COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Analysis of Worst-case of Loss and Crosstalk for Planar Silicon Photonic Switch Architectures

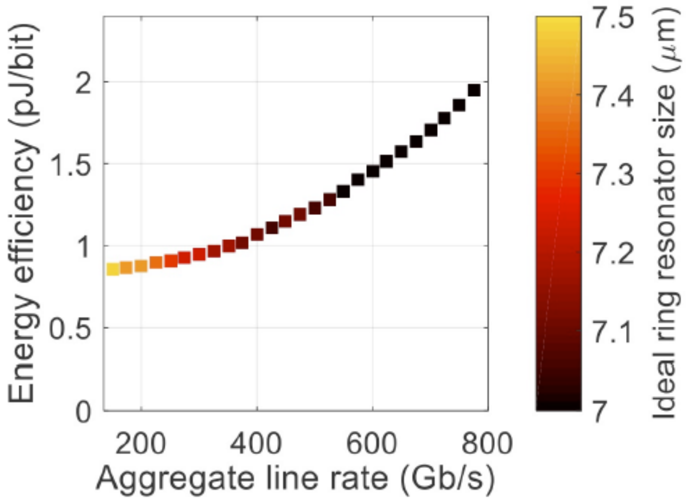
OPTICS 2017, March 27th-31st, Lausanne, Switzerland

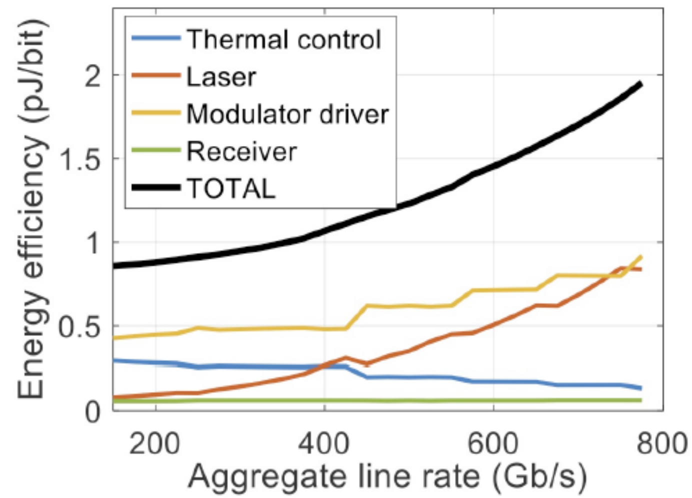
Meisam Bahadori, *Sébastien Rumley*

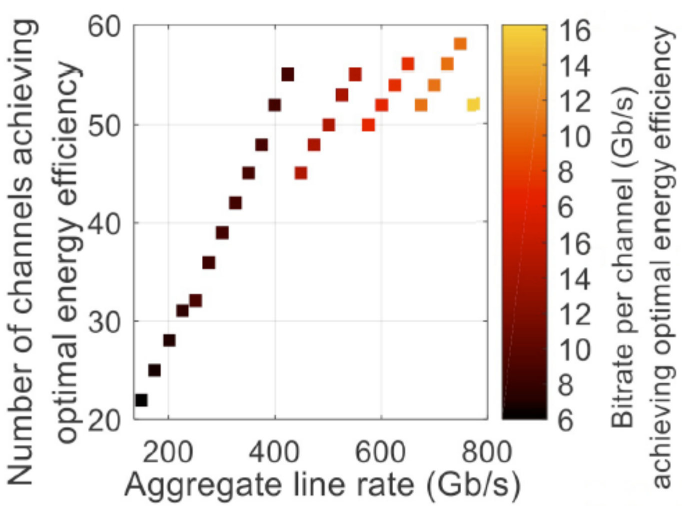
Lightwave Research Lab, Columbia University
New York, NY, USA

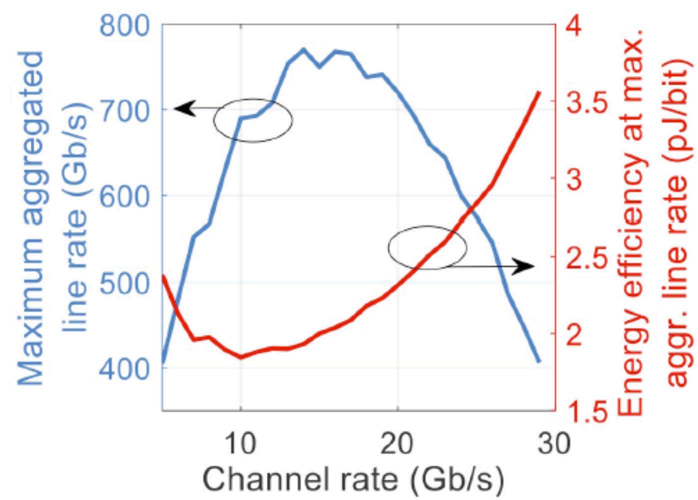
COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Fresh out the oven









[M. Bahadori et al. Energy-Performance Optimized Design of Silicon Photonic Interconnection Networks for High-Performance Computing, DATE, Lausanne, March 2017]

The bright side of optical switching

2.07Petabit/s

Optical Fiber Conference, 2001

1296-port MEMS Transparent Optical Crossconnect with 2.07Petabit/s Switch Capacity

R. Ryf, J. Kim, J. P. Hickey, A. Gnauck, D. Carr, F. Pardo, C. Bolle, R. Frahm , N. Basavanhally , C. Yoh, D. Ramsey, R. Boie , R. George, J. Kraus, C. Lichtenwalner, R. Papazian, J. Gates, H. R. Shea, A. Gasparyan, V. Muratov, J.E. Griffith, J.A. Prybyla, S. Goyal, C.D. White, M.T. Lin , R. Ruel, C. Nijander, S. Arney, D. T. Neilson, D. J. Bishop, P. Kolodner, S. Pau, C. Nuzman, A. Weis, B. Kumar, D. Lieuwen, V. Aksyuk, D. S. Greywall, T.C. Lee, H.T. Soh, W.M. Mansfield, S. Jin, W.Y. Lai, H.A. Huggins, D.L.Barr, R.A. Cirelli, G. R. Bogart, K. Teffeau, R. Vella, H. Mavoori, A. Ramirez, N.A.Ciampa, F.P. Klemens, M.D. Morris, T. Boone, J.Q. Liu, J.M.Rosamilia, C. R. Giles

Lucent Technologies
791 Holmdel-Keyport Rd
Holmdel, NJ 07733
Email: ryf@lucent.com

Agere Systems
600 Mountain Ave.
Murray Hill, NJ07974-0636

Abstract
A 1296-port MEMS transparent optical crossconnect with 5.1dB+/-1.1dB insertion loss at 1550nm is reported. Measured worst-case optical crosstalk in a fabric was ñ38dB and nominal switching rise/fall times were 5msec. A **2.07Petabit/s** switch capacity was verified upon cross-connecting a forty-channel by 40Gb/s DWDM data stream through a prototype fabric.

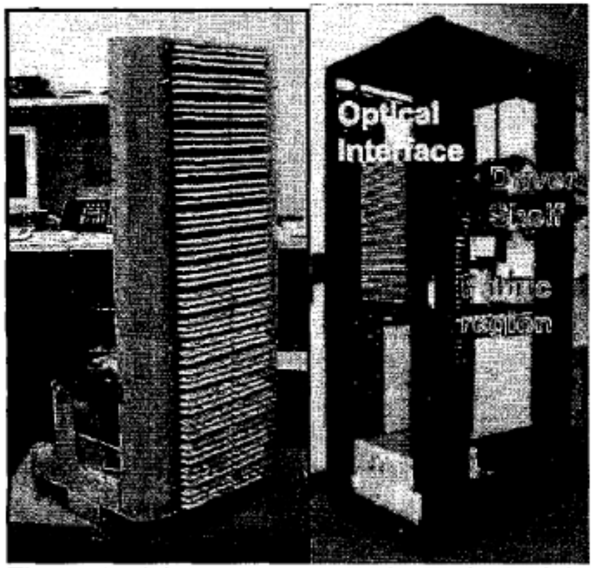


Figure 2. (a)1296-port optical switch module and (b) 648-port folded-geometry optical crossconnect prototype.

The bright side of optical switching

- Optical switches can have astonishing aggregate bandwidths
  - Absence of signal introspection
  - Possibility to receive dense WDM signals (> 1 Tb/s) on each port
- Translates into alluring $/Gb/s figures:

Type	# ports	Bandwidth per port	Total bandwidth	Price	Price per Gb/s	
Calient S320 Mems switch	320 ports	400Gbps (with 16 wavelengths at 25G – CDAUI-16 signaling)	128 Tb/s	$40k	**0.3 $/Gb/s**	
Mellanox SB7790	36 ports	100 Gbps (4xEDR Infiniband)	3.6 Tb/s	$12k	3.3 $/Gb/s	

→ Optical switching ~10 cheaper than electrical packet routing…

Optical switch vs. Electrical packet router

Optical switch: Bufferless

Electrical packet router: Buffered

- (Random access) buffering plays a crucial role
  - Without buffering, end-to-end scheduling required (no stopping in the exchange)
  - With buffering, scheduling made link after link
- Packet routers also allow differentiated QoS, error correction, etc.

→ **Packet router offers much higher "value" than optical switches**

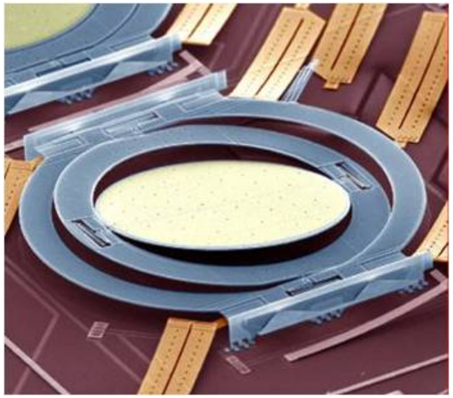
→ My personal guess:

→ at least 30x more value (for 10ns optical switching time)

→ $/Gb/s and mW/Gb/s figures of optical switches must be at least 30x lower

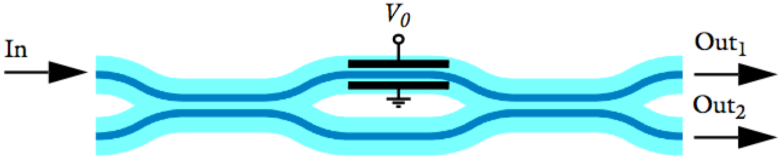
Technologies for optical switches

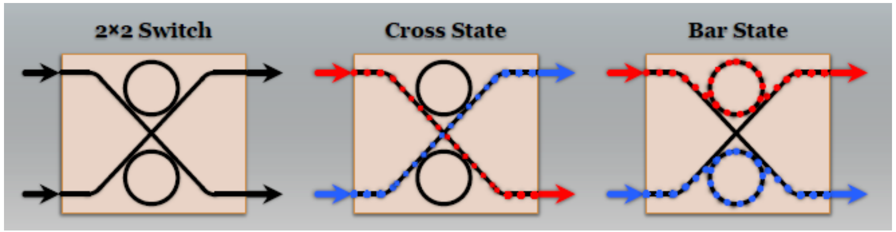
- MEMS-based Switches



- Broadband
- Low loss
- Bulky
- Slow
- Unaffordable?

- SiP-based Switches





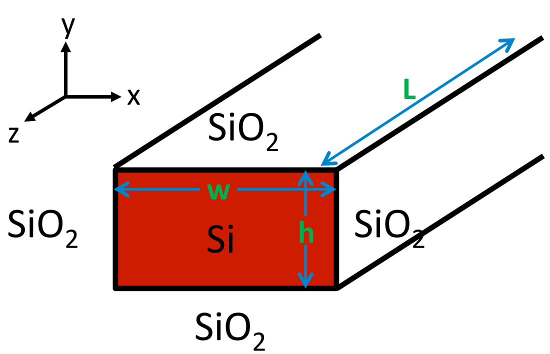
- Based on optical waveguides
- Fast (carrier based)
- Somewhat Broadband (MZI) **OR** Wavelength Selective (ring)
- Lossy
- High integration

  = cheap ?

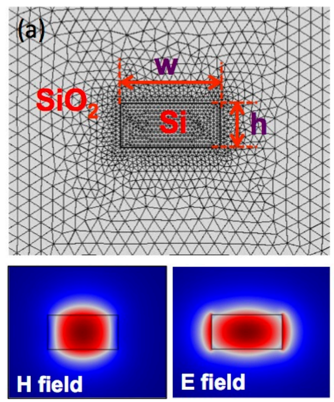
MZI SiP Switch: Analysis & Design

Wg Mode Analysis → Compact Model

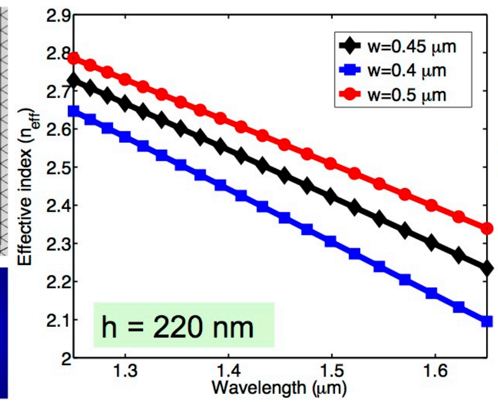
$$[S] = \begin{bmatrix} 0 & \sqrt{L}\exp(-j\phi) \\ \sqrt{L}\exp(-j\phi) & 0 \end{bmatrix}$$



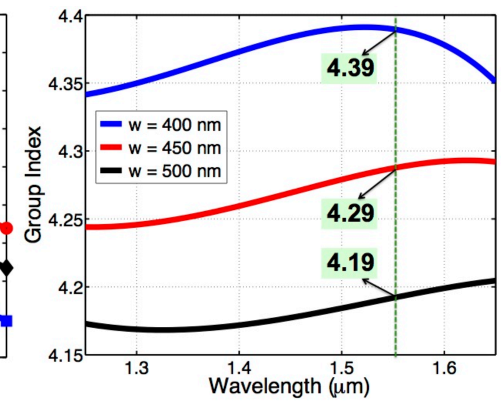
Strip Waveguide



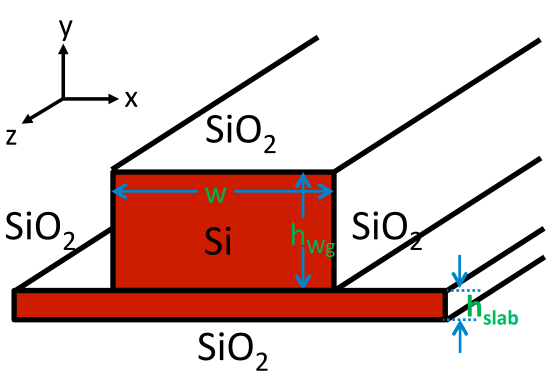
Mode Profile



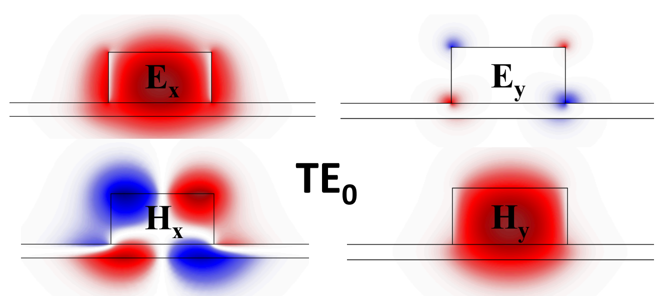
Effective Index



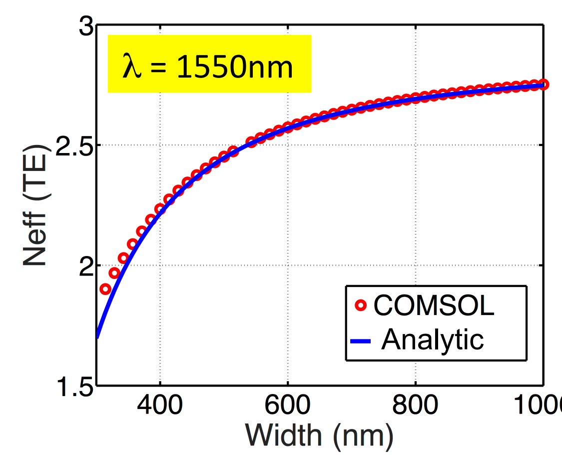
Group Index



Ridge Waveguide

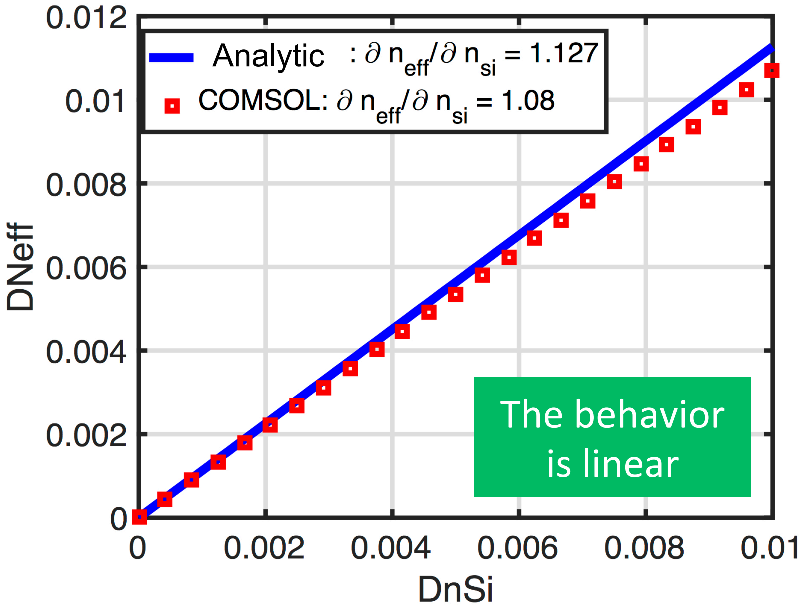


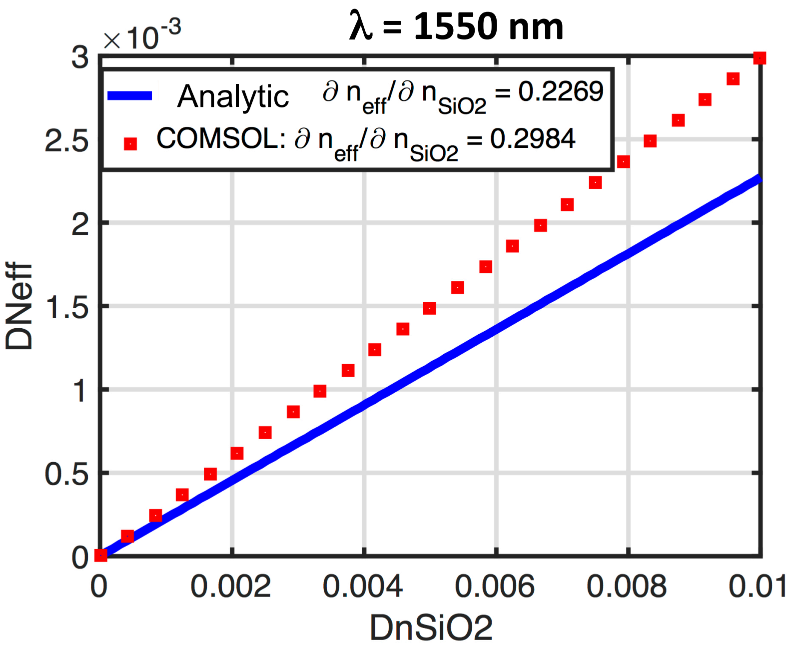
Mode Profile



MZI SiP Switch: Analysis & Design

Wg Mode Analysis → Compact Model → Perturbation Effects





MZI SiP Switch: Analysis & Design

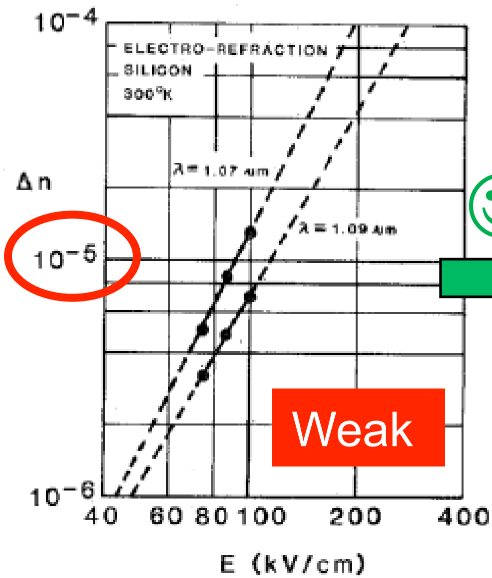
Wg Mode Analysis → Compact Model → Perturbation → Plasma Dispersion

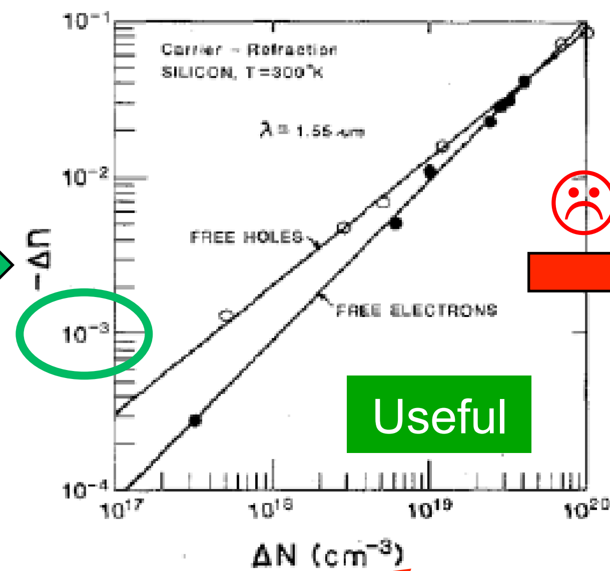
$$\Delta n_{\mathrm{Si}} = f(\Delta N, \Delta P) \qquad \Delta \alpha_{\mathrm{Si}} = g(\Delta N, \Delta P)$$

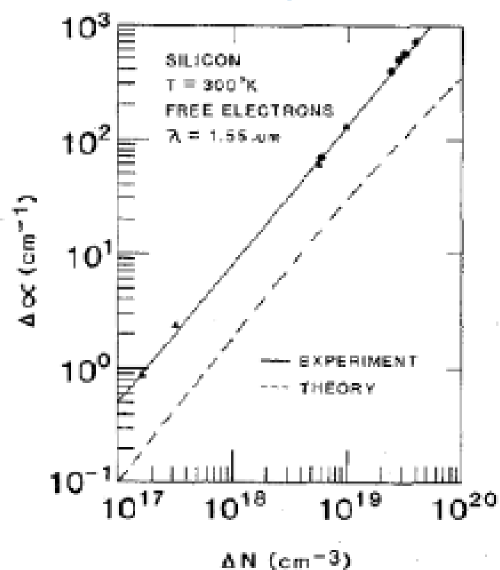
Electro-refraction

Change of index of refraction

Change of optical absorption



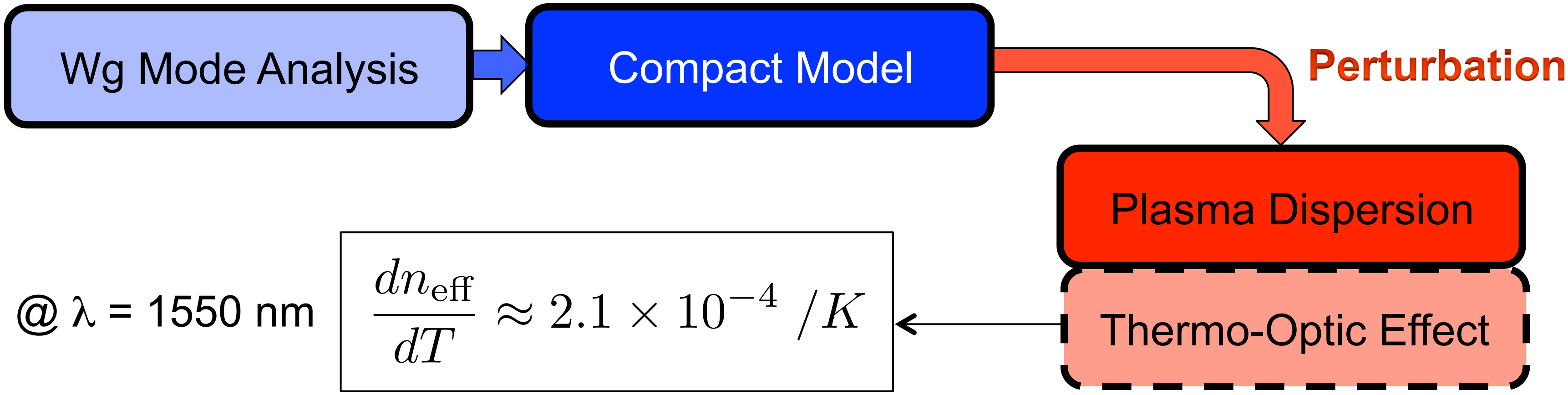




Charge carrier density

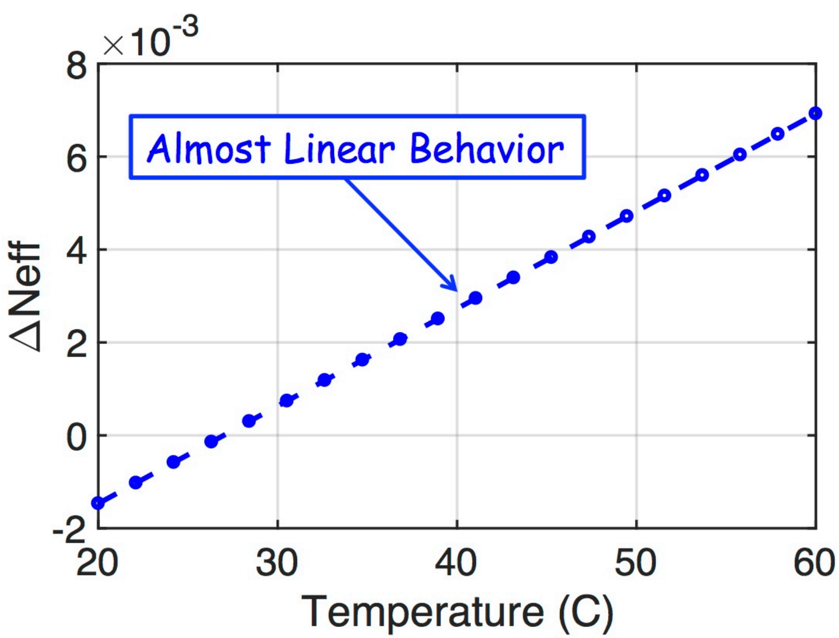
[*R. Soref, IEE JQE, Vol Q-23, No1, 1987*]

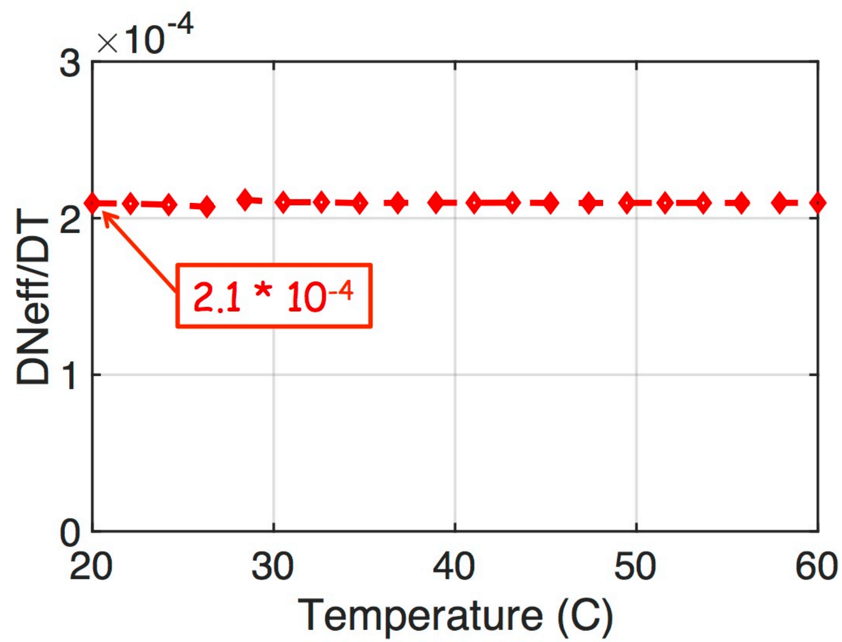
MZI SiP Switch: Analysis & Design



@ λ = 1550 nm

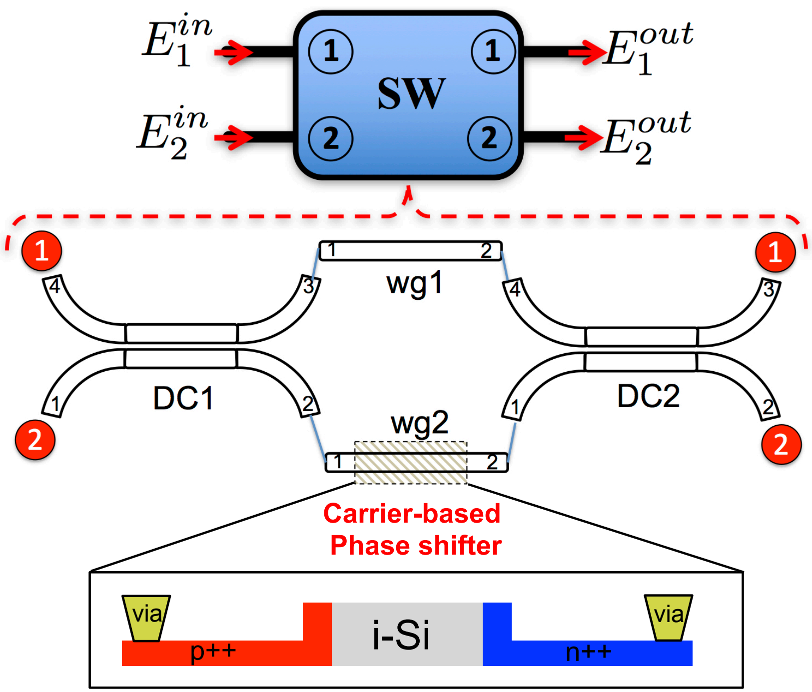
$$\frac{dn_{\text{eff}}}{dT} \approx 2.1 \times 10^{-4} \ /K$$





MZI SiP Switch: Analysis & Design

Wg Mode Analysis → Compact Model → (Perturbation) → Plasma Dispersion / Thermo-Optic Effect → Optical Response → Optimize 2x2 MZI



Phase Shifter

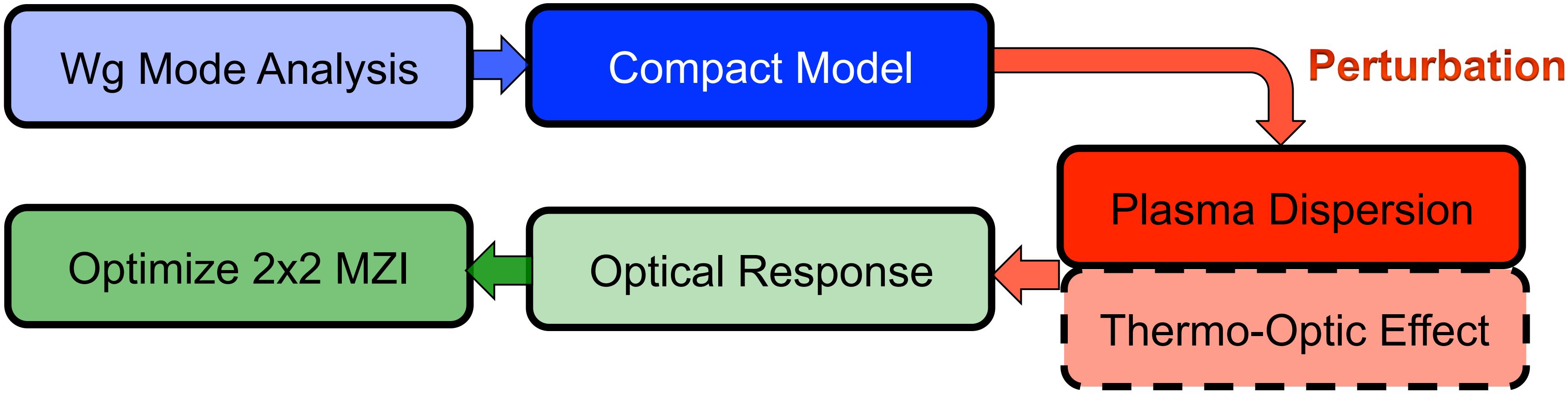
$$\Delta n_{\mathrm{Si}} = f(\Delta N, \Delta P)$$
$$\Delta \alpha_{\mathrm{Si}} = g(\Delta N, \Delta P)$$

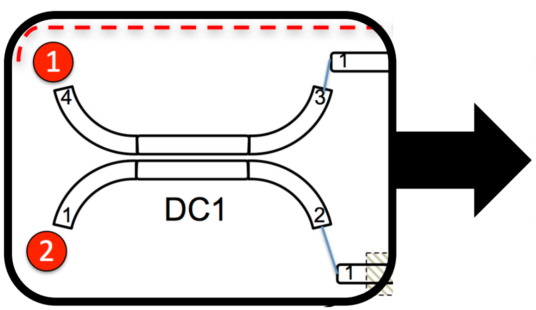
$$\Delta n_{\mathrm{eff}} = \frac{\partial n_{\mathrm{eff}}}{\partial n_{\mathrm{Si}}} \times f(\Delta N, \Delta P)$$

1.1

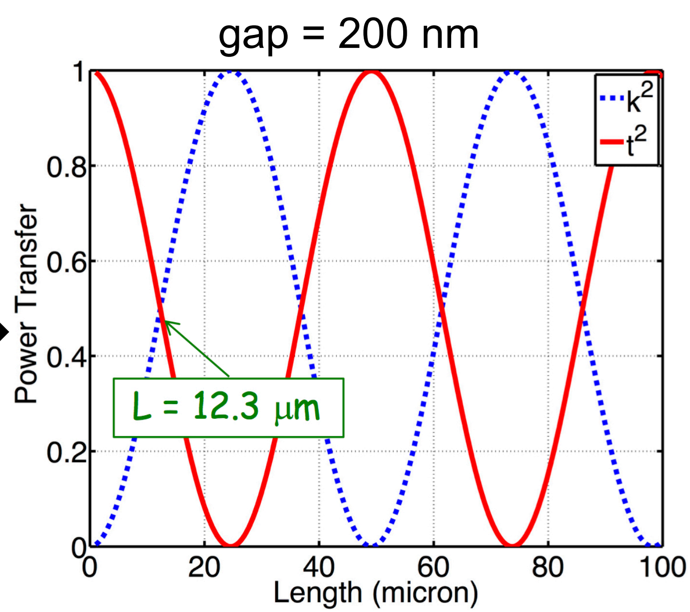
$$\Delta \phi = \frac{2\pi}{\lambda} \times \Delta n_{\mathrm{eff}} \times \mathrm{length}$$

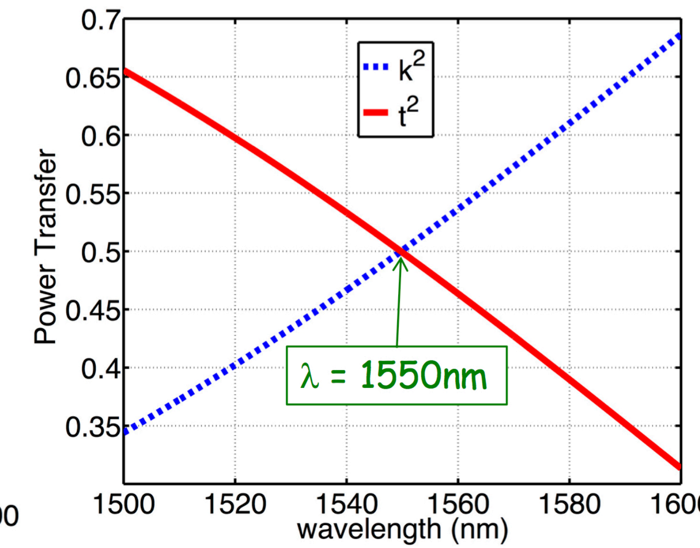
MZI SiP Switch: Analysis & Design



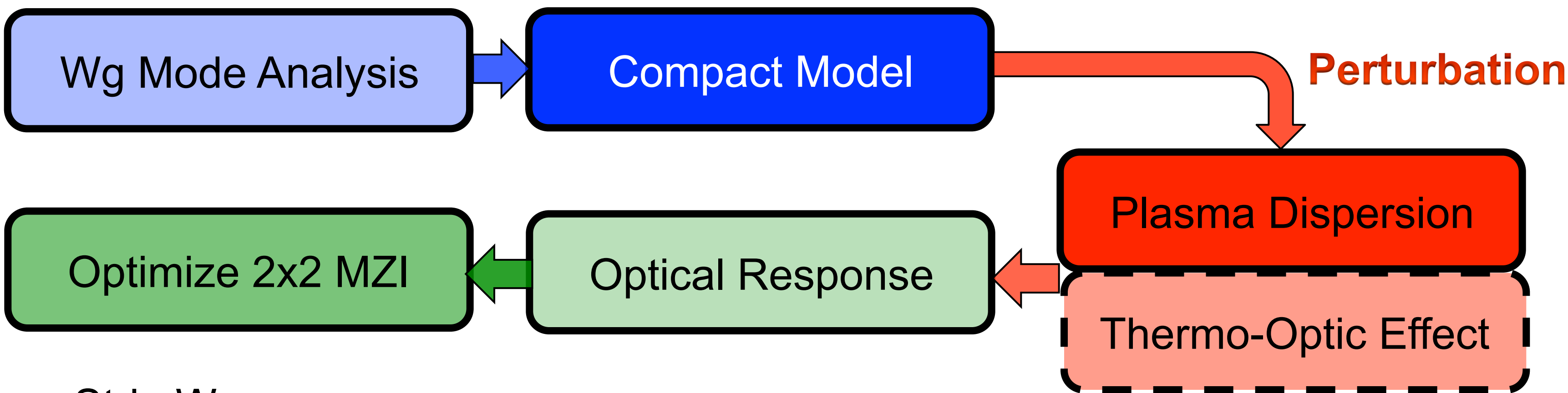


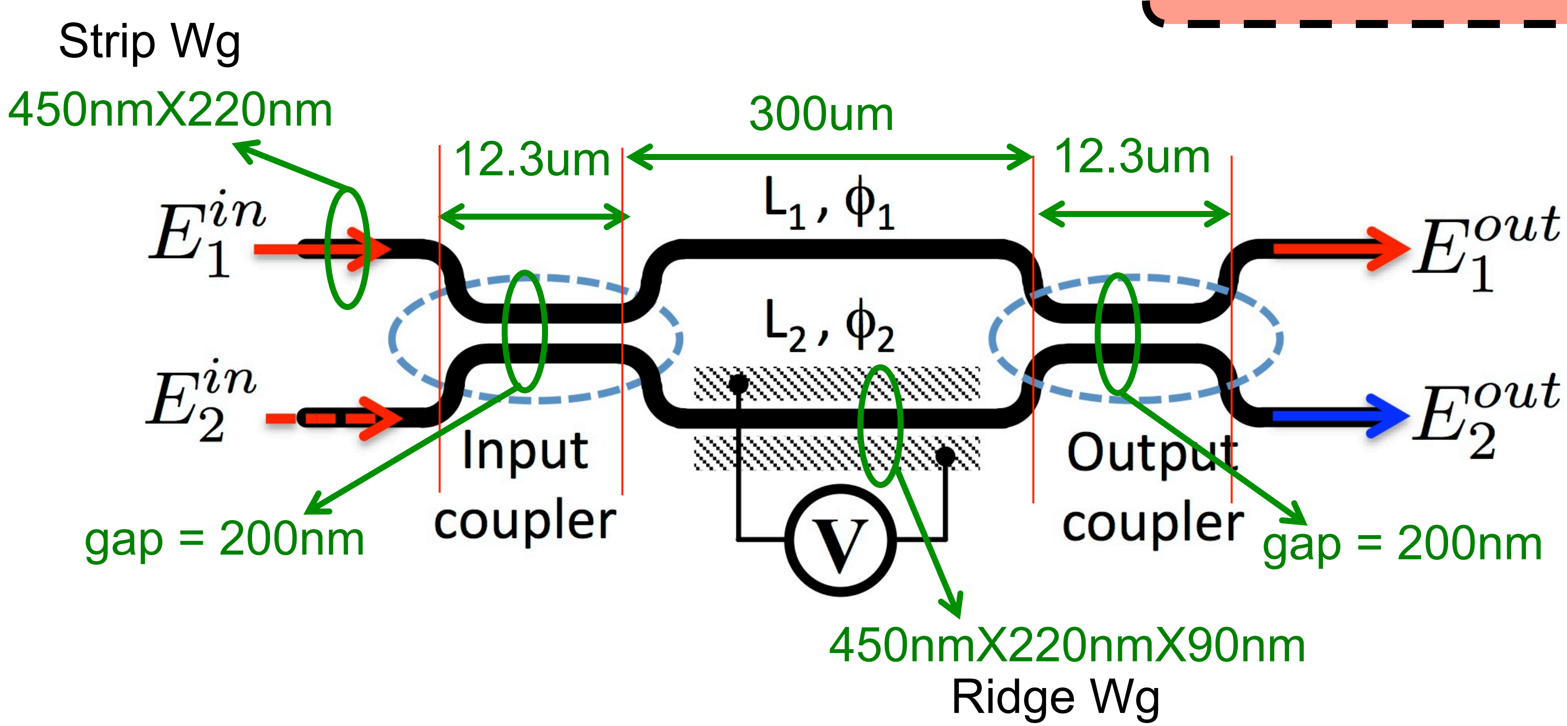
50-50 Coupler
@ λ = 1550 nm



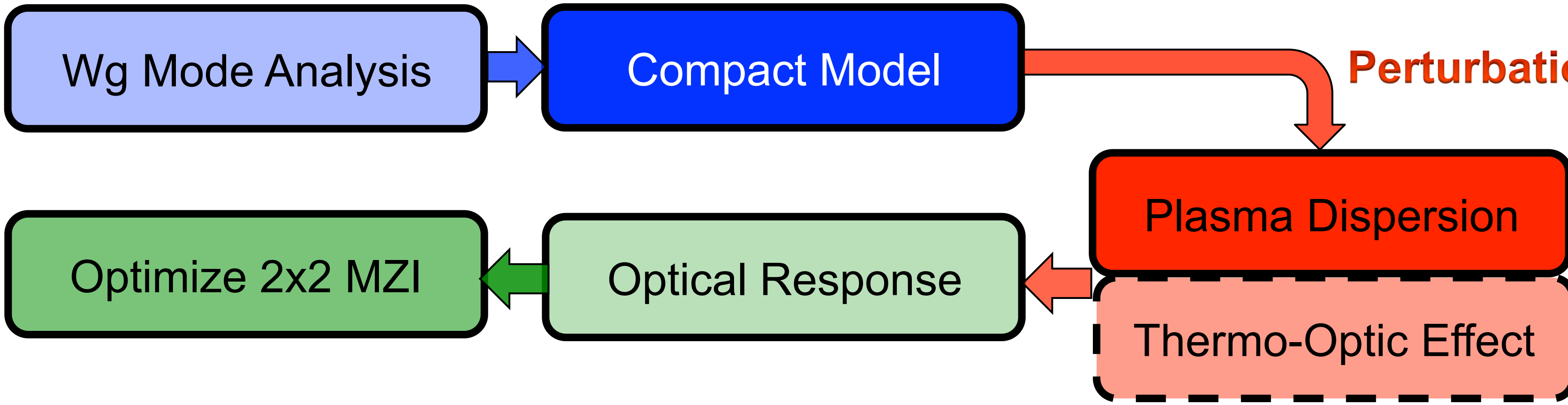


MZI SiP Switch: Analysis & Design

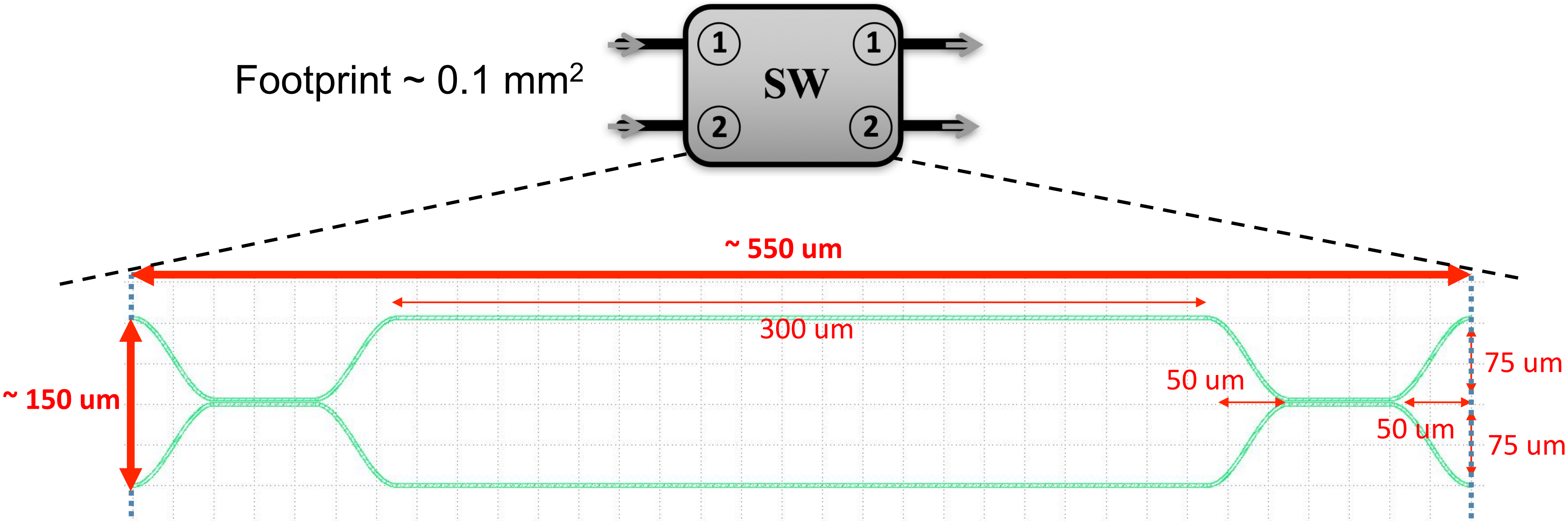




MZI SiP Switch: Analysis & Design

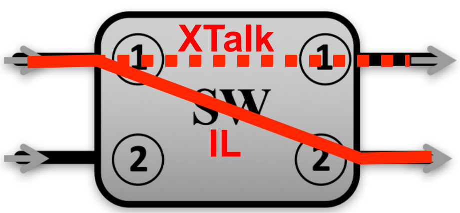


Footprint ~ 0.1 mm^2

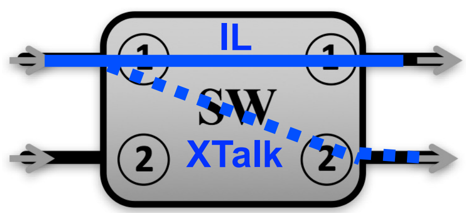


Optimization of MZI Switch

Topology		2x2 Switch			
Excitation Port		**Port 1 IN**		**Port 2 IN**	
Parameter		Value	State	Value	State
Insertion Loss (dB)	Port 1 OUT	1.2	B	0.25	C
	Port 2 OUT	0.25	C	1.2	B
Xtalk Level (dB)	Port 1 OUT	-∞	C	-19.8	B
	Port 2 OUT	-19.8	B	-∞	C



Cross-state



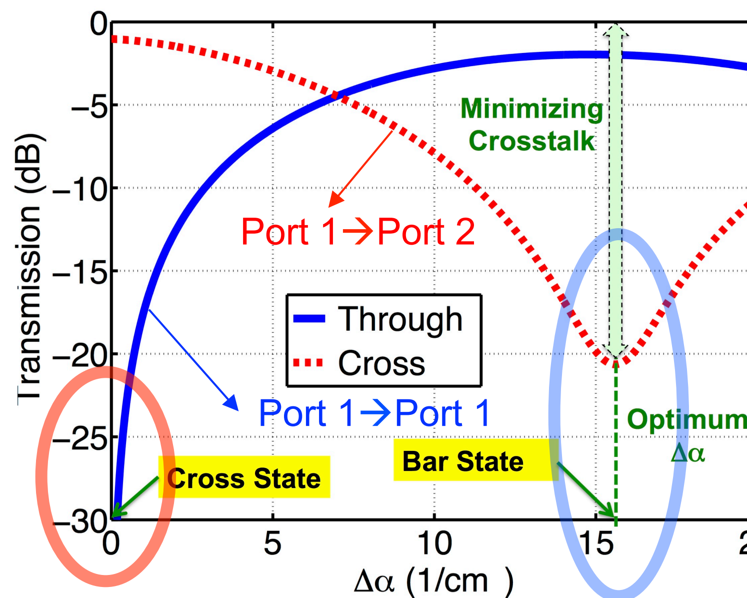
Bar-state

$$\mathrm{CROSS} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \mathrm{BAR} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Transition Matrix of MZI Switch

IBM MZI paper: [*N. Dupuis et al., IEEE JLT vol. 33, no. 17, 2015*]

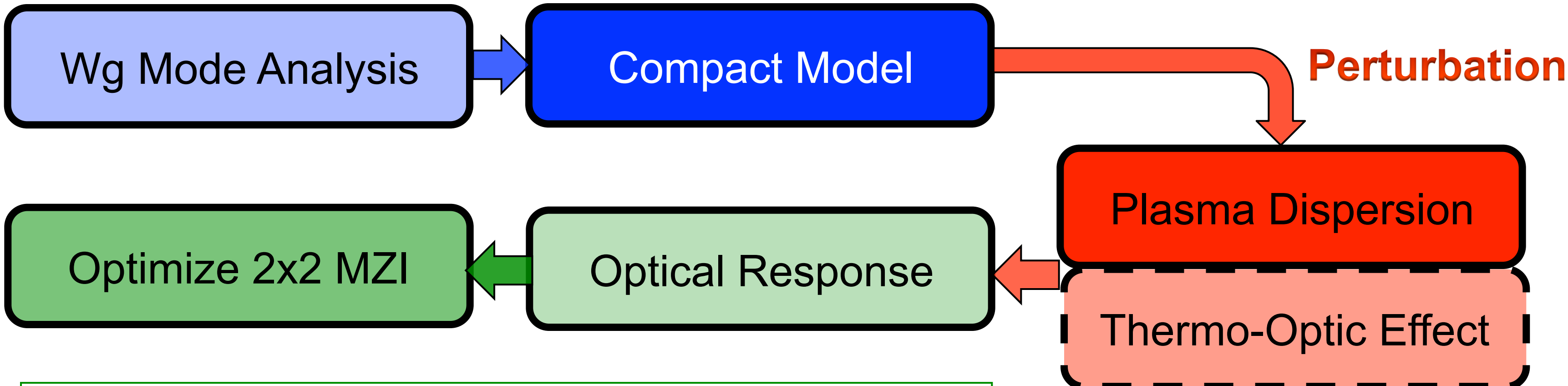
IL = 1.6 dB , Xtalk = -21.1 dB



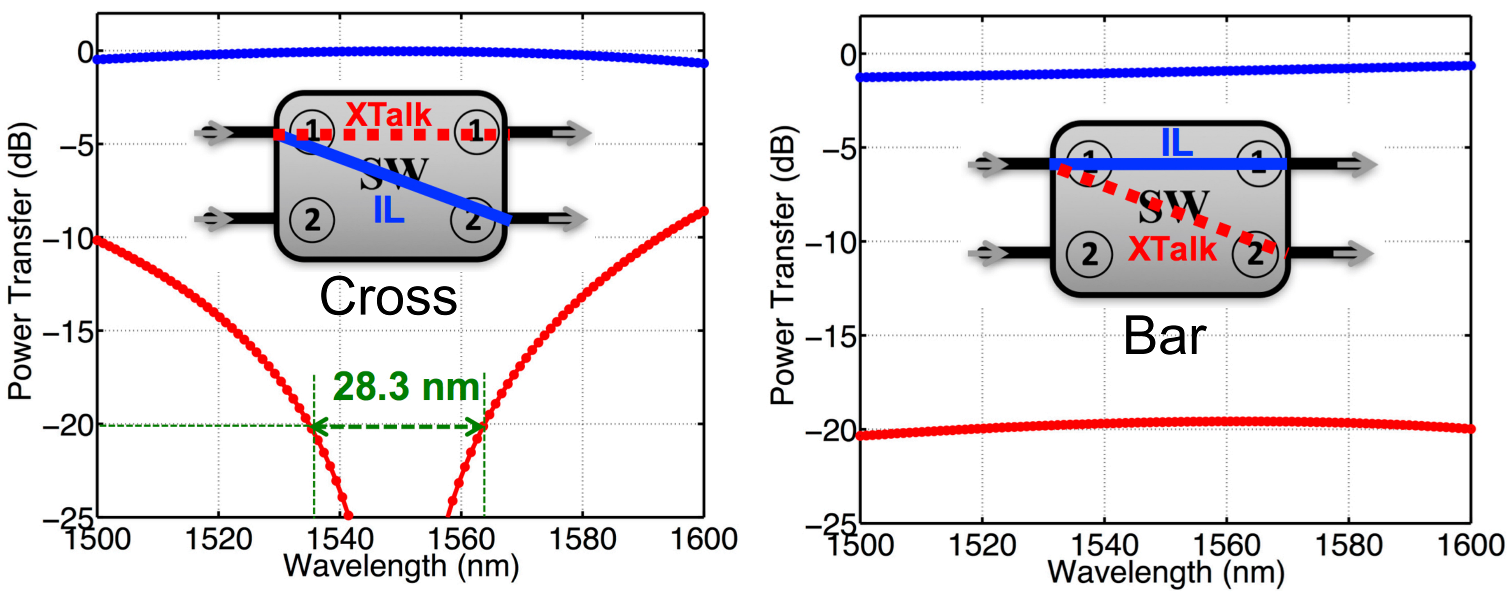
(excess loss of phase shifter)

$$\left|\frac{E_{\mathrm{out}}}{E_{\mathrm{in}}}\right|_{\text{(phase shifter)}} = e^{-(\Delta\alpha \times l)/2}$$

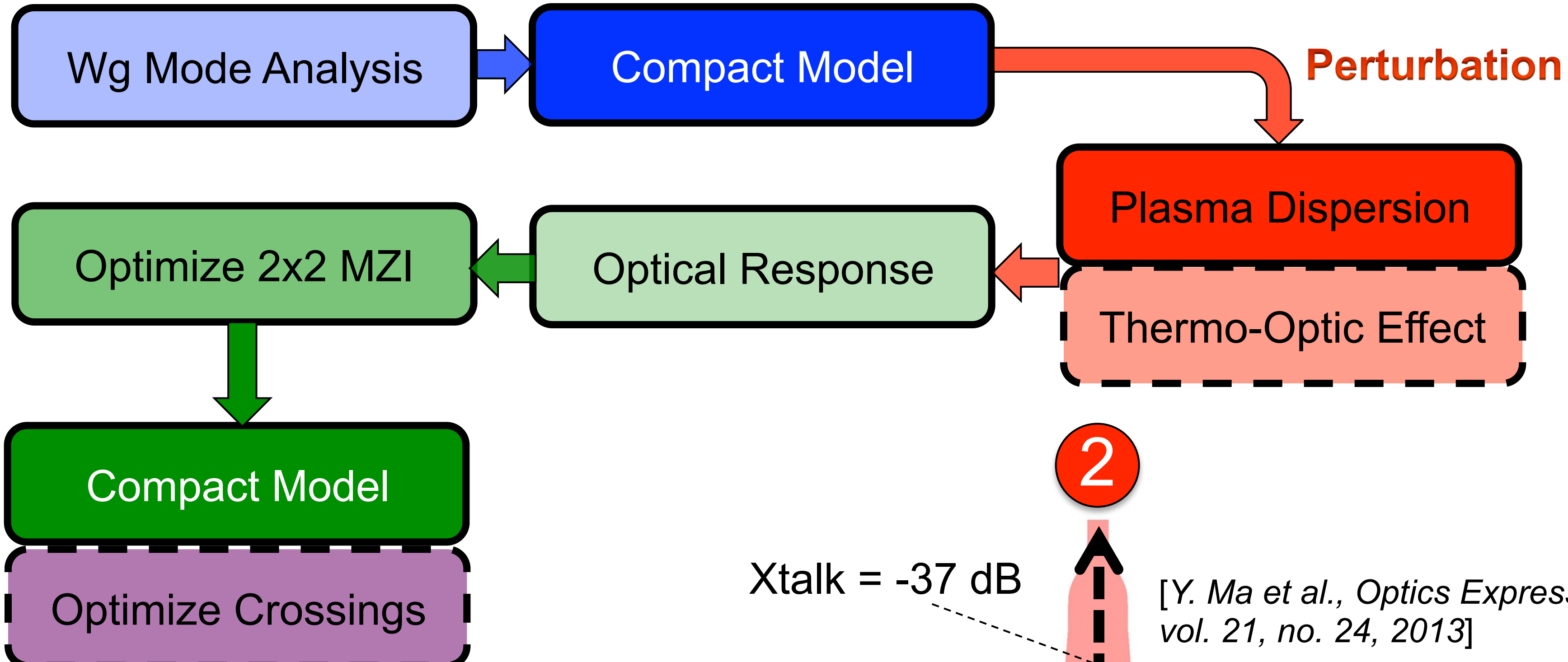
MZI SiP Switch: Analysis & Design



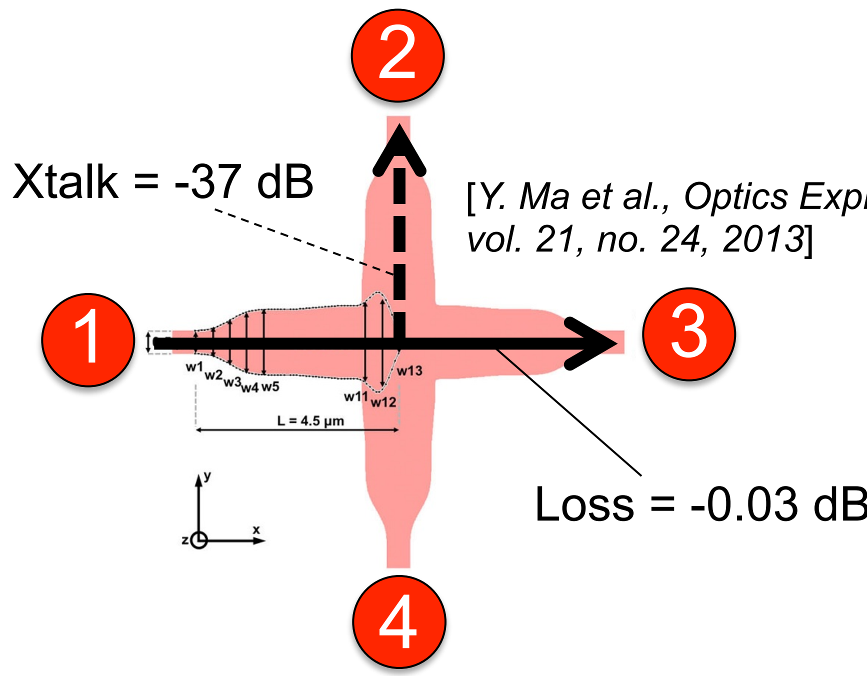
-20 dB Xtalk Bandwidth of our design = 28.3 nm



MZI SiP Switch: Analysis & Design

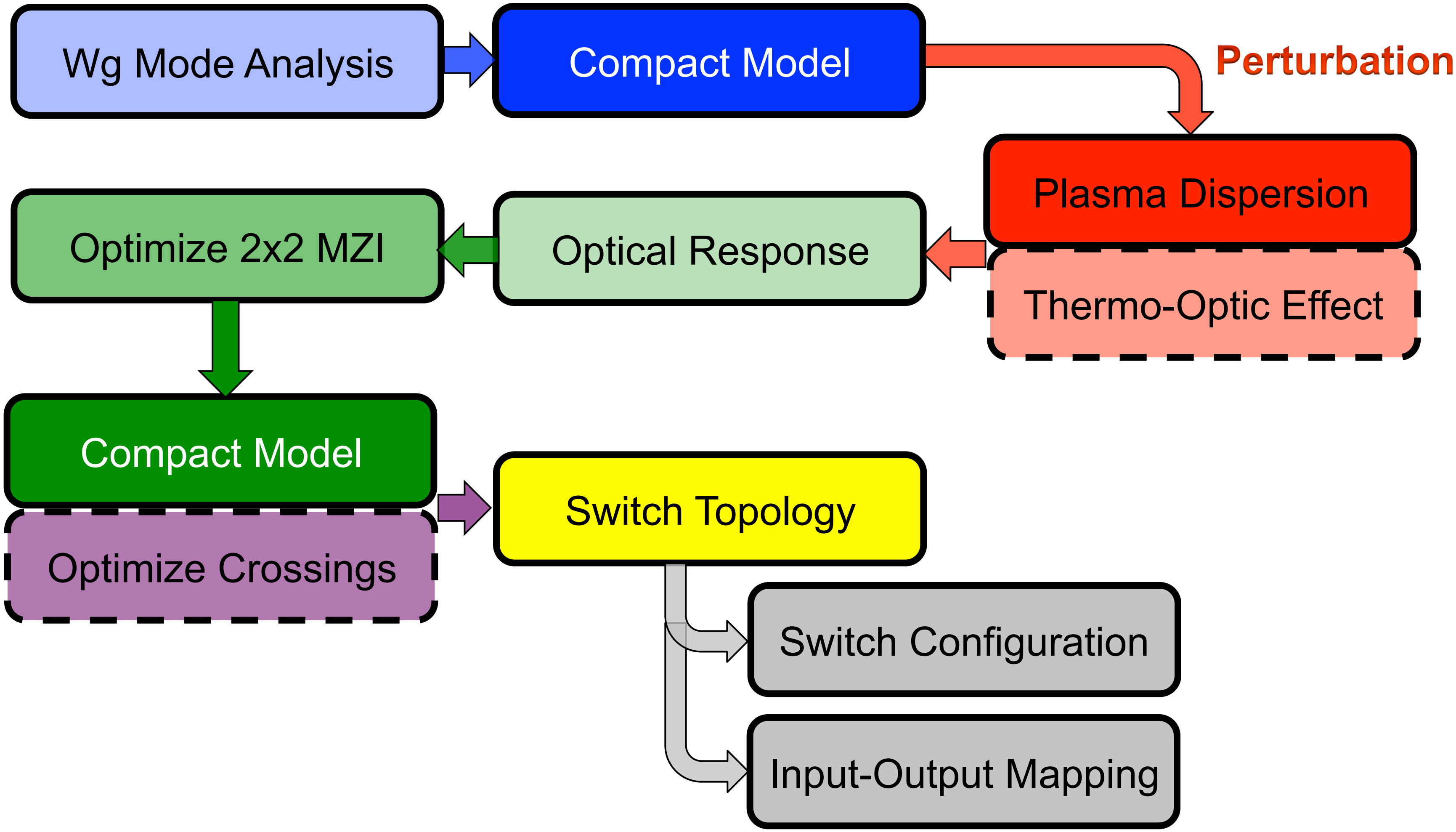


$$\mathbf{S} = \begin{bmatrix} R & jx & t & -jx \\ jx & R & -jx & t \\ t & -jx & R & jx \\ -jx & t & jx & R \end{bmatrix}$$

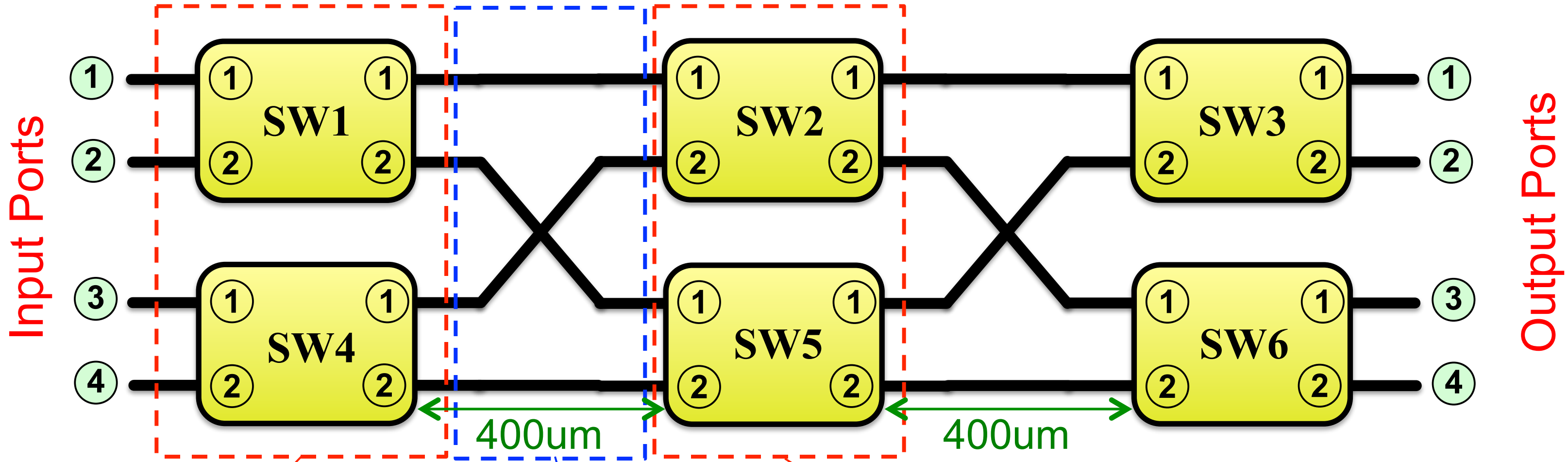


[*Y. Ma et al., Optics Express, vol. 21, no. 24, 2013*]

MZI SiP Switch: Analysis & Design



4x4 Benes Topology: Configuration → Mapping



First Column

Shuffle

Second Column

. . .

SW1, SW4

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Transition Matrix
C1

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix
X1

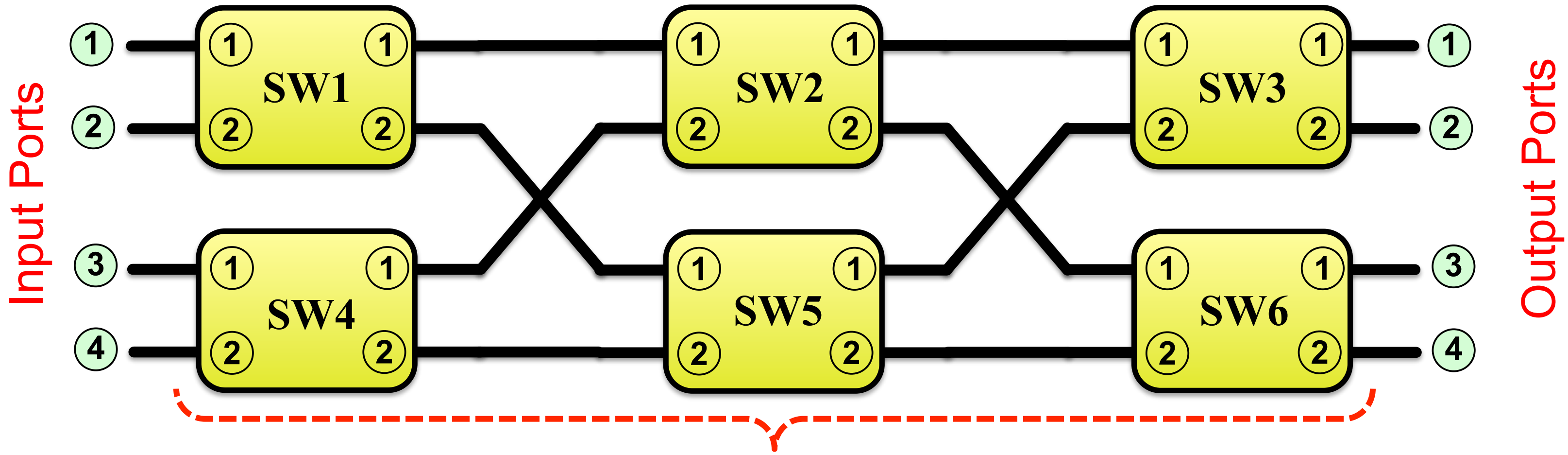
SW2, SW5

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Transition Matrix
C2

. . .

4x4 Benes Topology: Configuration → Mapping



$$\mathbf{T} = \mathbf{C3} \times \mathbf{X2} \times \mathbf{C2} \times \mathbf{X1} \times \mathbf{C1}$$

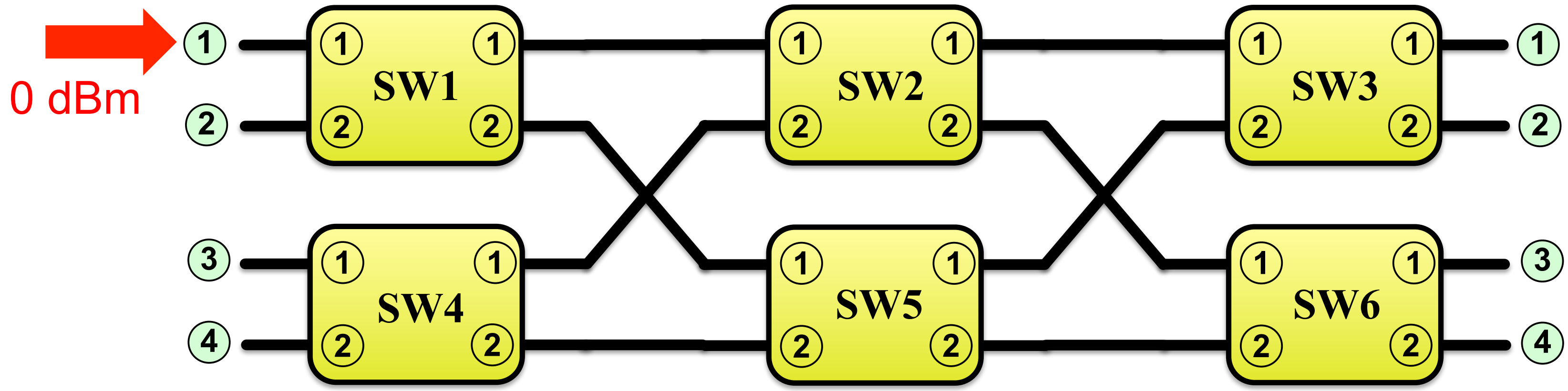
- Which output port does input Port 1 goes to?

One element is "1" and the rest is "0"

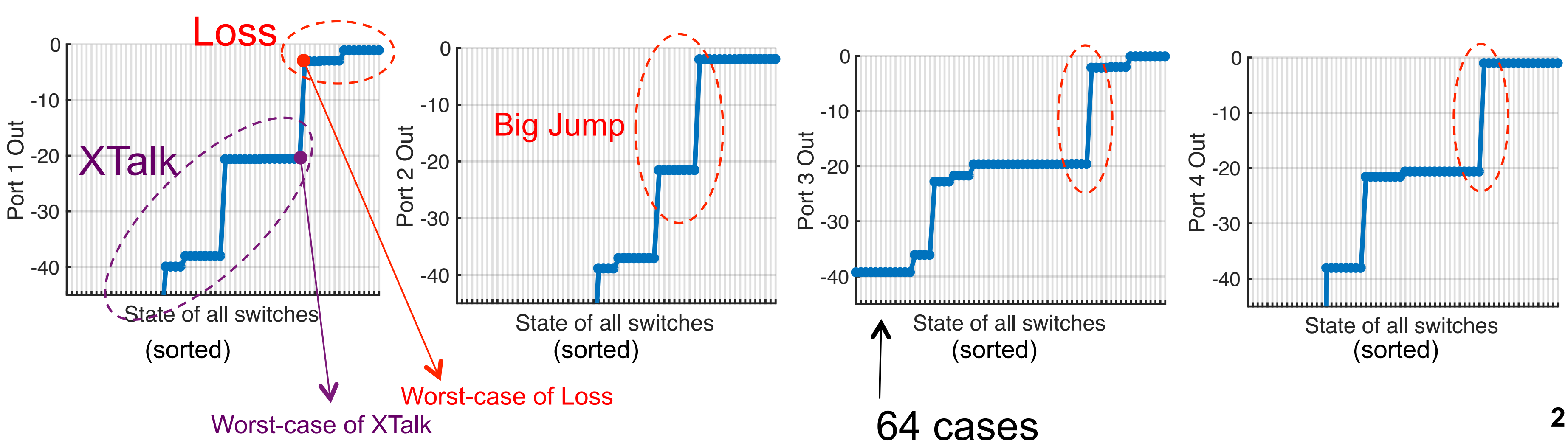
Output port 1
Output port 2
Output port 3
Output port 4

$$\begin{pmatrix} ? \\ ? \\ ? \\ ? \end{pmatrix} = \mathbf{T} \times \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

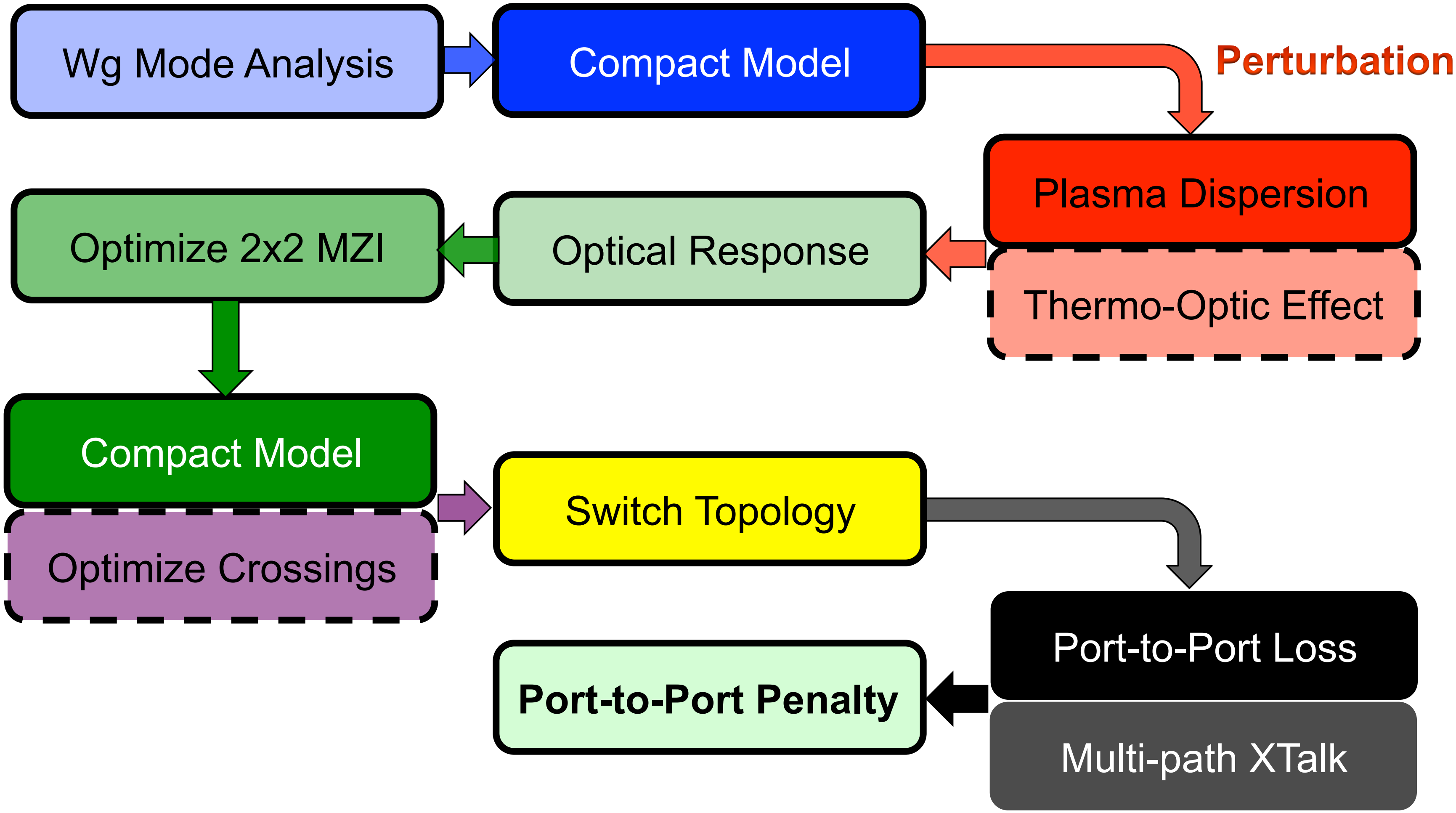
4x4 Benes: Let's Excite Input Port 1



- Monitor the output ports for all switch configurations
- Sort the switch configurations based on the output power of each port

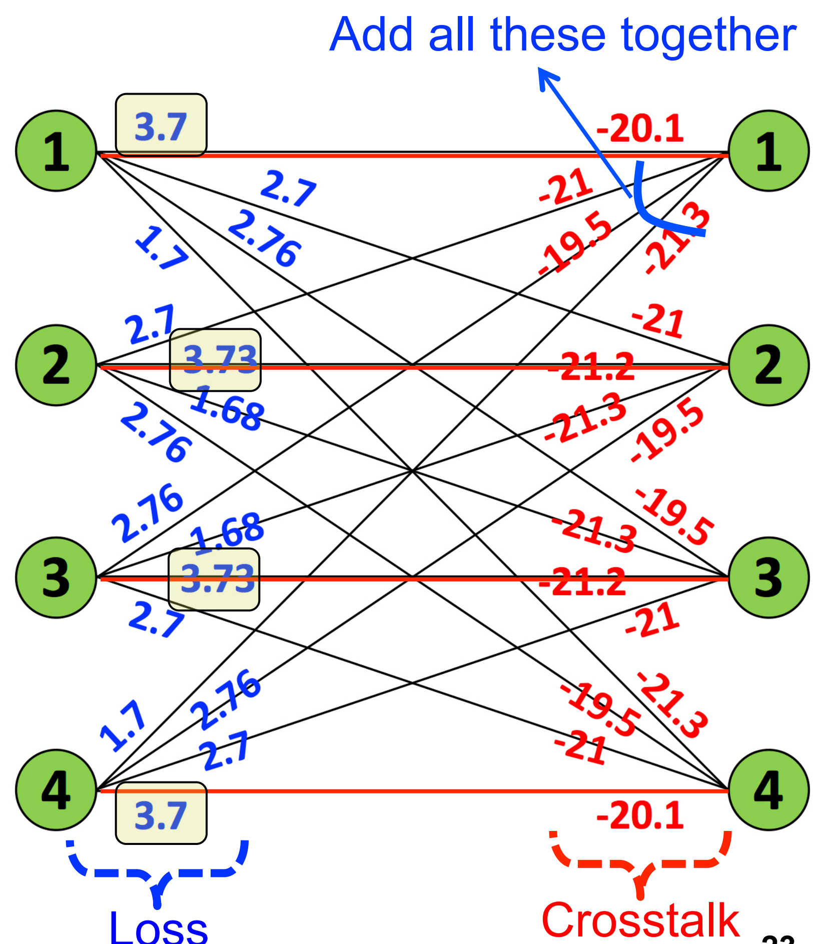


MZI SiP Switch: Analysis & Design

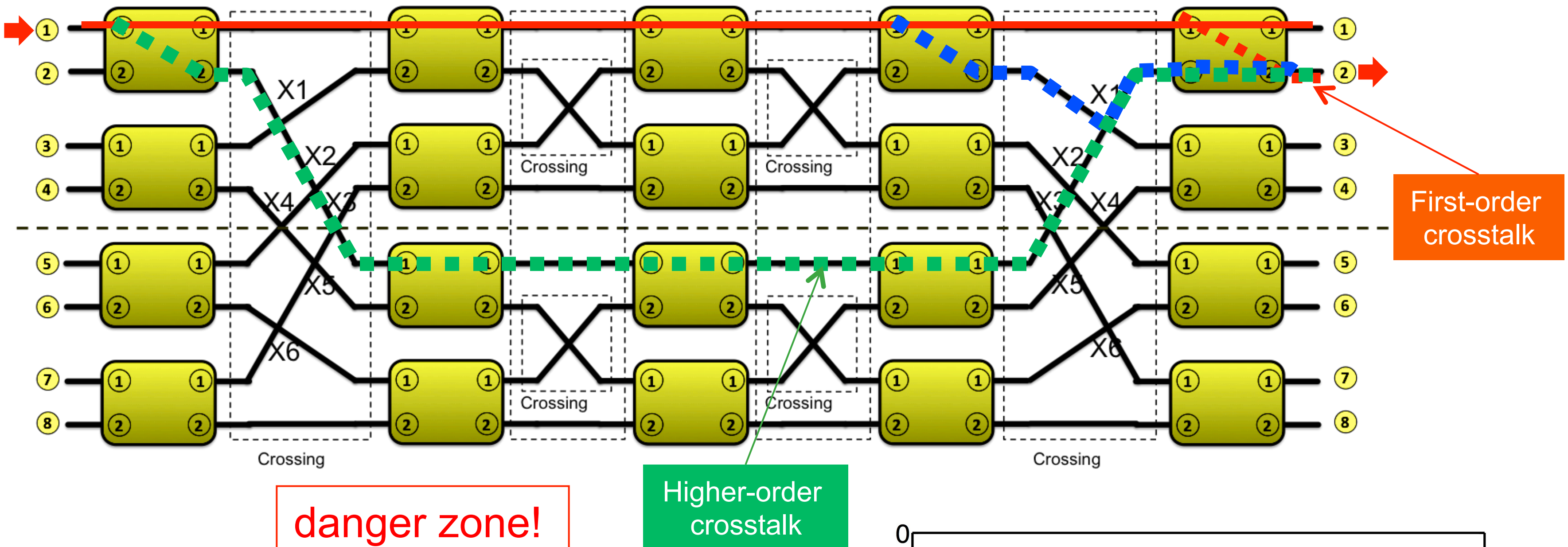


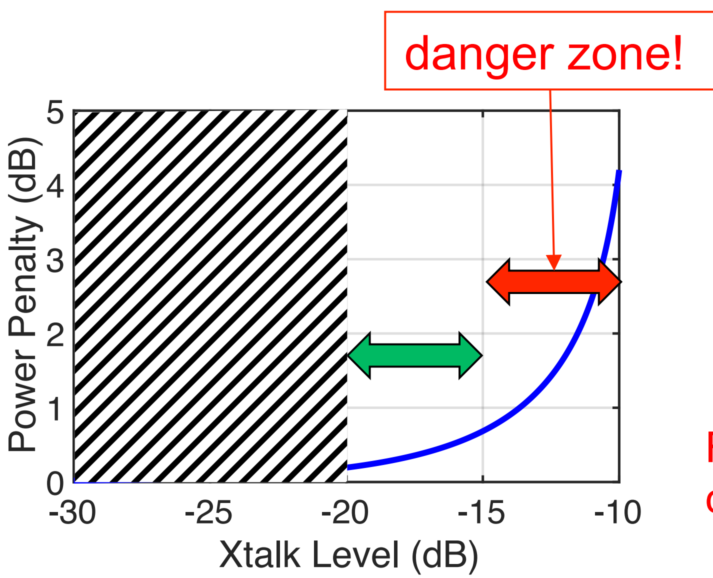
4x4 Benes Topology: Worst-case Graph

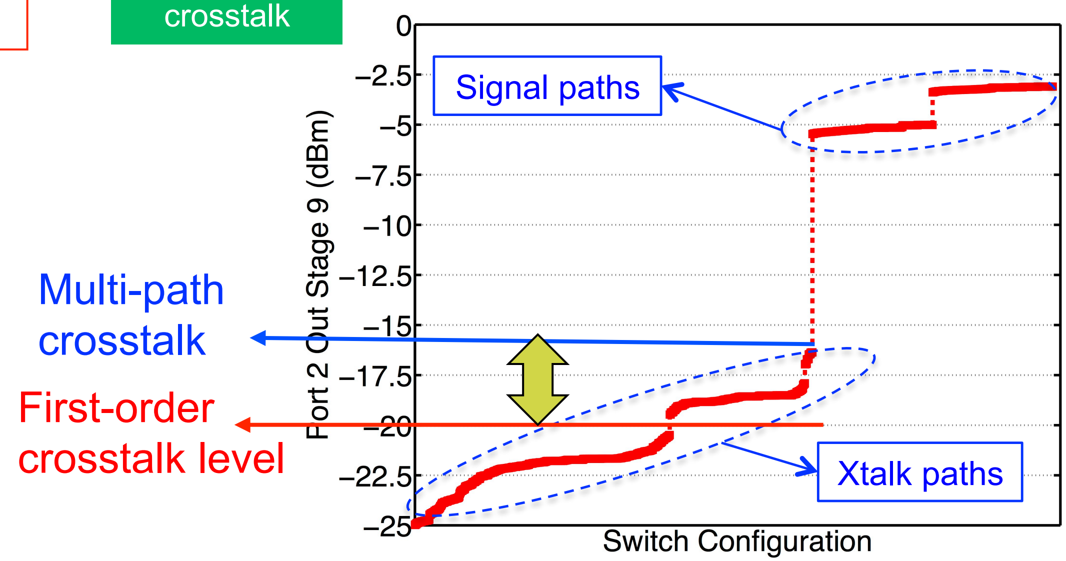
- What is the use of this (input, output) graph?
- Assume the (1,2,3,4)→(1,2,3,4) mapping
  - Worst case of insertion loss ?
    - ➢ 3.73 dB: (2→2) and (3→3)
  - Worst case of crosstalk?
    - ➢ Add all the worst-case crosstalks
    - ➢ (-21dB)+(-19.5dB)+(-21.3dB)
  - ➢ Resulting SNR: 12dB



8x8 Benes Topology: Higher-order Crosstalk



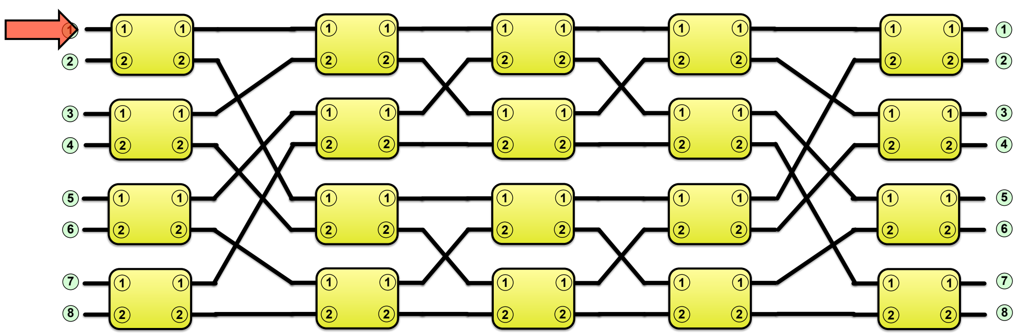


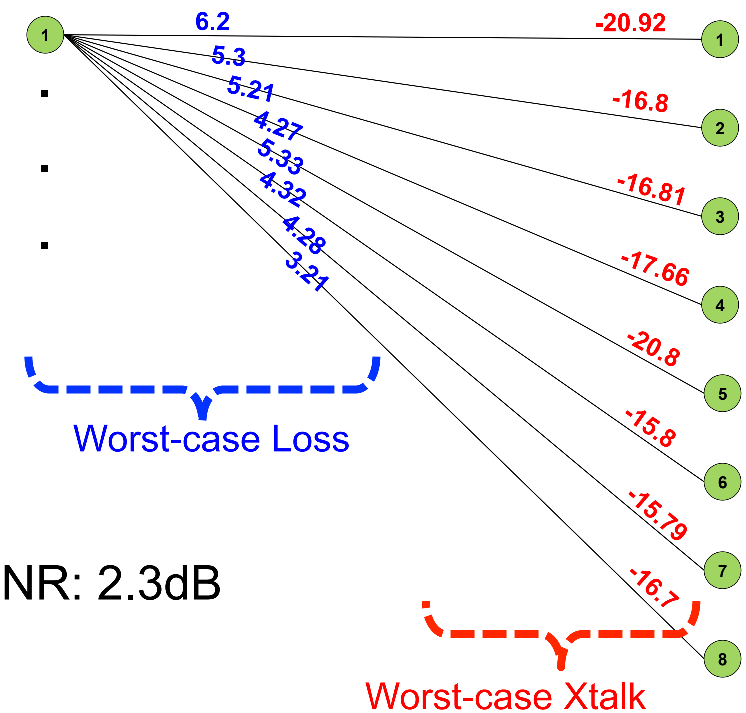


[Bahadori *et al.*, IEEE JLT, vol. 34 no.17, 2016]

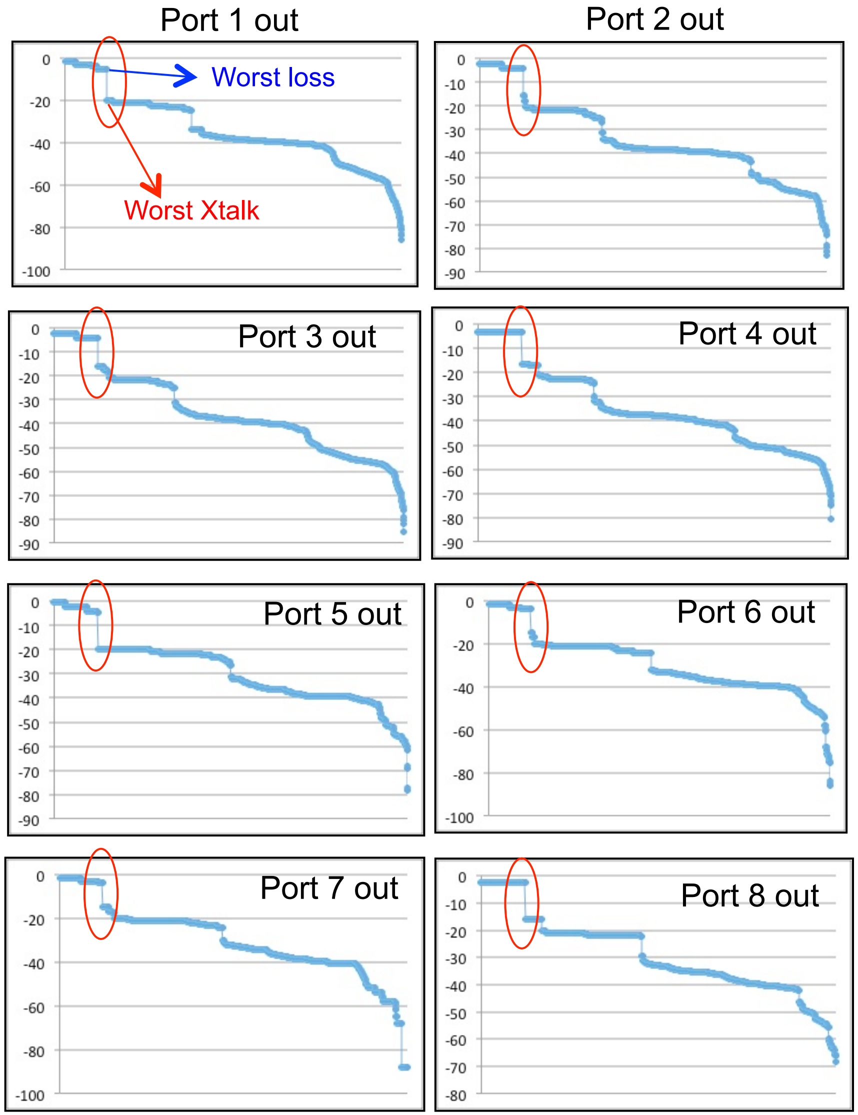
8x8 Benes Topology: Worst-Case Graph

- Example: Let's excite input port 1





SNR: 2.3dB



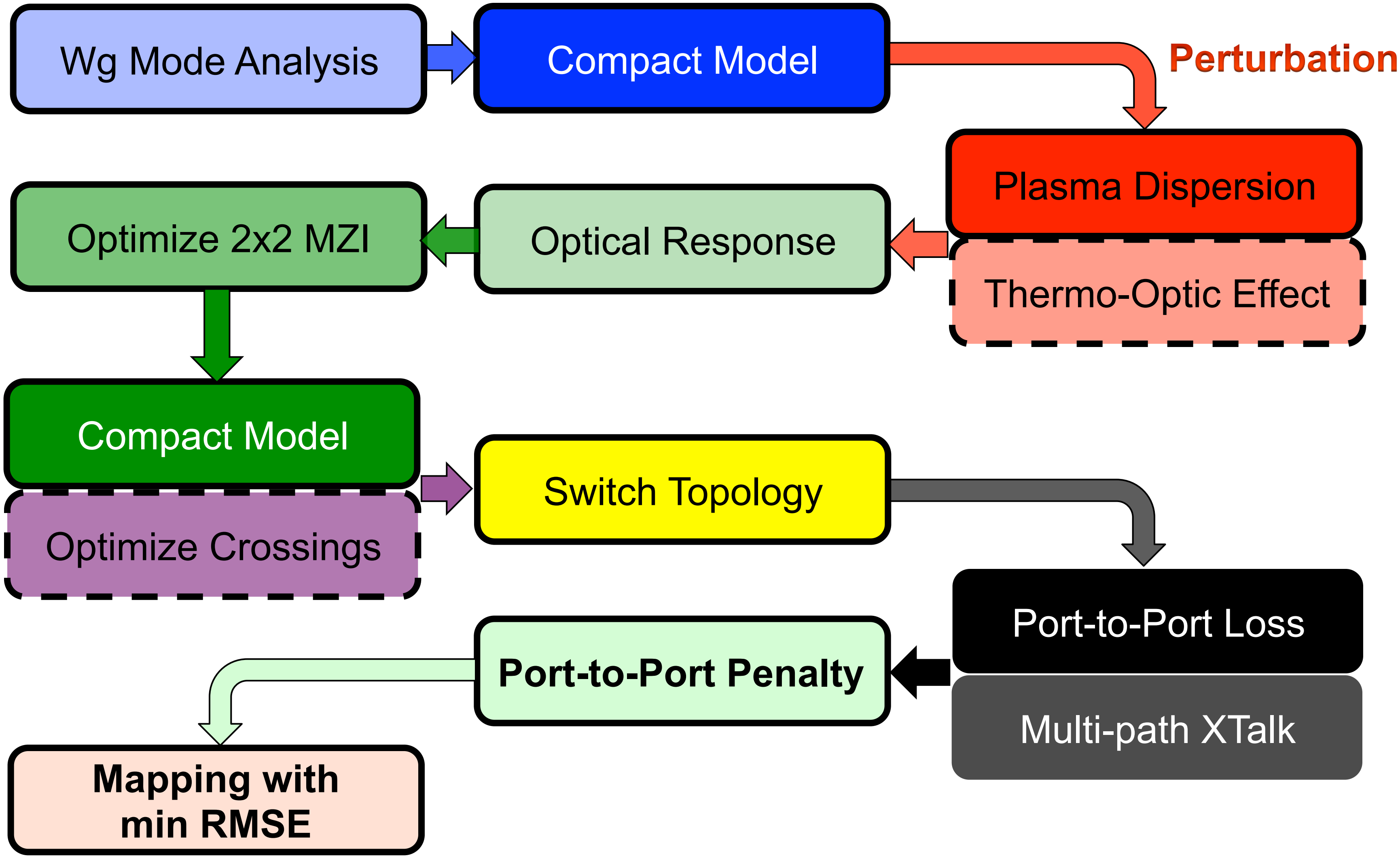
4x4 Benes Topology: Configuration to Mapping

- 2^6 = 64 Switch Configurations, 4! = 24 mappings → ***mappings repeat***

BB|BB|BB (1,2,3,4)	BC|BB|BB (1,2,4,3)	BB|BC|BB (1,4,3,2)	BC|BC|BB (1,4,2,3)	BB|BB|BC (1,2,4,3)	BC|BB|BC (1,2,3,4)	BB|BC|BC (1,3,4,2)	BC|BC|BC (1,3,2,4)
CB|BB|BB (2,1,3,4)	CC|BB|BB (2,1,4,3)	CB|BC|BB (4,1,3,2)	CC|BC|BB (4,1,2,3)	CB|BB|BC (2,1,4,3)	CC|BB|BC (2,1,3,4)	CB|BC|BC (3,1,4,2)	CC|BC|BC (3,1,2,4)
BB|CB|BB (3,2,1,4)	BC|CB|BB (3,2,4,1)	BB|CC|BB (3,4,1,2)	BC|CC|BB (3,4,2,1)	BB|CB|BC (4,2,1,3)	BC|CB|BC (4,2,3,1)	BB|CC|BC (4,3,1,2)	BC|CC|BC (4,3,2,1)
CB|CB|BB (2,3,1,4)	CC|CB|BB (2,3,4,1)	CB|CC|BB (4,3,1,2)	CC|CC|BB (4,3,2,1)	CB|CB|BC (2,4,1,3)	CC|CB|BC (2,4,3,1)	CB|CC|BC (3,4,1,2)	CC|CC|BC (3,4,2,1)
BB|BB|CB (2,1,3,4)	BC|BB|CB (2,1,4,3)	BB|BC|CB (2,4,3,1)	BC|BC|CB (2,4,1,3)	BB|BB|CC (2,1,4,3)	BC|BB|CC (2,1,3,4)	BB|BC|CC (2,3,4,1)	BC|BC|CC (2,3,1,4)
CB|BB|CB (1,2,3,4)	CC|BB|CB (1,2,4,3)	CB|BC|CB (4,2,3,1)	CC|BC|CB (4,2,1,3)	CB|BB|CC (1,2,4,3)	CC|BB|CC (1,2,3,4)	CB|BC|CC (3,2,4,1)	CC|BC|CC (3,2,1,4)
BB|CB|CB (3,1,2,4)	BC|CB|CB (3,1,4,2)	BB|CC|CB (3,4,2,1)	BC|CC|CB (3,4,1,2)	BB|CB|CC (4,1,2,3)	BC|CB|CC (4,1,3,2)	BB|CC|CC (4,3,2,1)	BC|CC|CC (4,3,1,2)
CB|CB|CB (1,3,2,4)	CC|CB|CB (1,3,4,2)	CB|CC|CB (4,3,2,1)	CC|CC|CB (4,3,1,2)	CB|CB|CC (1,4,2,3)	CC|CB|CC (1,4,3,2)	CB|CC|CC (3,4,2,1)	CC|CC|CC (3,4,1,2)

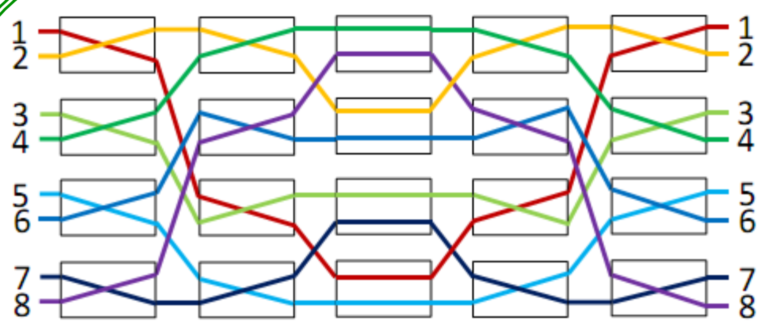
Mapping: (1,2,3,4) → (?,?,?,?)

MZI SiP Switch: Analysis & Design



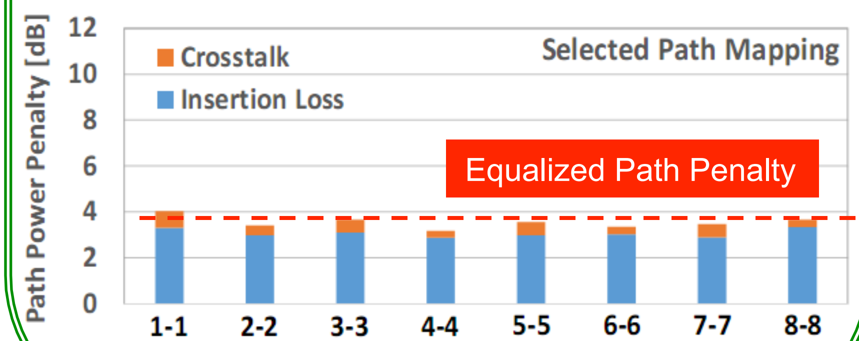
8x8 Benes Topology: Equalized Penalty Mapping

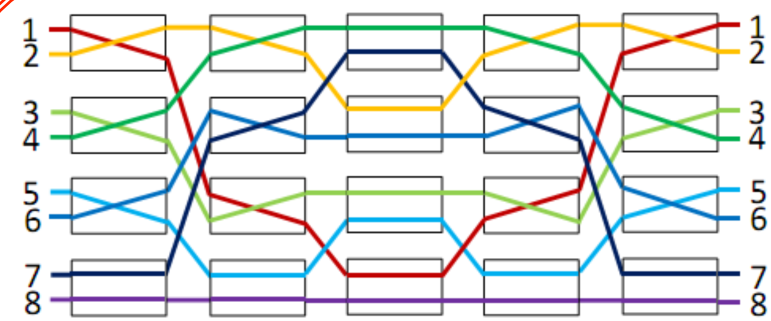
- How much improvement can we gain for (1,2,3,4,5,6,7,8)→(1,2,3,4,5,6,7,8)? **~ 6 dB !**



Minimum RMSE case

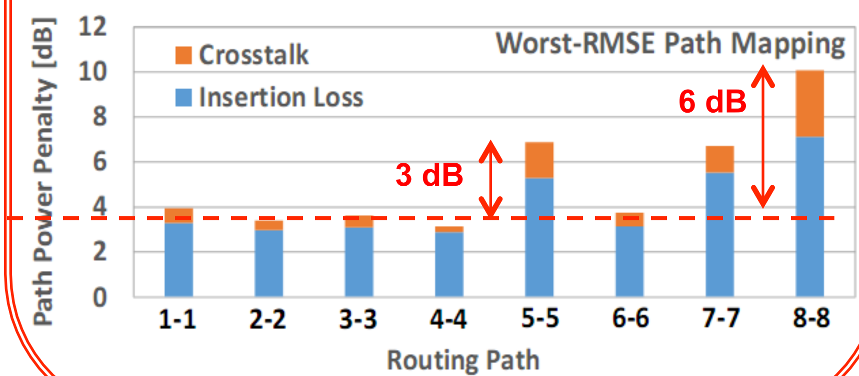
CCCC|CCCC|BBBB|CCCC|CCCC





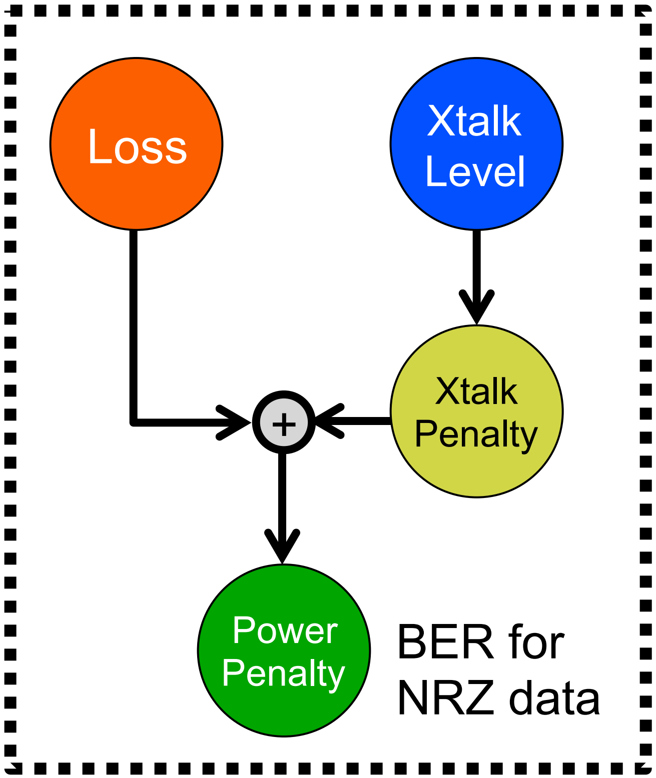
Maximum RMSE case

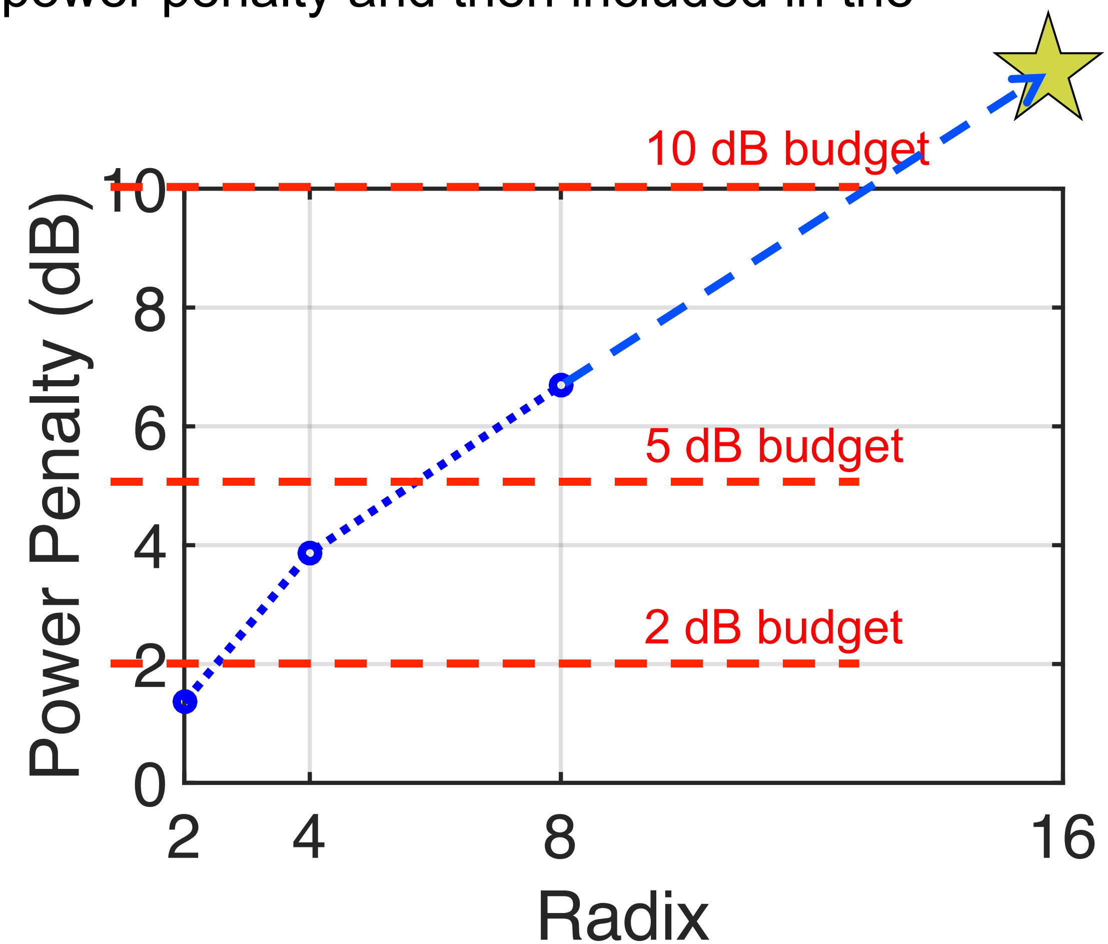
CCCB|CCCB|BBBB|CCCB|CCCB



Scalable Switches: Limitation of Power Budget

- **<u>Insertion Loss</u>** is directly included in the power budget
- Crosstalk is translated into power penalty and then included in the power budget

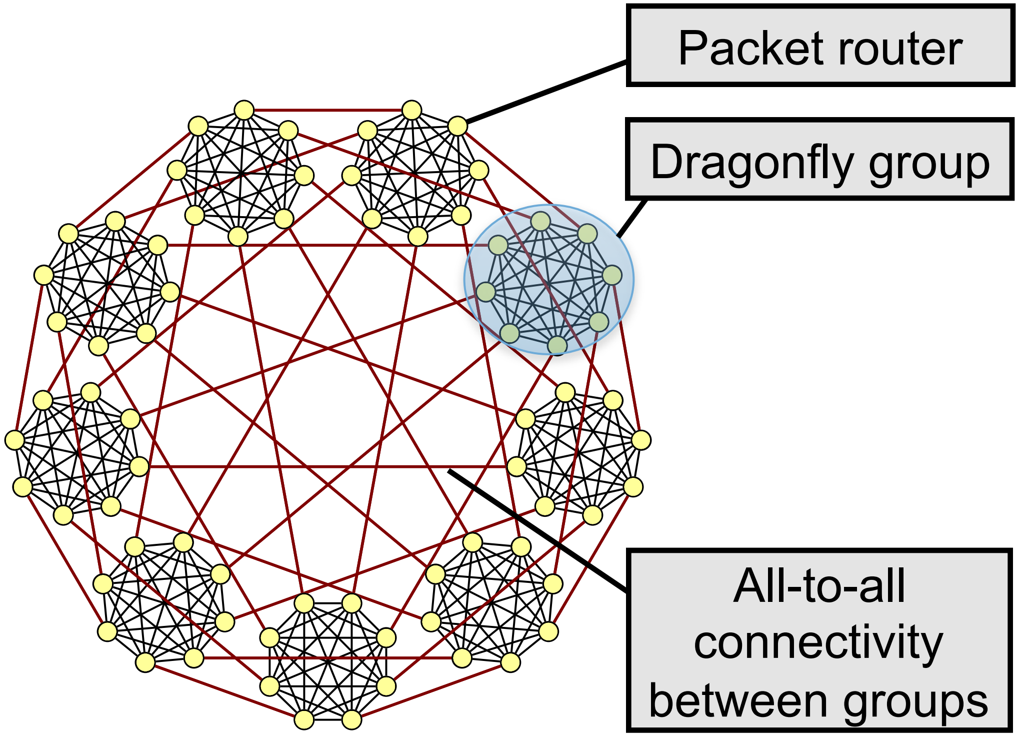




Introducing Flexfly

- **Flexfly: bandwidth steering in Dragonfly topologies with low cost, low-radix Silicon Photonics switches**

- Dragonfly topology:
  - Inside groups: all-to-all
    - short distance electrical links
  - All-to-all between the groups
    - Long distance optical links
  - Advantages:
    - Scalability
    - Low-diameter
  - Disadvantage:
    - Diluted inter-group bandwidth (to save on optical links)



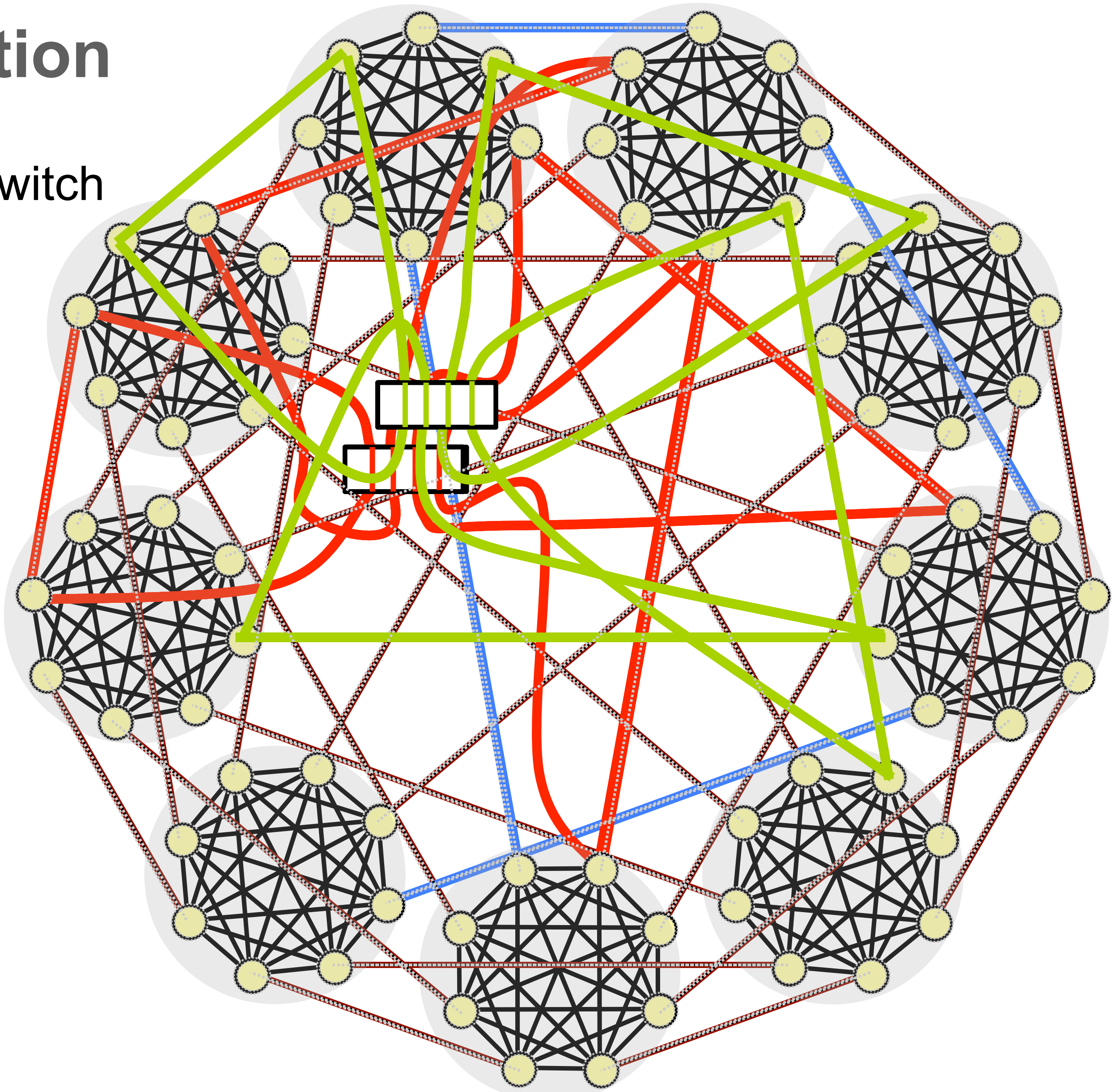
[1] K. Wen, et al. "Flexfly: Enabling a Reconfigurable Dragonfly through Silicon Photonics", SuperComputing, Best Student Paper award, 2016

Flexfly construction

- *R*: radix of the optical switch

1) Consider group-to-group links only
2) Create sets of *R* optical links
3) Assign a switch to each set of links

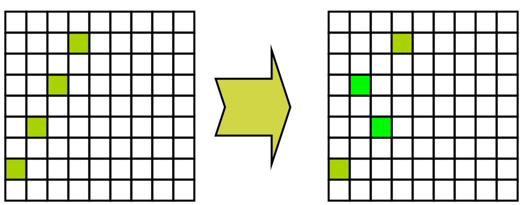
The above happen at machine design time



Flexfly construction

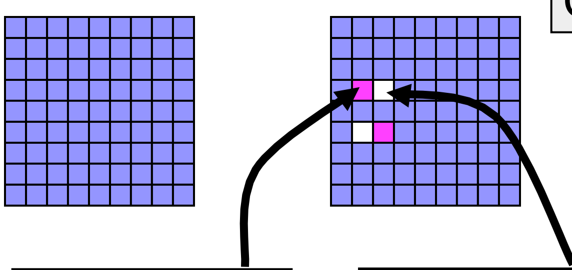
Before running parallel application:

4) Configure each switch to steer bandwidth



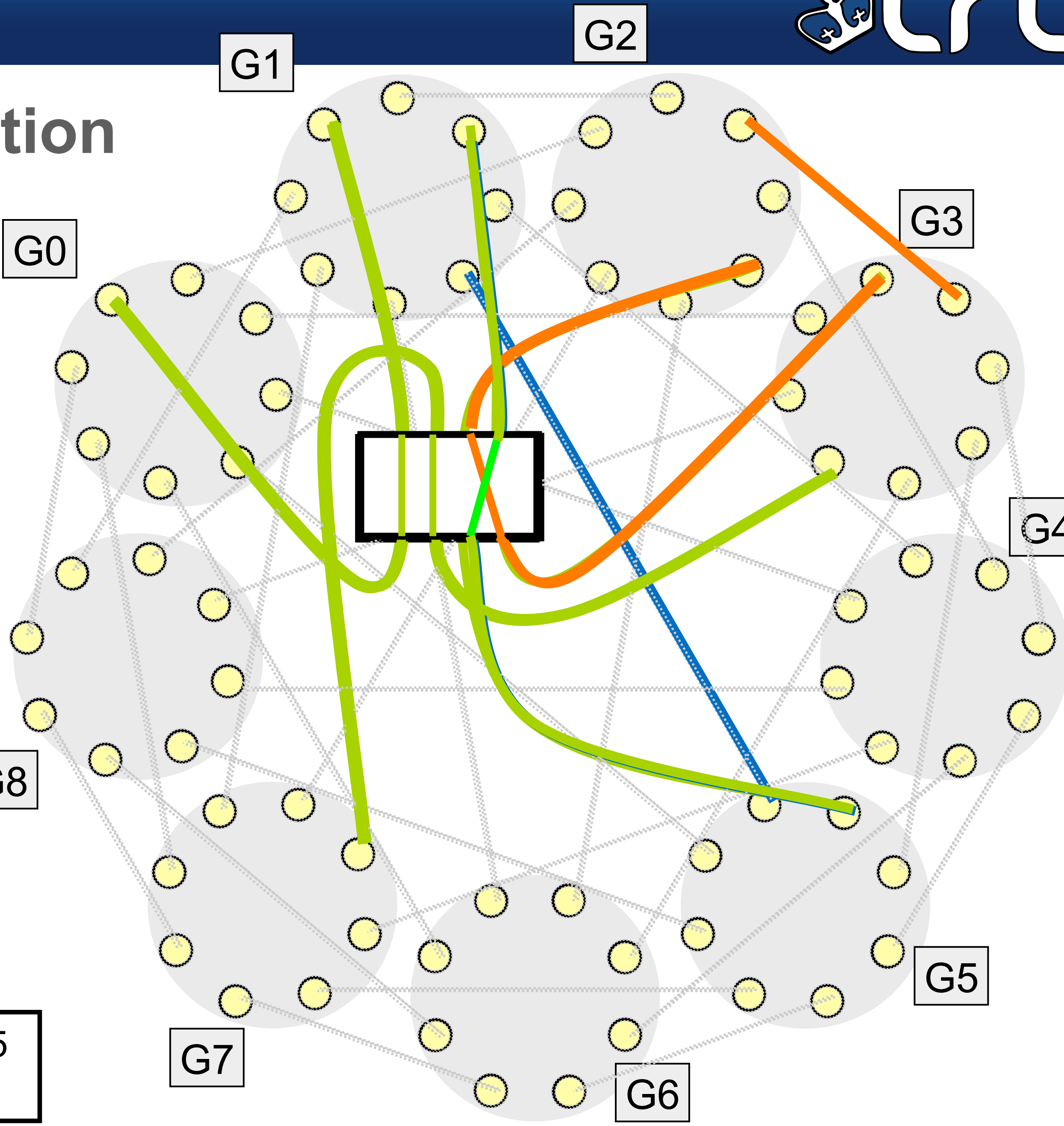
Original group-to group bandwidth distribution

New bandwidth distribution

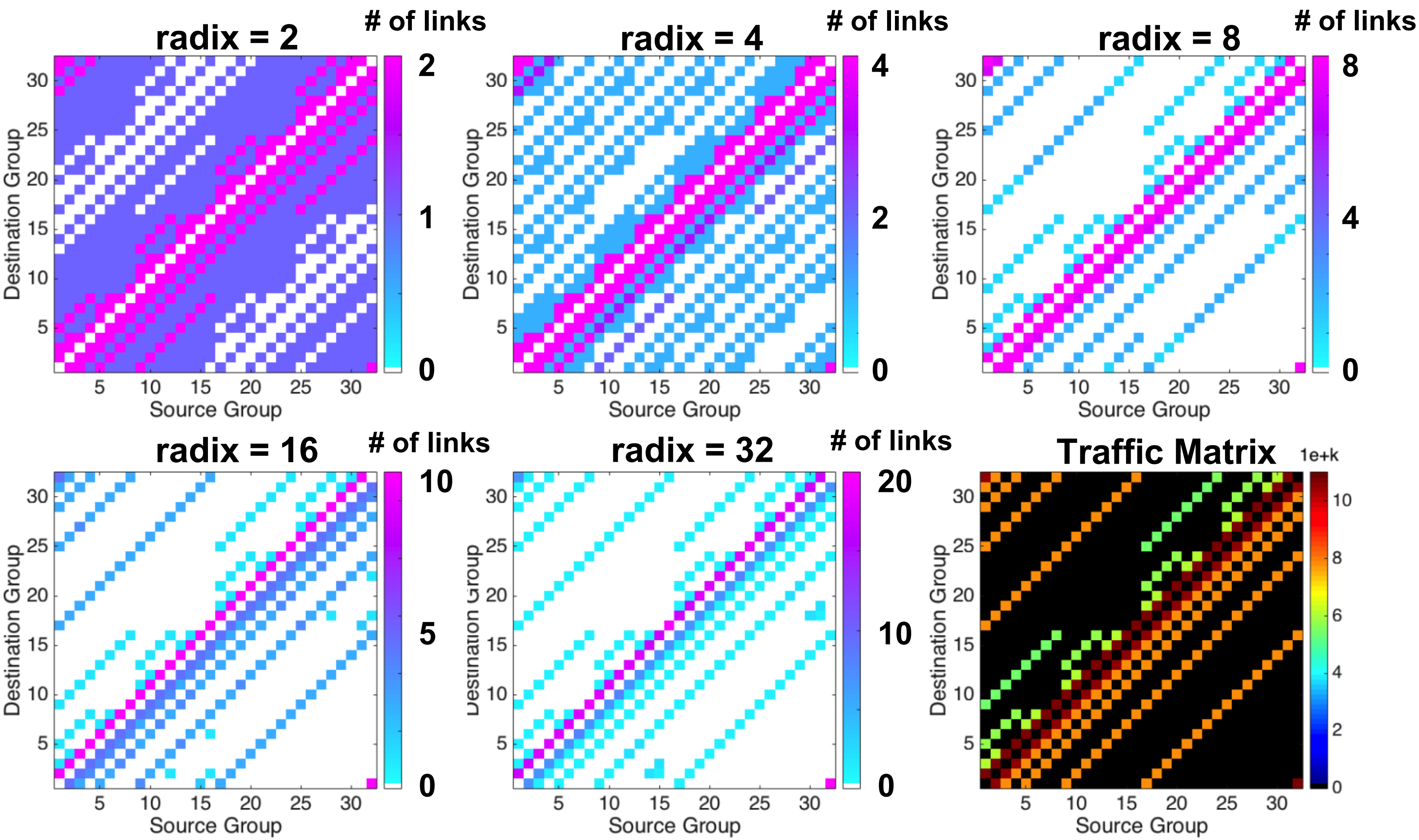


1 – 5 enjoys doubled BW

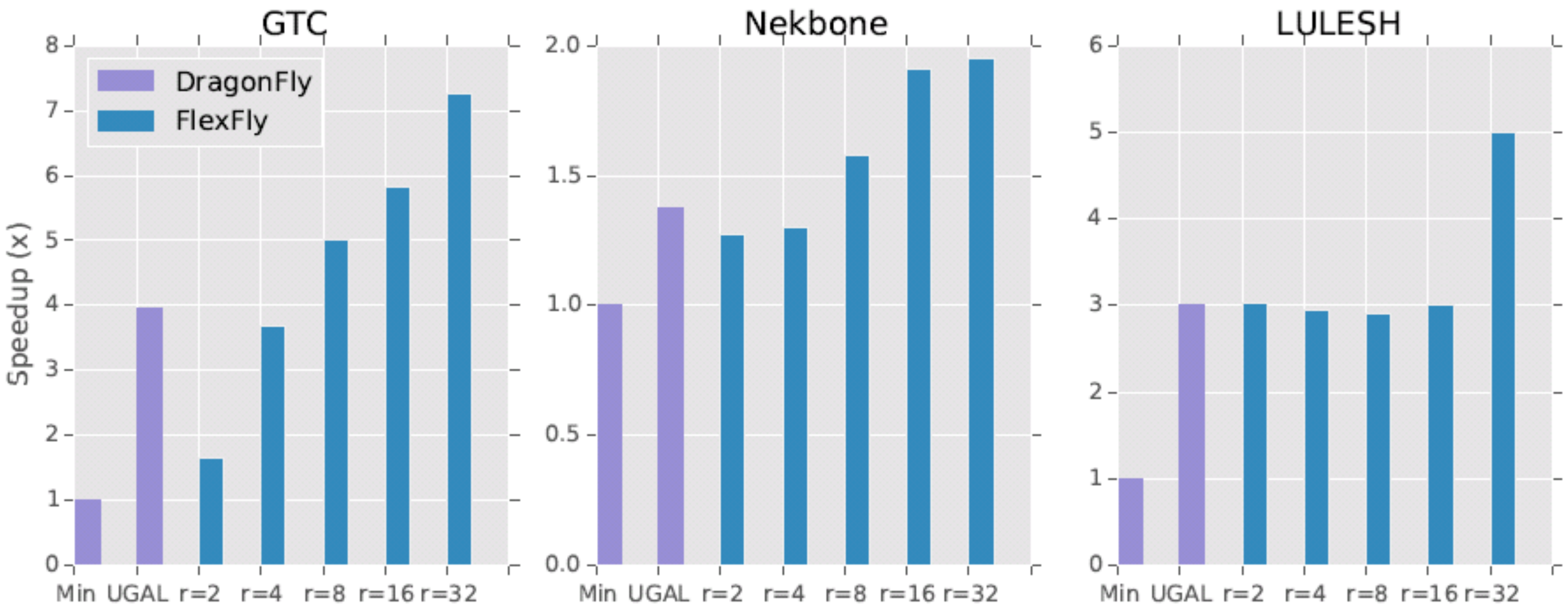
BW of 2 – 5 “stolen”



Adapting topology for GTC application



Flexfly – simulated performance



Thank you!

Columbia University

LIGHTWAVE RESEARCH LABORATORY

arpa·e
CHANGING WHAT'S POSSIBLE

Sandia National Laboratories

U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

SRC® Semiconductor Research Corporation

Hewlett Packard Enterprise

CISCO

ROCKPORT

AIM photonics
AMERICAN INSTITUTE for MANUFACTURING INTEGRATED PHOTONICS

NASA

LPS
The Laboratory for Physical Sciences