4 に示すように、コンバイナを用いた 2 分木ベースの累算器の伝搬遅延は $\tau_{com} \log_2 N_L$ である。したがって、 D_{opt} に $\tau_{com} \log_2 N_L$ を足した値が近似積和演算器のクリティカルパス遅延となる。両方式のクリティカルパス遅延を図 8 に示す。簡単のため、 $n_X = m_X = n_W = m_W = 8$ としている。横軸はレイヤ L のニューロン数 N_L である。“Optical”は光速で累算を行うため、足し合わせる乗算結果数が増大してもほぼ一定の伝搬遅延を維持し、 $N_L = 8$ の際およそ 31 ps となる。他方、“Electrical”は N_L の増大と共にクリティカルパス上の並列加算器数が増大するため、“Optical”と比べて

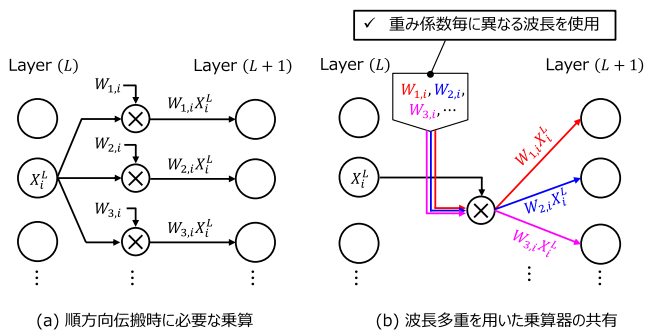


図 9 波長分割多重技術を用いた乗算器の共有.

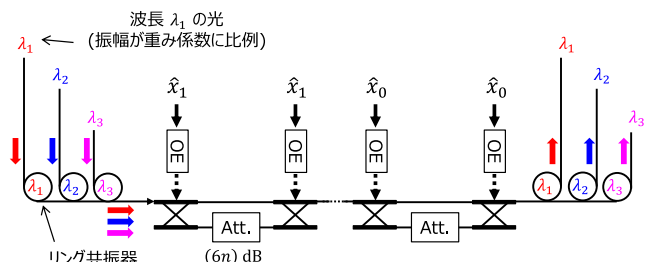


図 10 波長分割多重技術を適用した近似乗算器.

急速に遅延が悪化する。その結果、 $N_L \geq 4$ 以上の規模の回路で、“Optical”は“Electrical”より1桁以上高速に動作する。

4. 光ニューラルネットワークへの適用

4.1 波長分割多重を用いた回路素子数削減

波長分割多重 (Wavelength Division Multiplex: WDM) を用いた回路素子数削減手法のコンセプトを図9に示す。図9(a)は、レイヤ L のニューロンから得られる値 X_i^L がレイヤ $L+1$ の各ニューロンへ伝搬する状況を示している。重み係数 W_j^L が分配先のニューロンに応じて変化するため、各分配先で乗算結果が異なる一方、被乗数 X_i^L は常に一定である。この特徴を利用し、図6(b)に示すような回路素子削減手法を提案する。被乗数 $2^{\hat{x}}$ を1個の並列乗算器に入力し、添え字 i に応じて異なる波長を持つ光信号を入力する。各信号強度を重み係数に応じて予め変調する。波長の異なる光信号は互いに干渉し合わないため、独立して並列乗算を行うことができる。光回路を用いた実装例を図10に示す。図6に示した回路にマイクロリング共振器をが追加されている。マイクロリング共振器は、共振器の半径に応じて、特定の波長のみ進行方向を曲げる光素子である。マイクロリング共振器を用いて光信号を並列乗算器にまとめることで、回路素子数を効率的に削減できる。3.2節で述べた通り、レイヤ $L, L+1$ のニューロン数をそれぞれ N_L, N_{L+1} とすると、単純に並列乗算器を実装した場合 $N_L \times N_{L+1}$ 個並列乗算器が必要である。他方、波長多重技術を活用することで、回路実装に必要な光並列乗算器数は N_L まで削減される。

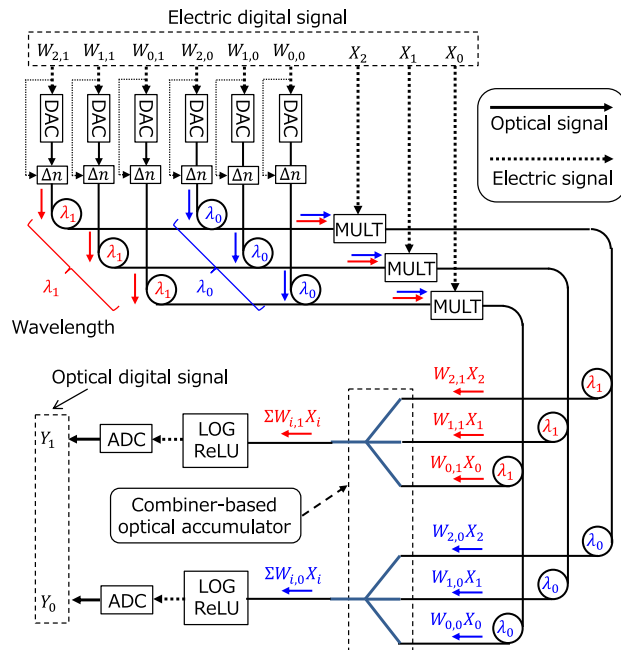


図 11 光ニューラルネットワークの構成例.

4.2 全体像

対数量子化と波長多重技術を適用したニューラルネットワークの構成例を図11に示す。図11は、レイヤ $L, L+1$ のニューロン数がそれぞれ3, 2の場合であり、レイヤ L とレイヤ $L+1$ の間で行われる積和演算回路を示している。第3章で述べた通り、“DAC”、“MULT”、“Combiner-based optical accumulator”でそれぞれDA変換、近似並列乗算、累算を行っている。また、重み係数 W_j^L に与える光信号の波長をそれぞれ λ_1, λ_0 に振り分けることで、近似並列乗算器を共有し、光素子数を削減することができる。

積和演算の結果、振幅変調されたアナログ光が得られる。この振幅値を (i) Rectified Linear Unit (ReLU) を用いた活性化関数へ入力し、(ii) 対数演算を行い、(iii) AD変換して次のレイヤへ値を与える、ことでニューラルネットワークを構成する。累算器を通過した光信号の振幅値から演算結果を求めることができるため、ホモダイン検波を通して光の振幅値を電流に変換する。ホモダイン検波回路の構成を図12に示す。“アナログ光信号”部分より、累算器を通過した光信号が入力される。入力された光信号の電界強度 A に比例した電流がフォトディテクタ (PD) を通って出力される。入力された光信号が参照光と位相 π ずれる場合、PDが出力する電流の向きが逆向きになり、マイナスの値

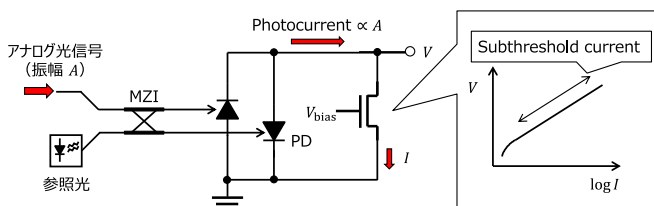


図 12 対数演算を同時に行うホモダイン検波回路.

を取得することができる。生成された光電流がトランジスタを通ることで、 A に応じた電圧値が V より出力される。ここで、トランジスタのゲート電圧にしきい値電圧より小さな電圧 V_{bias} を与えることで、対数演算を行うことができる。これは、トランジスタに印加されるゲート電圧が極端に低い場合、ドレイン-ソース間の電流値は、ドレイン-ソース間の電圧値の指数関数に比例するためである。得られた電圧値を例えば [16] で提案されている光 AD コンバータに入力することで (i) と (iii) を実現できる。上記構成の動作検証は今後の課題である。

5. 結論

集積ナノフォトニクス技術の登場により、低レイテンシ動作するオンチップ回路の設計が可能になった。本稿では、ニューラルネットワークに適した近似並列乗算器と累算器の提案を行った。近似演算を導入することで、クリティカルパス上の OE 変換器の個数を高々 1 個に抑えつつ、従来の CMOS 型演算器より 1 桁以上高速な演算を実現可能であることを示した。また、波長多重技術を用いることで、ニューロン数の 2 乗に比例した個数が必要であった並列乗算器を、ニューロン数の 1 乗のレベルまで削減できることを示した。本稿の今後の課題として、電力や面積の見積もり、シミュレーションを通じた動作検証等が挙げられる。

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現・利活用を基軸とする次世代フォトニクスの基盤技術」(JPMJCR15N4) の助成により行われた。

参考文献

- [1] A. Ceyhan, M. Jung, S. Panth, S. K. Lim, and A. Naeemi, "Impact of Size Effects in Local Interconnects for Future Technology Nodes: A Study Based on Full-Chip Layouts," in *IEEE International Interconnect Technology Conference*, May 2014, pp. 345–348.
- [2] K. Nozaki, A. Shakoov, S. Matsuo, T. Fujii, K. Takeda, A. Shinya, E. Kuramochi, and M. Notomi, "Ultralow-energy electro-absorption modulator consisting of InGaAsP-embedded photonic-crystal waveguide," *APL Photonics*, vol. 2, no. 5, p. 056105, 2017.
- [3] X. Wu, Y. Ye, W. Zhang, W. Liu, M. Nikdast, X. Wang, and J. Xu, "UNION: A Unified Inter/Intra-Chip Optical Network for Chip Multiprocessors," in *IEEE/ACM International Symposium on Nanoscale Architectures*, June 2010, pp. 35–40.
- [4] J. Hardy and J. Shamir, "Optics Inspired Logic Architecture," *Opt. Express*, vol. 15, no. 1, pp. 150–165, Jan 2007.
- [5] T. Asai, Y. Amemiya, and M. Koshihara, "A Photonic-Crystal Logic Circuit Based on the Binary Decision Diagram," in *Int'l Workshop on Photonic and Electromagnetic Crystal Structures*, Mar 2000, pp. T4–14.
- [6] C. Condrat, P. Kalla, and S. Blair, "Logic Synthesis for Integrated Optics," in *Great Lakes Symposium on Great Lakes Symposium on VLSI*, ser. GLSVLSI '11, 2011, pp. 13–18.
- [7] Q. Xu and R. Soref, "Reconfigurable optical directed-logic circuits using microresonator-based optical switches," *Opt. Express*, vol. 19, no. 6, pp. 5244–5259, Mar 2011.
- [8] Z. Wang, Z. Ying, S. Dhar, Z. Zhao, D. Z. Pan, and R. T. Chen, "Optical switches based carry-ripple adder for future high-speed and low-power consumption optical computing," in *Conference on Lasers and Electro-Optics*. Optical Society of America, 2017, p. STh1N.2.
- [9] T. Ishihara, A. Shinya, K. Inoue, K. Nozaki, and M. Notomi, "An Integrated Nanophotonic Parallel Adder," *J. Emerg. Technol. Comput. Syst.*, vol. 14, no. 2, pp. 26:1–26:20, Jul. 2018.
- [10] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. B.-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, "Deep Learning with Coherent Nanophotonic Circuits," in *Nature Photonics*, vol. 11, Jun 2017, pp. 441–446.
- [11] A. N. Tait, T. F. de Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Neuromorphic Photonic Networks using Silicon Photonic Weight Banks," in *Scientific Reports volume 7, Article number: 7430*, Aug 2017.
- [12] D. Miyashita, E. H. Lee, and B. Murmann, "Convolutional neural networks using logarithmic data representation," *CoRR*, vol. abs/1603.01025, 2016. [Online]. Available: <http://arxiv.org/abs/1603.01025>
- [13] E. H. Lee, D. Miyashita, E. Chai, B. Murmann, and S. S. Wong, "LogNet: Energy-efficient neural networks using logarithmic computation," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 5900–5904.
- [14] S. Kita, K. Nozaki, K. Takata, A. Shinya, and M. Notomi, "Silicon Linear Optical Logic Gates for Low-Latency Computing," in *Conference on Lasers and Electro-Optics*, May 2018, pp. 1–2.
- [15] K. Nozaki, S. Matsuo, T. Fujii, K. Takeda, M. Ono, A. Shakoov, E. Kuramochi, and M. Notomi, "Photonic-crystal nano-photodetector with ultrasmall capacitance for on-chip light-to-voltage conversion without an amplifier," *Optica*, vol. 3, no. 5, pp. 483–492, May 2016.
- [16] Y. Imai, T. Ishihara, H. Onodera, A. Shinya, S. Kita, K. Nozaki, K. Takata, and M. Notomi, "An Optical Parallel Multiplier Using Nanophotonic Analog Adders and Optoelectronic Analog-to-Digital Converters," in *2018 Conference on Lasers and Electro-Optics (CLEO)*, May 2018, pp. 1–2.