# Evaluation of Vector-Matrix Multiplier using optical device

Koji Sato

Kyushu-University

# Agenda

- Background
- Purpose
- Introduction of MZI
- Evaluation
  - Area
  - latency
- Summary

# Background

- Needs for processing more data with low latency
  - ✓Cyber physical system
  - ✓Data center
  - ✓Embedded system

- CMOS LSI's technical and physical limitation
  - ✓Ohmic loss
  - ✓Leakage current
  - ✓Too high frequency

# Background

- Vector-matrix multiplier(VMM)
  - VMM is used in many system (e.g. neural network)

- Linear transformation with optical device
  - Low latency, and low power consumption
  - MZI VMM perform at almost light speed

# Purpose

- Evaluation of Vector-Matrix Multiplier(VMM) using optical devices

- Comparing optical VMM and other VMMs
  - Other VMM
    - ASIC
    - GPU
    - CPU

# Introduction of MZI VMM

- Singular Value Decomposition(SVD)

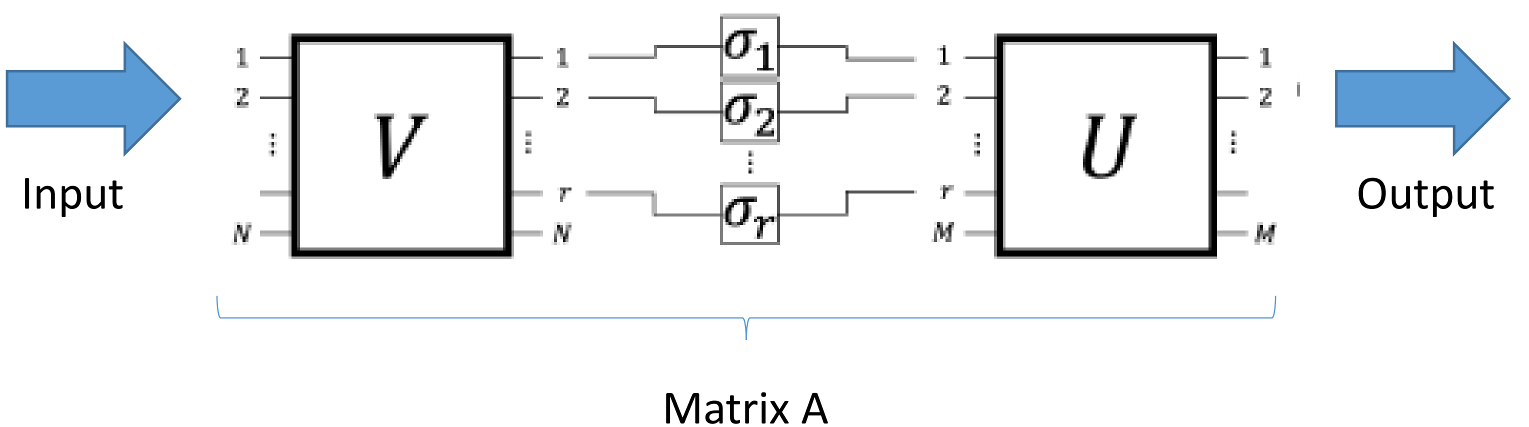
  M × N matrix (A) can be decomposed as :

$$A = U\Sigma V$$

  U : M × M unitary matrix

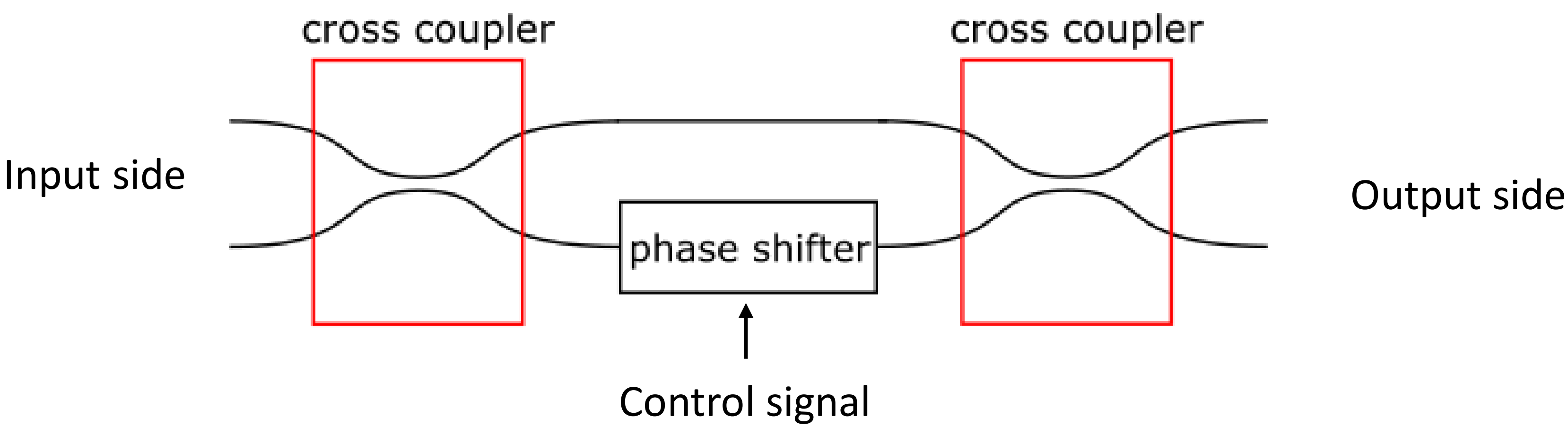
  V : N × N unitary matrix

  Σ : M × N diagonal matrix with non-negative real number



# Introduction of MZI VMM

- MZI(Mach-Zenhder Interferometer)

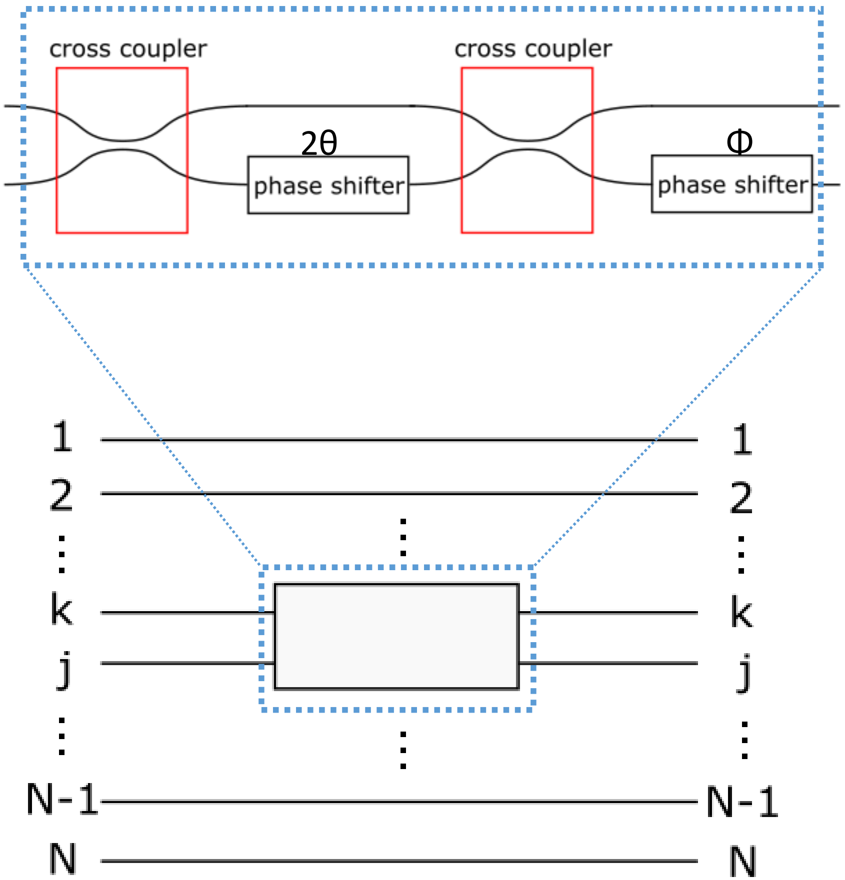



Cross coupler : coupling two of input lightwave

Phase shifter : shift lightwave’s phase depending on control signal

# Introduction of MZI VMM

- MZI unitary matrix transformation




This component corresponds to a following matrix transformation:

$$T(\theta, \phi) = \begin{pmatrix} e^{i\phi} sin\theta & e^{i\phi} cos\theta \\ cos\theta & -sin\theta \end{pmatrix}$$

$$T_{m,n}(\theta, \phi) = \begin{bmatrix} 1 & 0 & \cdots & & & 0 \\ 0 & 1 & & & & \vdots \\ \vdots & & e^{i\phi} sin\theta & e^{i\phi} cos\theta & & \vdots \\ \vdots & & cos\theta & -sin\theta & & \vdots \\ \vdots & & & & 1 & 0 \\ 0 & \cdots & & & 0 & 1 \end{bmatrix}$$

# Introduction of MZI VMM

# Introduction of MZI VMM

# Area : evaluation

- MZI VMM
  - Calculate from model formula

$$S = \mathrm{S_S} \times \mathrm{N} + \mathrm{S_{MZI}} \times N(N-1) + S_{AMP} \times N + S_{PD} \times N$$

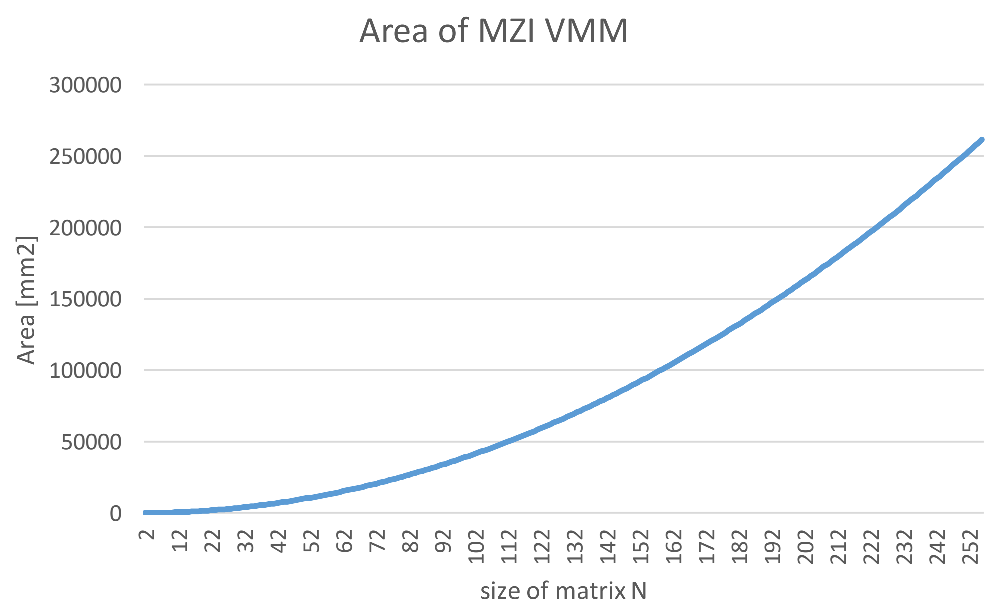
$S_S$ :size of a light source

$S_{MZI}$ : size of a MZI

$S_{AMP}$ : size of amplifier

$S_{PD}$ : size of photo detector

N; size of matrix



# Latency :method of Evaluation

- MZI VMM
  - Calculate from model formula

$$L = \frac{n}{c} l N_{pass} + L_{AMP} + L_{PD}$$

n : refractive index

c : speed of light

l : length of MZI

$N_{pass}$ : the max number of MZI that light must pass

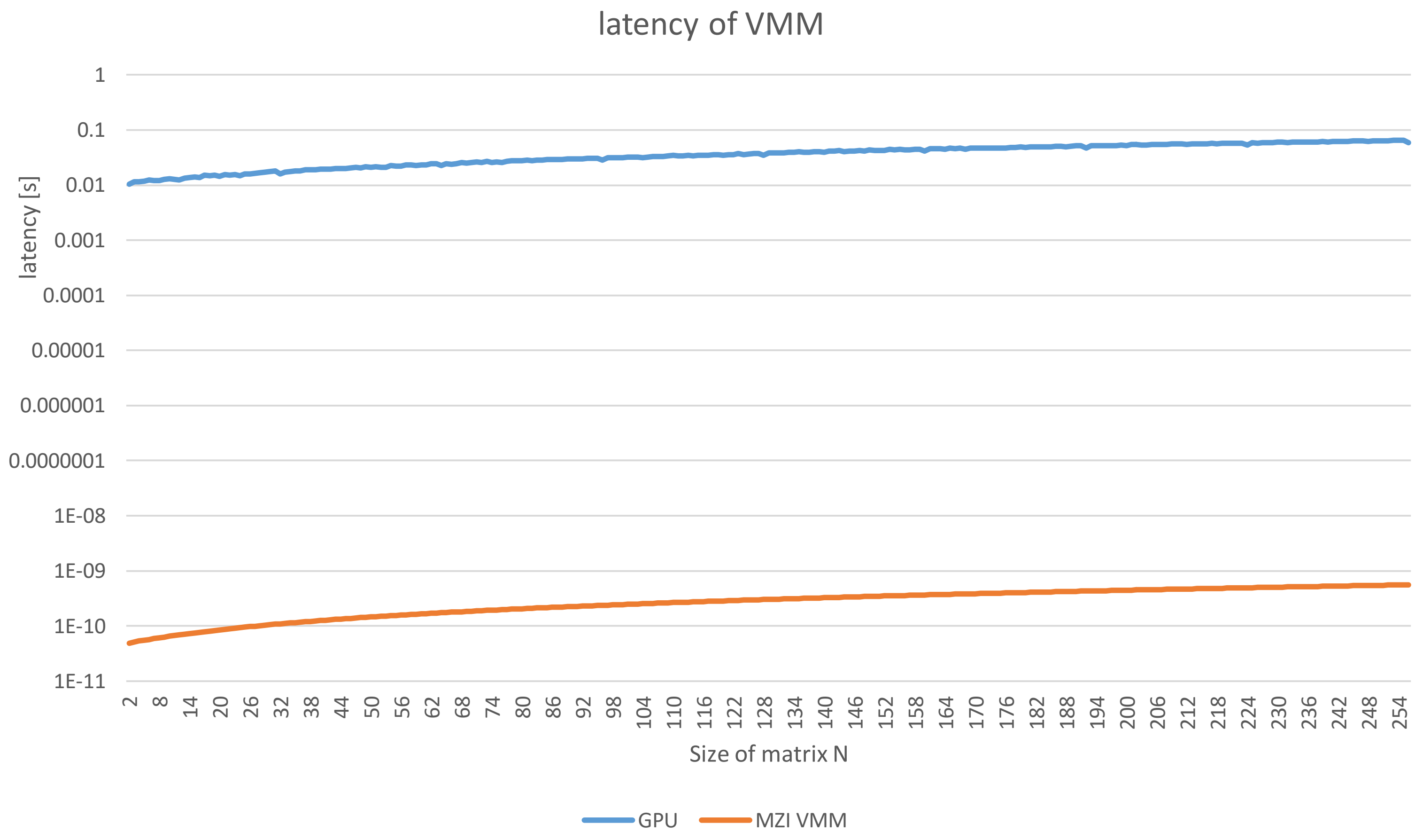
$L_{AMP}$ : latency of amplifier

$L_{PD}$ : latency of photo detector

# Latency :method of Evaluation

- GPU
  - Use library CUBLAS(CUDA Basic Linear Algebra subprograms)
  - Run on NVIDEA Tesla k20m (354nodes)
    - 345.6GFLOPS
    - Memory 128GB
    - Bandwidth 102.4GB
  - Calculate 400times and get the average time

# Evaluation


latency of VMM
latency [s]
1
0.1
0.01
0.001
0.0001
0.00001
0.000001
0.0000001
1E-08
1E-09
1E-10
1E-11
2 8 14 20 26 32 38 44 50 56 62 68 74 80 86 92 98 104 110 116 122 128 134 140 146 152 158 164 170 176 182 188 194 200 206 212 218 224 230 236 242 248 254
Size of matrix N
GPU
MZI VMM

# Plan

- Evaluate accuracy of MZI VMM
  - Survey about noise of optical devices
    - Optical amplifier
    - Phase shifter
    - Photo detector
  - Simulate on Optisystem

- Evaluate performance of vector-matrix multiplication in other method
  - ASIC
  - CPU

# Summary

- Introduce MZI VMM
- Evaluate vector-matrix multiplication
  - MZI VMM
  - GPU
- Plan to survey MZI VMM's noise