Requirements and novel devices for optical interconnects

David Miller

Stanford University
http://ee.stanford.edu/~dabm

Acknowledgements

Collaborators

- Stanford students
  - Ross Audet, Krishna Balram, Stephanie Claussen, Elizabeth Edwards, Onur Fidaner, Y. Ge, Ekin Kocabas, Yu-Hsuan Kuo, Salman Latif, Dany-Sebastian Ly-Gagnon, Jon Roth, Rebecca Schaevitz, Shen Ren, Yiwen Rong, Emel Tasyurek, Pierre Wahl, Justin White
- Stanford postdoc
  - Georgios Veronis
- Stanford faculty
  - Mark Brongersma, Shanhui Fan, Jim Harris, Krishna Saraswat
- Others
  - Ted Kamins, Takuo Tanemura

Funding

- FCRP/DARPA Interconnect Focus Center
- Intel
- DARPA EPIC, UNIC, and Optocenters programs
- AFOSR Plasmonics and Robust and Complex On-Chip Nanophotonics MURIs

For a copy of these slides, send e-mail to dabm@ee.stanford.edu

Main reference D. A. B. Miller, Proc. IEEE 97, 1166 - 1185 (2009)

Summary

Reasons why we have to use optical interconnects

- Energy
- Density

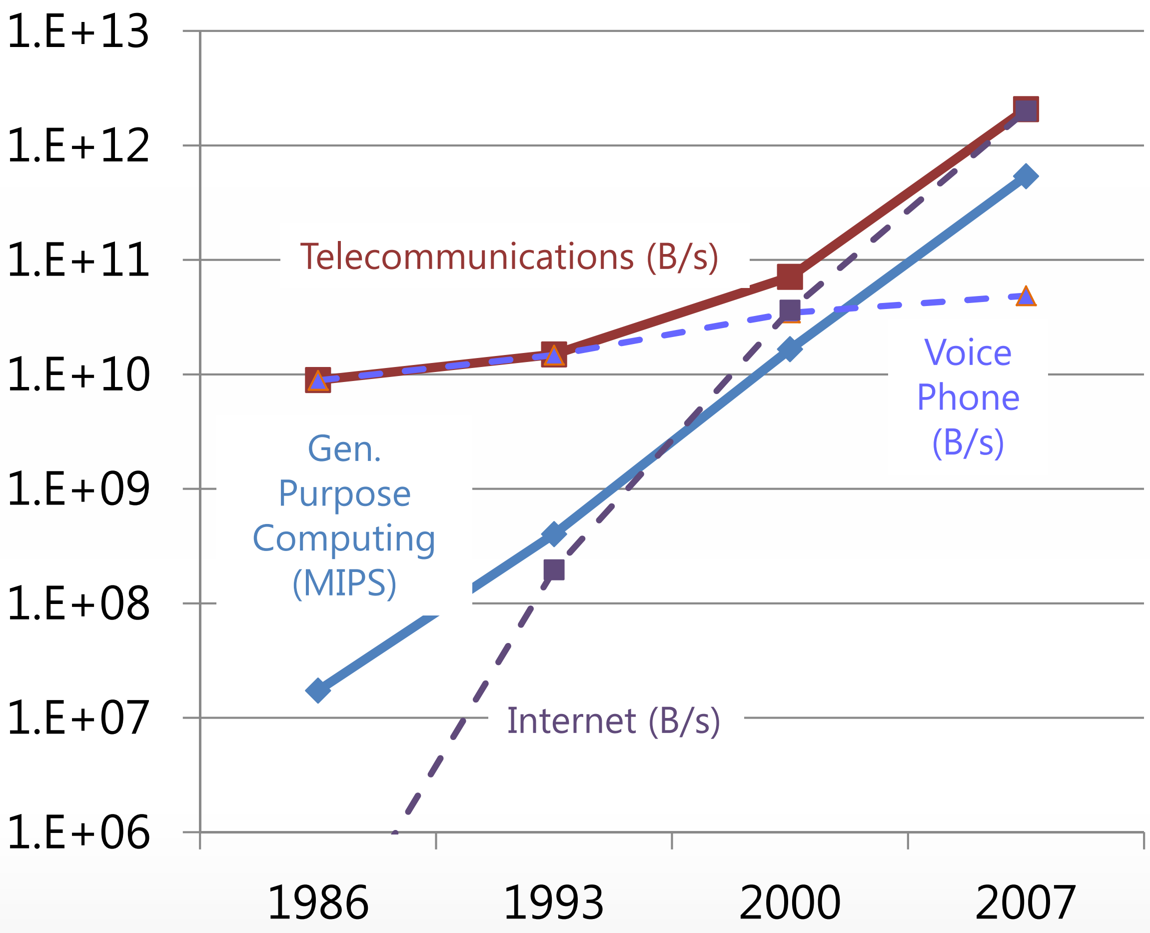
How are we going to do that?

Where are the opportunities?

For a copy of these slides, send e-mail to dabm@ee.stanford.edu

Main reference D. A. B. Miller, Proc. IEEE 97, 1166 - 1185 (2009)

Growth in information communication and processing



M. Hilbert and P. Lopez, "The World's Technological Capacity to Store, Communicate, and Compute Information," Science 332, 60-65 (2011)

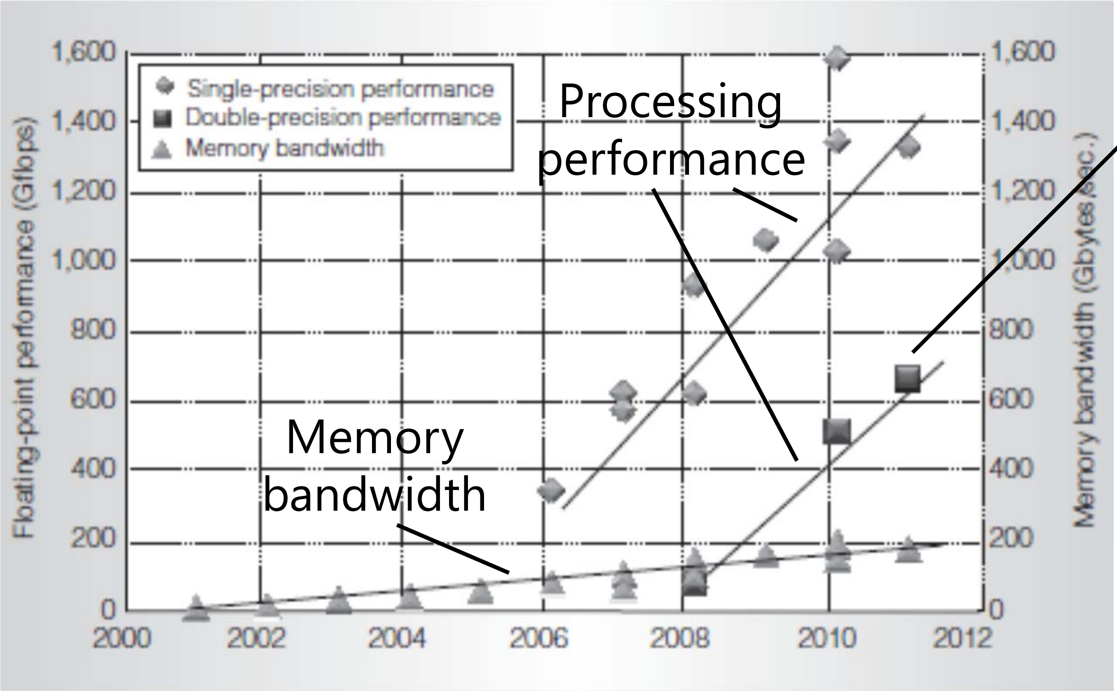
MIPS – million instructions per second ~ 3 - 6 instructions = 1 floating point operation (FLOP)

Both
- Internet traffic
- General purpose computing hardware

growing ~ 60 % per year
- ~ X 100 in 10 years

Hardware scaling challenge
- Energy
  - Energy per bit has to reduce
    - In communications and logic
    - Energy scaling not environmentally sustainable
      - Information technology already (2006) consumed ~ 1.5% of US electricity
- Communication
  - Communication density inside systems already at limits for electrical approaches

Memory bandwidth and floating point performance of graphics processor chips



Processing performance

Memory bandwidth

Current graphics processing chip (Nvidia *Fermi*)

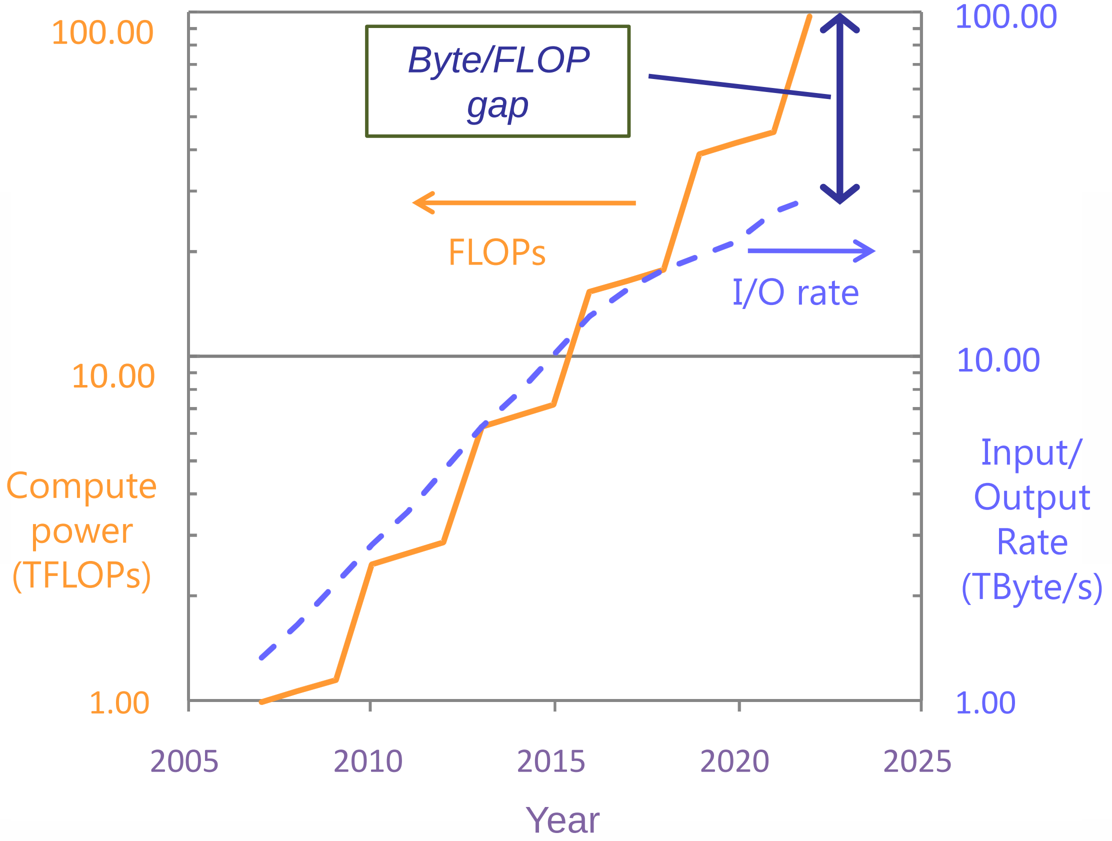
- 1.6 GHz internal clock rate
- 130 W
- 665 Gflops peak (double precision)
- 175 GB/s peak off-chip bandwidth
  - 1.4 Tb/s
  - Note that total 2010 internet traffic ~ 60 Tb/s

S. W. Keckler, W. J. Dally, B. Khailany, M. Garland, and D. Glasco, "GPUs and the Future of Parallel Computing," IEEE Micro 31, No.5, 7 – 17 (Sept./Oct. 2011)

Of the 225 pJ required for an operation

- 50 pJ is for logic and arithmetic operations
- Most of the rest is for interconnect to move the data on and off chip
- Accessing this data from external DRAM memory takes more than 10 nJ

Projected chip performance – bytes/FLOP



Compute power in floating point operations per second (FLOPs)

Scaled from 2007 chip

Input/Output rate from ITRS (International Technology Roadmap for Semiconductors) (scaling number)

(# Signal pins) x (off-chip clock rate)

"Device Requirements for Optical Interconnects to Silicon Chips," Proc. IEEE 97, 1166 - 1185 (2009)

Input/Output interconnect (I/O) rate does not keep up with ability of chip to calculate

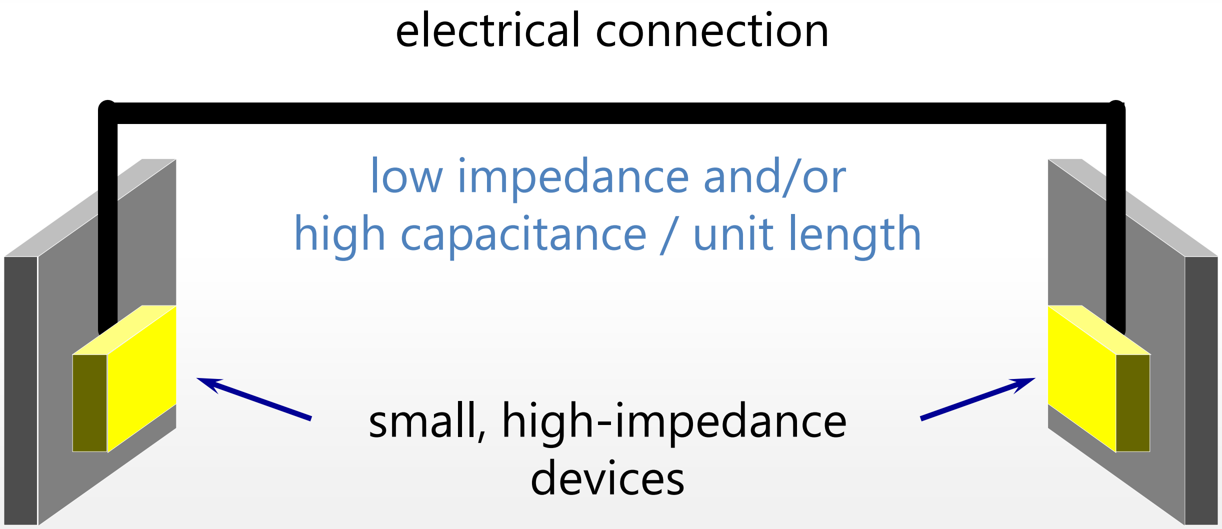
Ideal of 1 Byte of memory access for each floating point operation (FLOP) cannot be retained

- Byte/FLOP gap

Power dissipation in electrical interconnects

Dissipation in electrical interconnects is from charging and discharging wire capacitance

- Wires always have large capacitance per unit length
  - ~ 2 pF/cm
- Simple logic-level signaling results in specific dissipation
  - E.g., at 2pF/cm and a 2 cm chip, at 1 V on-off signaling
    - Dissipate at least ~ $\frac{1}{2}CV^2$ per bit sent across chip ~ 2pJ



Saving energy with optical interconnects – “quantum impedance conversion”

Suppose 1 nW optical beam

- classical voltage (e.g., from side to side of linearly polarized light beam)
  - 600 microvolts in 377 ohm impedance

Presume 1 eV photons

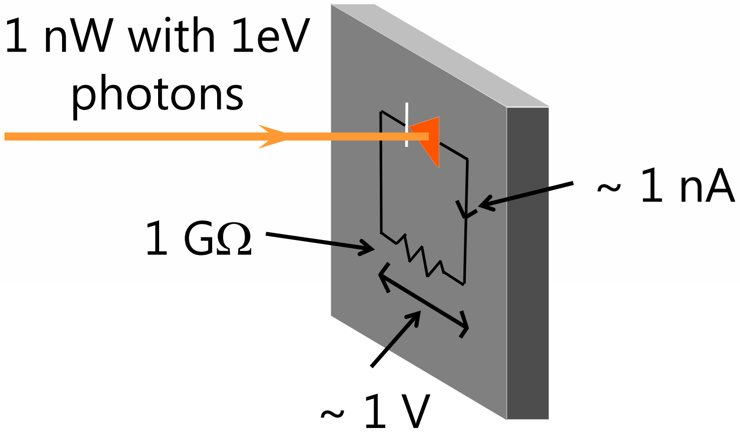
Presume simple photodiode operating with a 1 gigaohm load

- Then 1 nW of light can generate about 1 nA of current,
  - 1 V in 1 gigaohm load

Photodiode has performed “quantum impedance transformation”

- Consequence of photoelectric effect

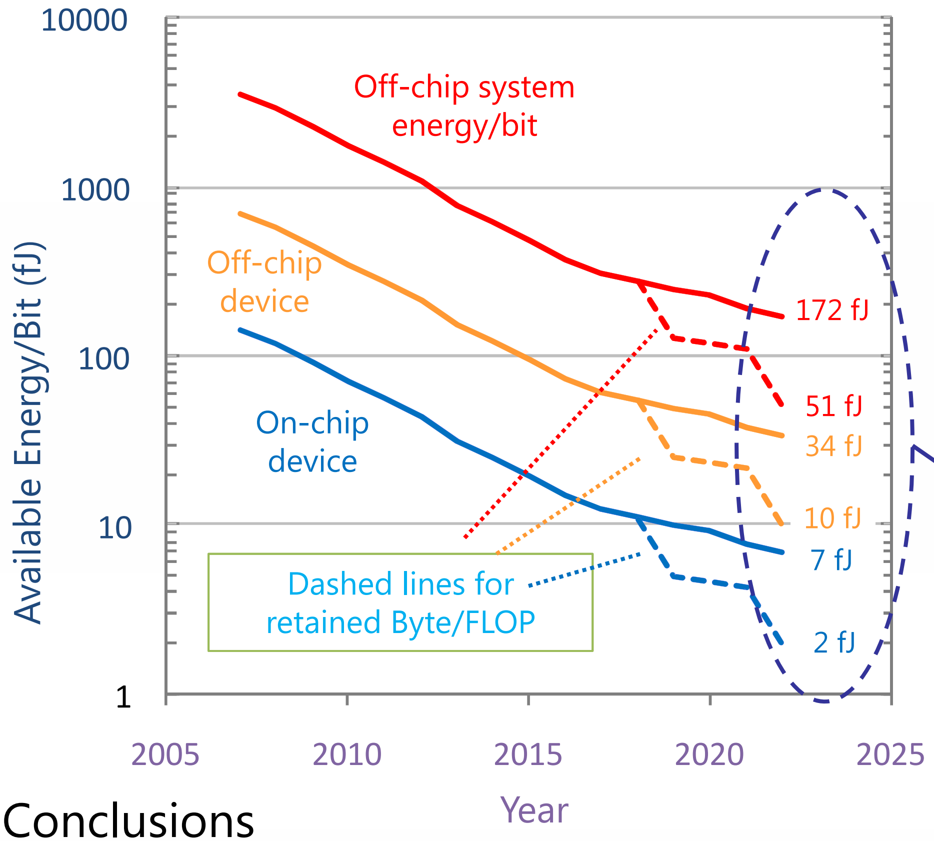
$$\text{Power} = \frac{\left(\text{Voltage}_{RMS}\right)^2}{\text{Impedance}}$$



Optics Letters 14, 146-148, (1989)

Optics only has to charge the photodetector and transistor to the logic voltage, not the interconnect line

Available energy for interconnects



Total chip power limited to ~ 200 W

Limits total system energy per bit

Dashed lines presume we retain Byte/FLOP ratio in the later years

Allowable total energy per bit communicated

based on 20% of power for each of off-chip and on-chip interconnect, and device energy being 20% of total energy/bit

Conclusions

- 50 – 100 fJ/bit system energies are low enough for off-chip interconnect
- Optical output devices need to have energies of ~ 10 fJ for desired I/O in later years
- Even lower energies desirable for on-chip interconnect

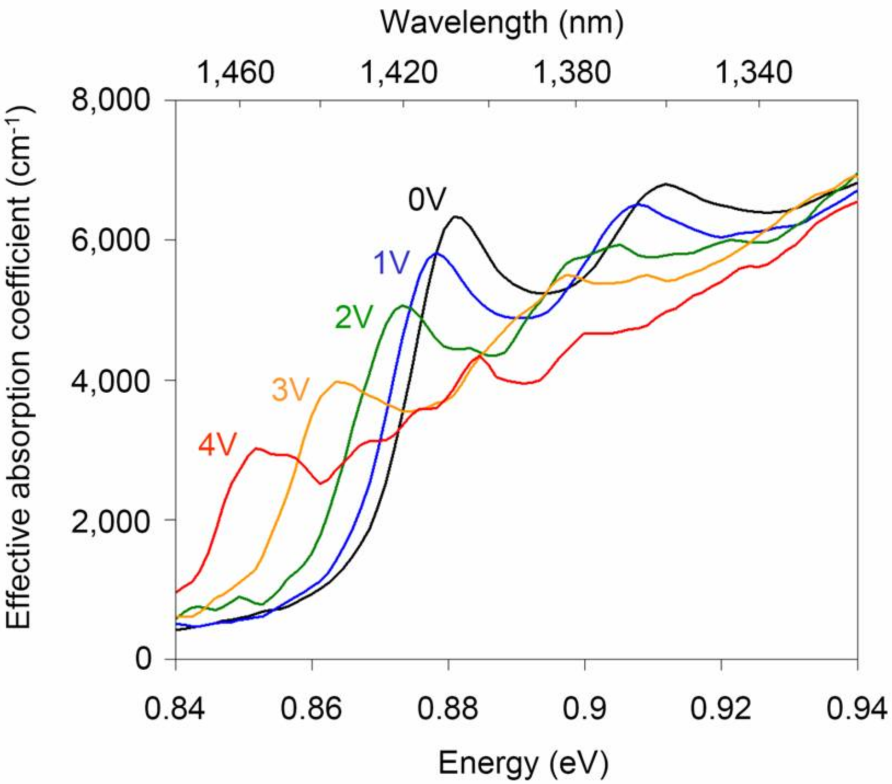
Approaches to reduce device energy

- Integrate optics and optoelectronics compatible with silicon CMOS
  - Germanium quantum well modulators
    - Very low energy devices potentially integrable with CMOS

- Exploit nanophotonics for compact very low energy devices
  - Nanophotonic resonators
  - Nanometallic enhanced photodetectors
    - Potentially very low capacitance for low total energy of communication

QCSE in germanium quantum wells on Si

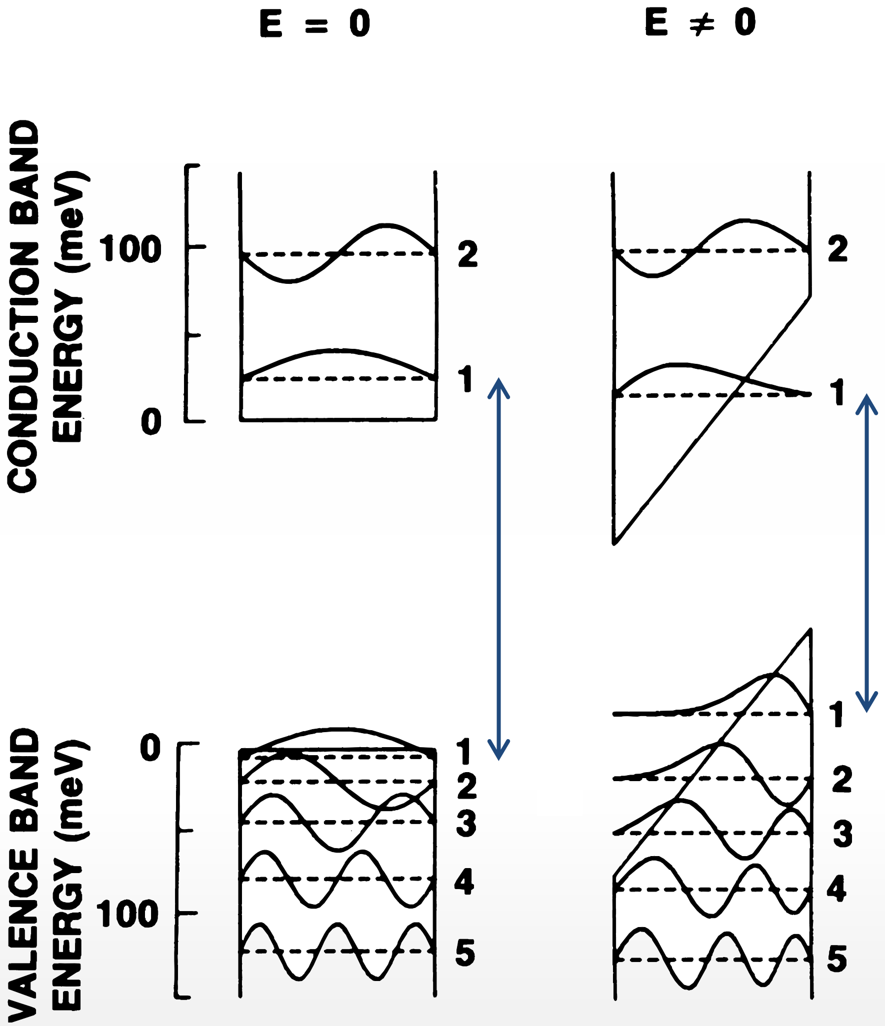
- Quantum-confined Stark effect (QCSE)
  - Strongest high-speed optical modulation mechanism
  - Used today for high-speed, low power telecommunications with III-Vs
- QCSE in germanium quantum wells on silicon substrates
  - compatible with CMOS fabrication
  - Can work over C-band at 1.55 μm
- Surprises
  - Works very well in "indirect gap" semiconductor
    - Uses Ge direct gap absorption



Y.-H. Kuo, Y.-K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. B. Miller & J. S. Harris, Nature 437, 1334-1336 (2005)

Funded by DARPA EPIC Program, Intel, MARCO/DARPA Interconnect Focused Research Center
J. Harris and D. Miller groups, Stanford University

Quantum-confined Stark effect



Electron and hole energy levels shift with field

- Causes, e.g., reduction in lowest (1 – 1) transition energy

See

“Quantum Mechanics for Scientists and Engineers,” (Cambridge, 2008)

Also upcoming

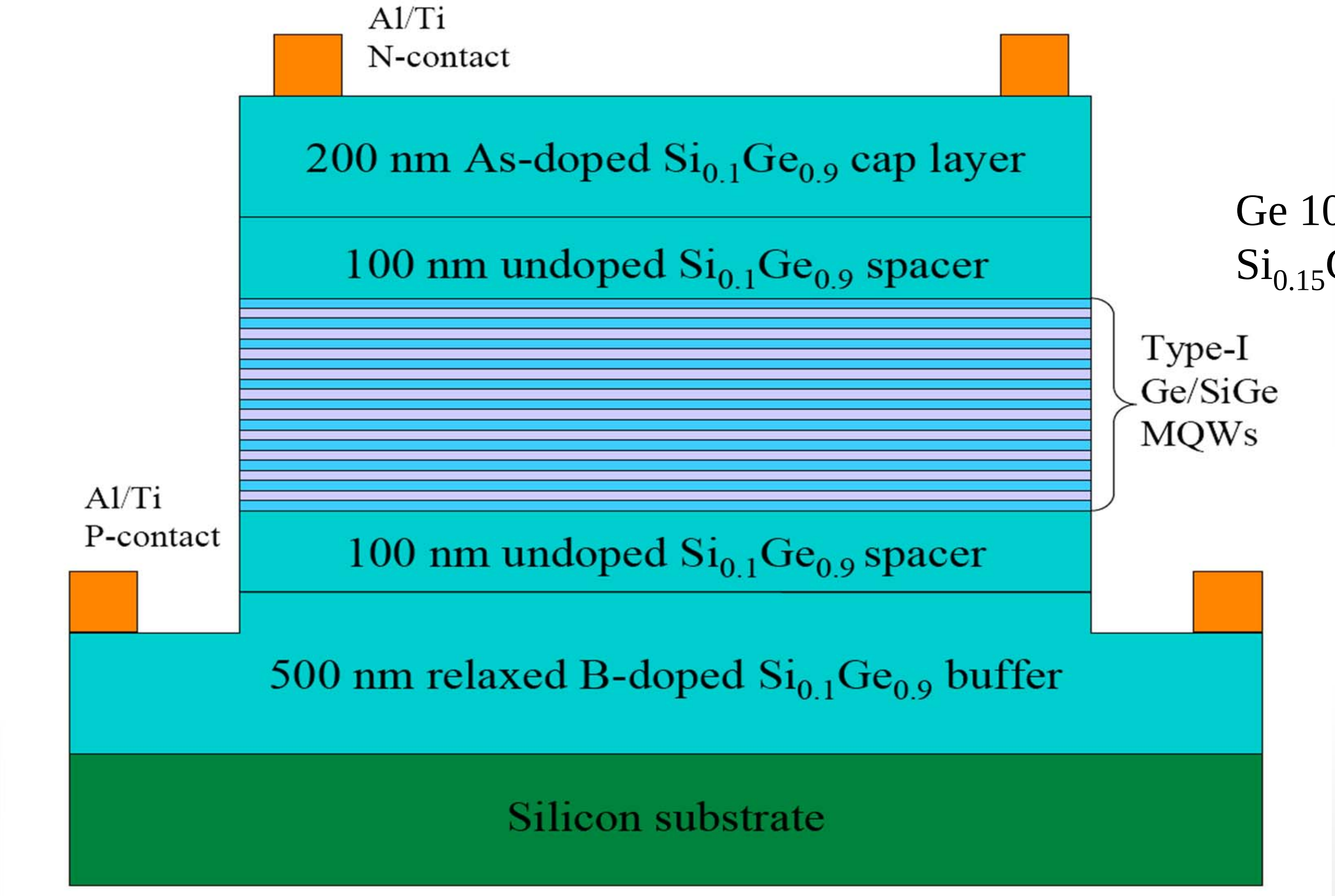
“Quantum Mechanics for Scientists and Engineers”

Massive open on-line course (MOOC)

See http://class.stanford.edu

Starting September 24, 2013

Ge quantum well modulator on Si

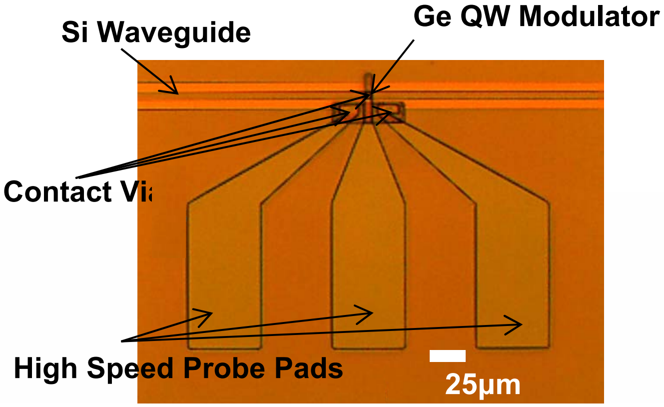


Ge 10nm/
$Si_{0.15}Ge_{0.85}$ 16nm

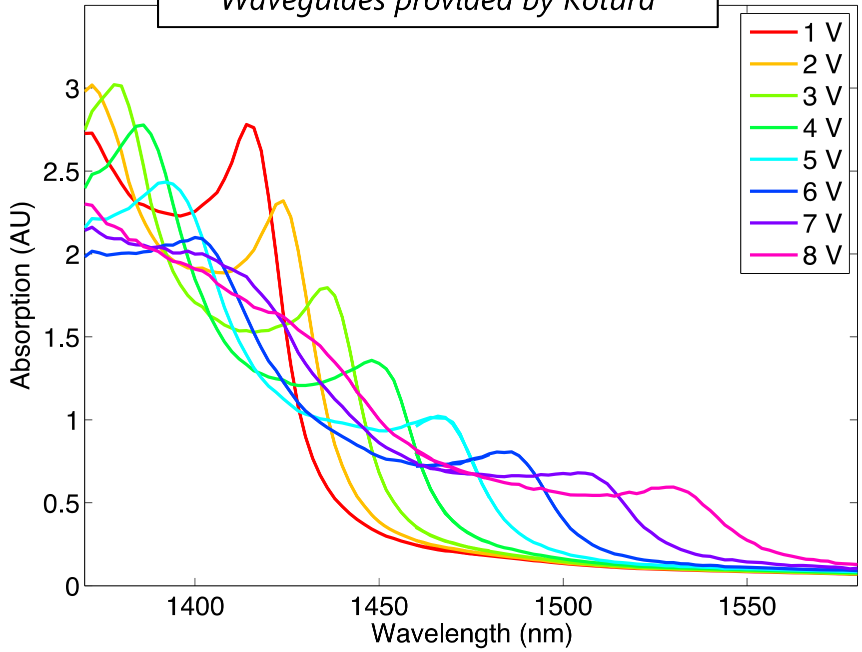
Ge quantum well waveguide-integrated modulator

- 10 microns long, 0.8 microns wide, 500 nm thick intrinsic region
  - No resonator
- Selective area growth of quantum wells in silicon on insulator (SOI) waveguides
  - voltage tunable device, even to 1550 nm
  - Capacitance ~ 3 fF
- 3 dB modulation with 4 V bias, 1 V swing, 1460 nm
  - Dynamic energy per bit ~ ***0.75 fJ***
- Tested to 7Gb/s (equipment limited)

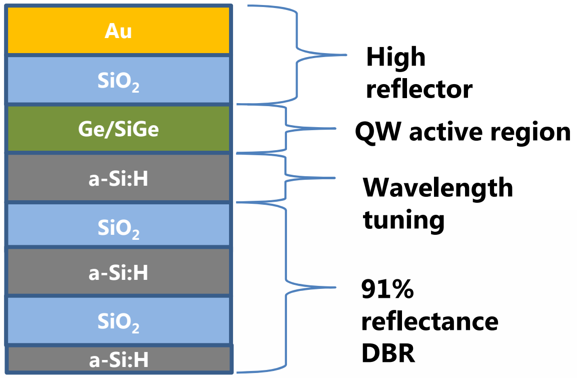
S. Ren et al., IEEE PTL 24, 461 – 463 (2012)
D. A. B. Miller, Optics Express 20, A293-A308 (2012)

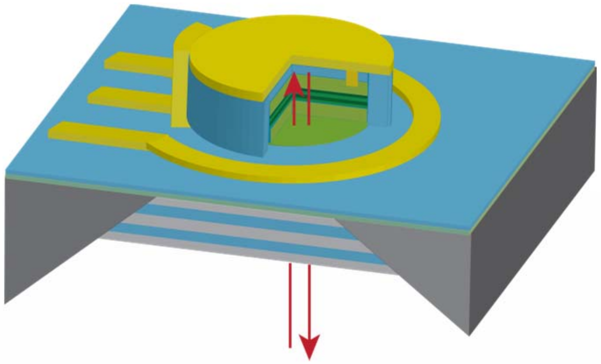


Waveguides provided by Kotura

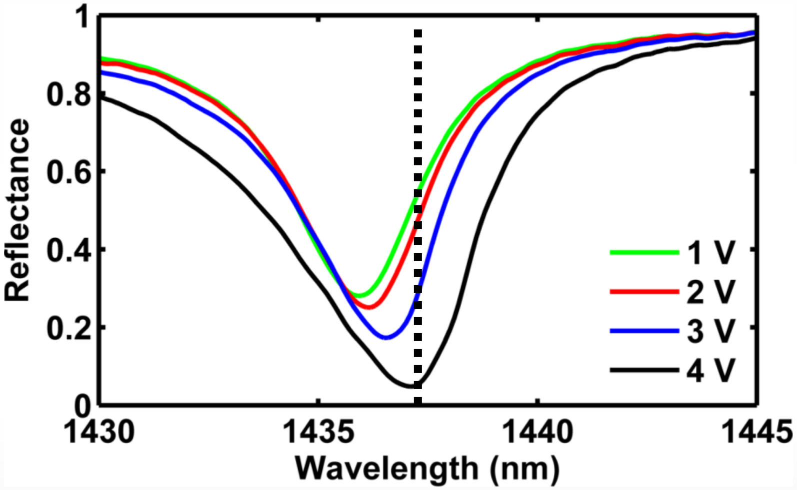


Ge quantum well reflection modulator on Si





Parameter	Value
Extinction ratio (2.5 V swing)	10.3 dB
Extinction ratio (1 V swing)	7.2 dB
Insertion loss (at 2V)	3.7dB
Resonator Q (low absorption state)	350

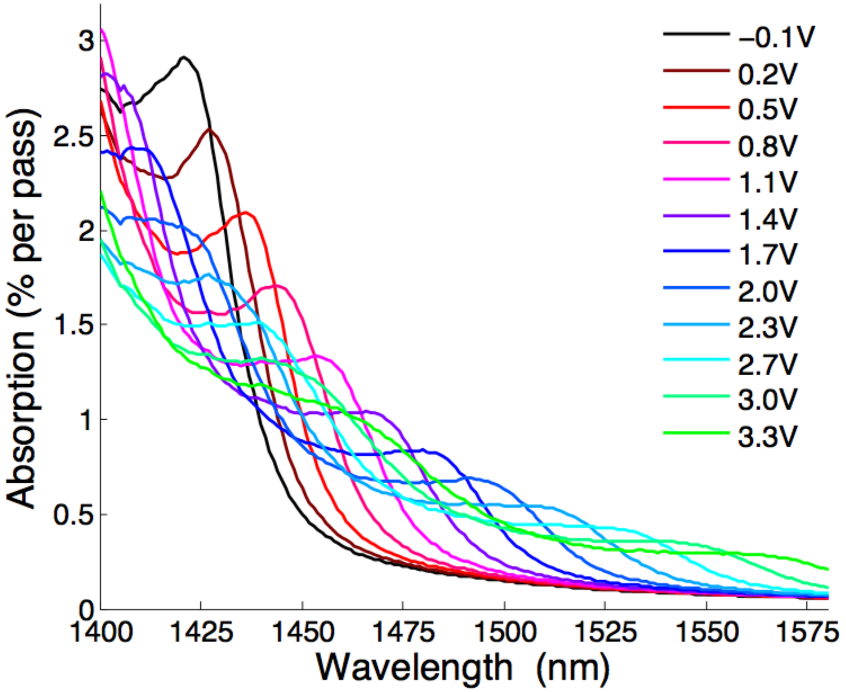


R. M. Audet, E. H. Edwards, K. C. Balram, S. A. Claussen, R. K. Schaevitz, E. Tasyurek, Y. Rong, E. I. Fei, T. I. Kamins, J. S. Harris, and D. A. B. Miller, "Surface-Normal Ge/SiGe Asymmetric Fabry-Perot Optical Modulators Fabricated on Silicon Substrates," J. Lightwave Technol. (to be published)

- QCSE shows strong modulation even for "surface-normal" operation
- Uses standard silicon substrate
- Asymmetric Fabry-Perot design
- Shows feasibility of surface-normal modulators compatible with silicon electronics
- Potential scalability to ~ 10 fJ/bit in micropillars

Low voltage modulator diode structure

Layer	
N-doped cap	$Si_{0.13}Ge_{0.87}$ As 6e18cm^{-3} 60nm
spacer	$Si_{0.13}Ge_{0.87}$ 60nm
barriers	$Si_{0.19}Ge_{0.81}$ 17nm
QW (x5)	100% Ge 12nm
spacer	$Si_{0.16}Ge_{0.84}$ 40nm
P-doped Buffer	$Si_{0.12}Ge_{0.88}$ Boron 4e19cm^{-3} 3x80nm
Si substrate	p+ doped



Electroabsorption contrast ratio >5dB over S- and C-bands with 1V drive swing

Virtual substrate (p-doped buffer) only 240 nm thick

Complete diode structure 560 nm thick

Low operating voltage minimizes photocurrent dissipation

E. H. Edwards, et al., Opt. Express **21**, 867 (2013)

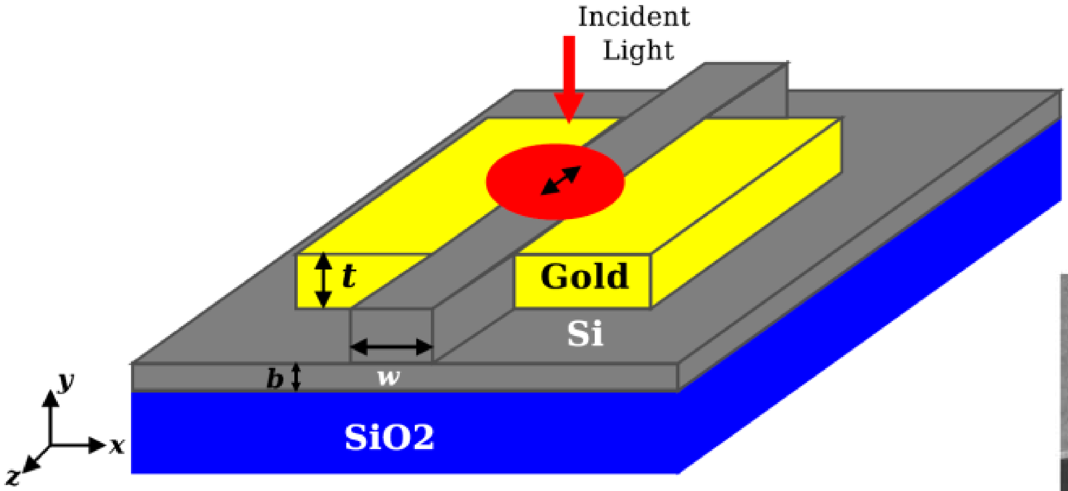
Hybrid Nanophotonic Photodetectors

Objective

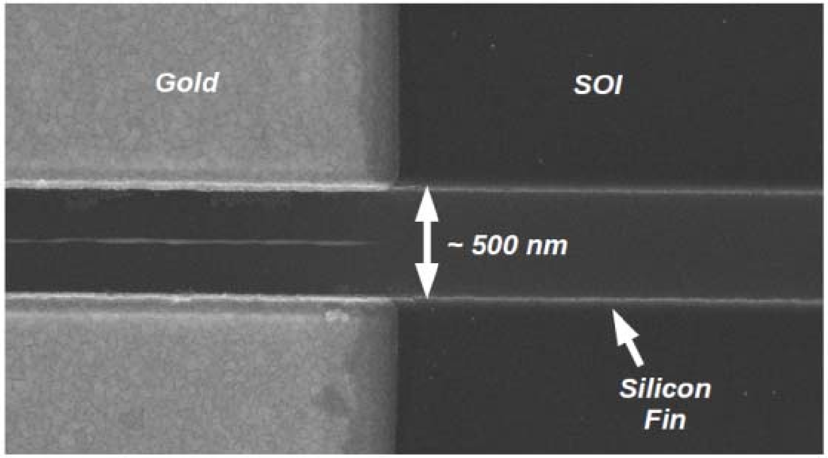
Demonstrate efficient photodetection into subwavelength structures compatible with CMOS processing techniques

Approach

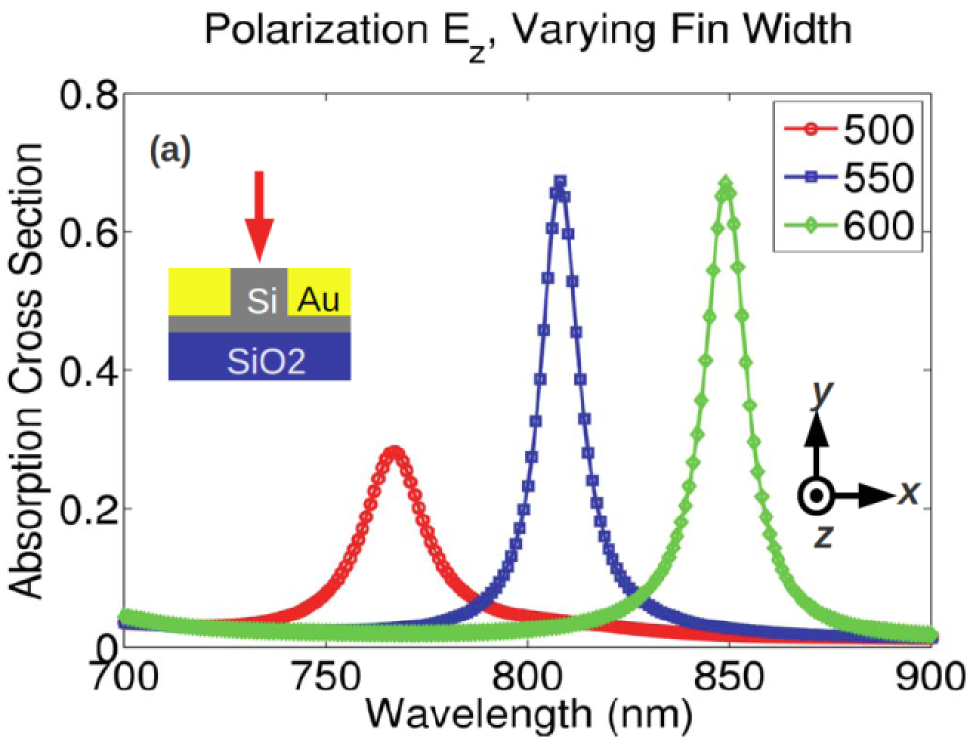
- Embed a silicon photodetector inside a nanometallic slit
- Use same metals for optical confinement and electrical connections
- Exploit resonances in the combined metal/dielectric resonance
  - enhance absorption and allow tunability of spectral response

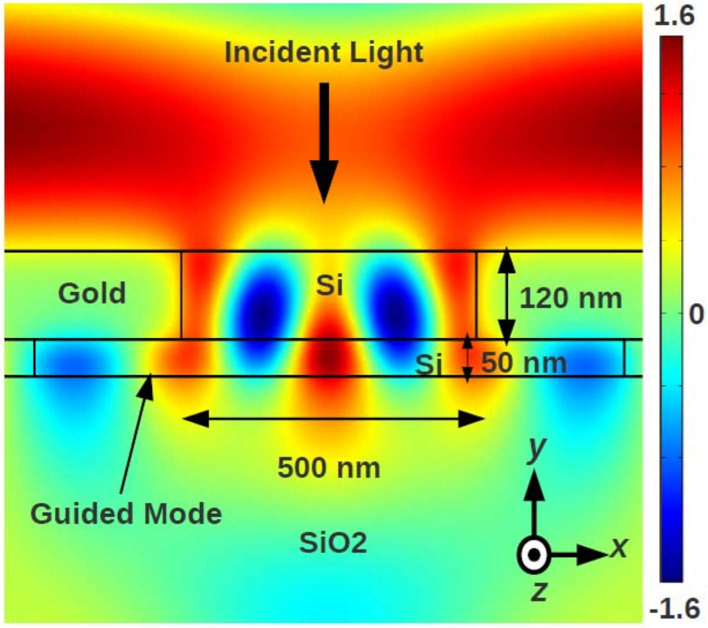


K. C. Balram and D. A. B. Miller, "Self-aligned silicon fins in metallic slits as a platform for planar wavelength-selective nanoscale resonant photodetectors," Opt. Express 20, 22735-22742 (2012)



Resonances in silicon fin in nanometallic slit - simulations

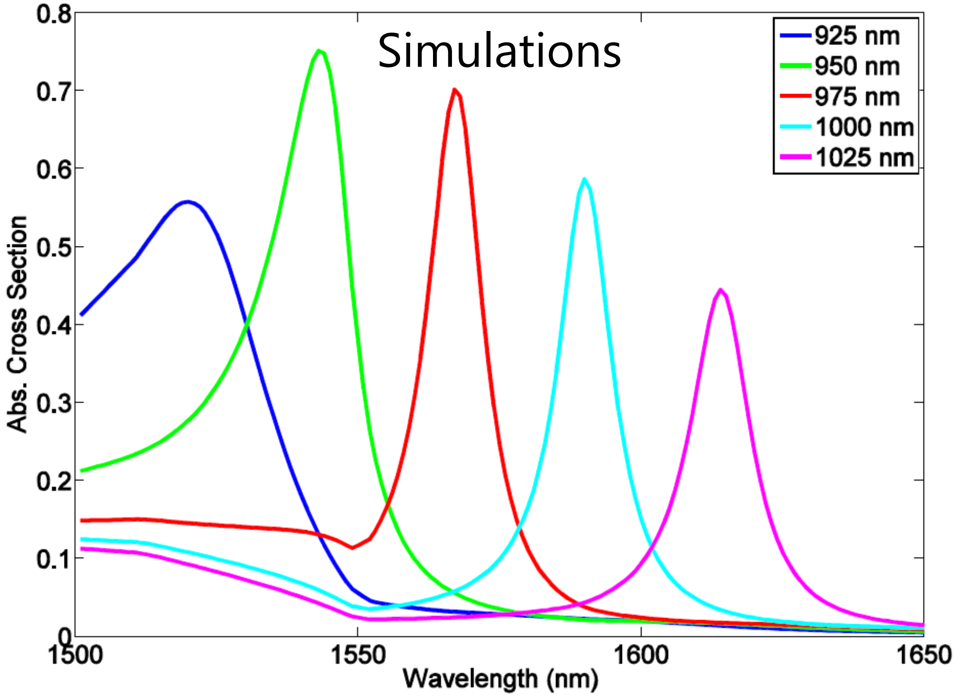


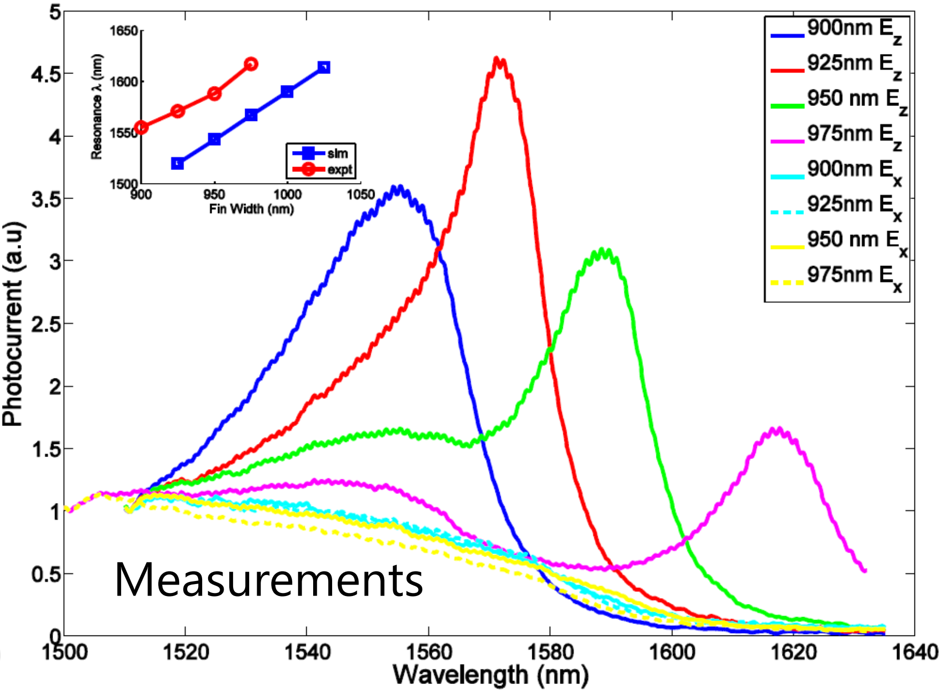


Shine light from the top on silicon fin

- Light polarized in z direction along the slit (out of the plane of the figure)
  - 5th order dielectric lateral Fabry-Perot resonance
    - Tunable by design of slit width
    - Very strong absorption
      - >60% of photons incident on slit absorbed in 170 nm thick Si at 850 nm

Resonances in germanium fin in nanometallic slit





Results

Good agreement between theory and experiment

Measured responsivity 1.2 A/W for 925 nm wide device in 280 nm thick

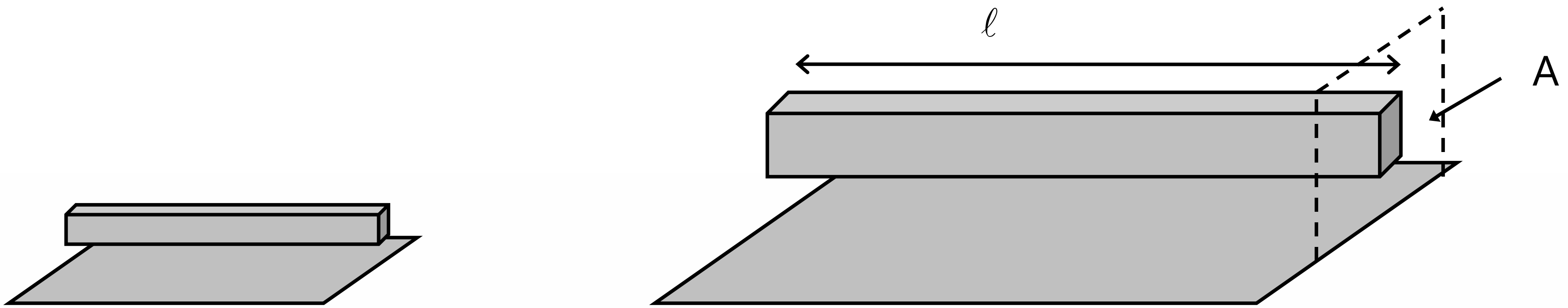
Combination of resonance and photoconductive gain

Key features

- Allows Ge detectors beyond 1550 nm
  Without strain or Sn incorporation
- Could permit coarse WDM splitting using different slit widths for different bands

K. C. Balram, R. M. Audet, and D. A. B. Miller,
Opt. Express **21**, 10228-10233 (2013)

Density problem in electrical interconnects



this wire

carries the same number of bits per second as

this wire

J. Parallel and Dist. Comp. 41, 4252 (1997)

Get universal form of scaling for simple digital connections (no repeaters, no multilevel modem techniques)

bit rate B $\propto$ A / ℓ^{2}

B ~ 10^{15} A / ℓ^{2} bits/s for LC lines

B ~ 10^{16} A / ℓ^{2} bits/s for RC lines

B ~ 10^{17} A / ℓ^{2} bits/s for equalized LC lines

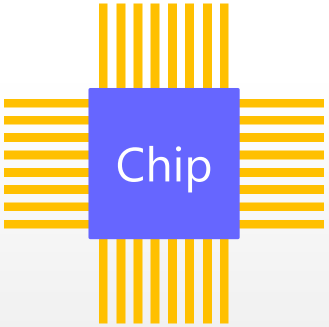
Once the wiring fills all space,

the capacity cannot be increased either by making the system smaller or making it larger

Optics completely avoids this scaling limitation

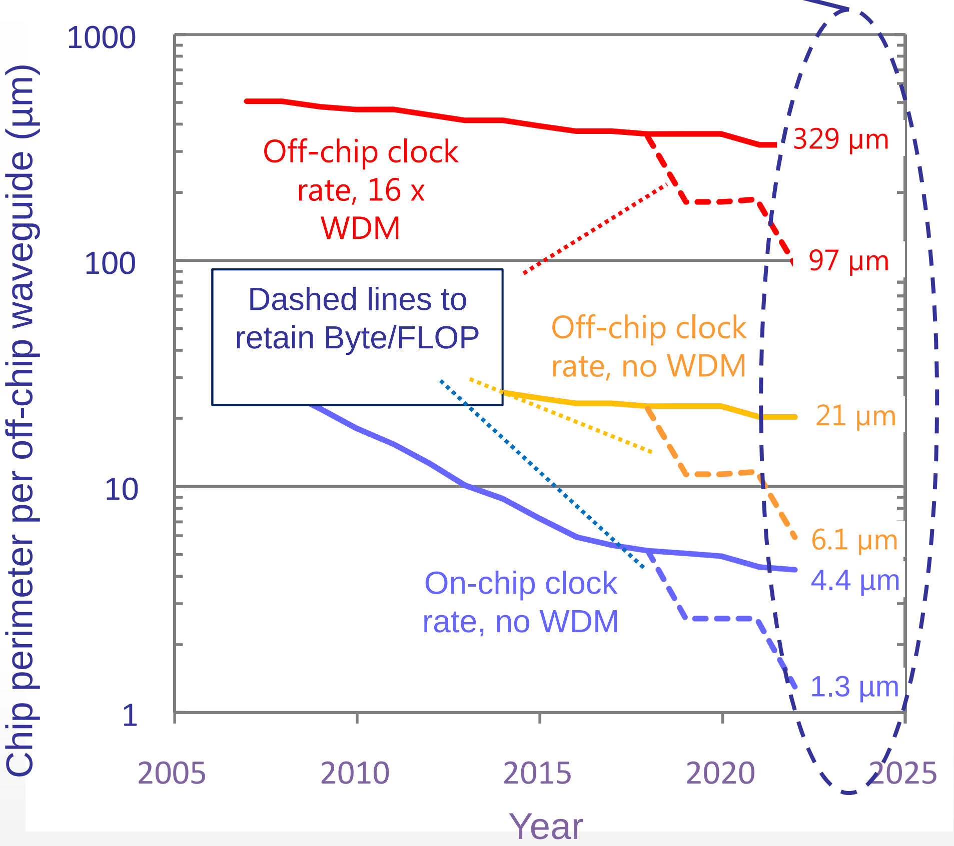
Waveguide optics - available width per waveguide

- Can we get enough optical channels off the edge of the chip in fibers or waveguides?
  - Yes, but for waveguides packed around chip perimeter we need
    - High clock speeds and wavelength division multiplexing (WDM) for optical fibers, or
    - dense waveguides or
    - spatial multiplexing



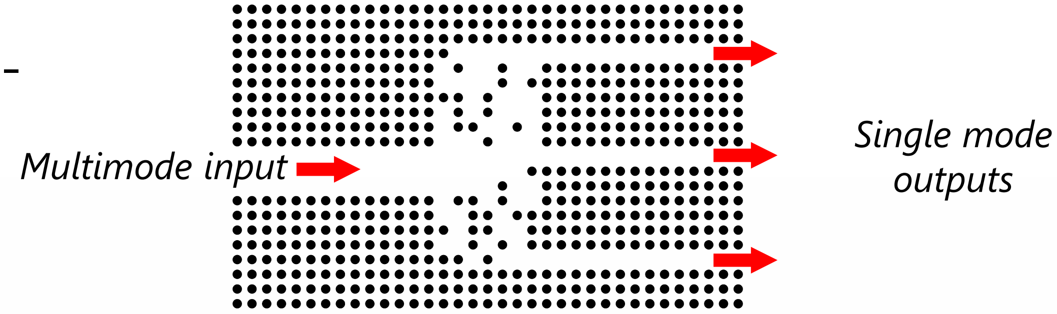
Waveguides or fibers to and from chip edge

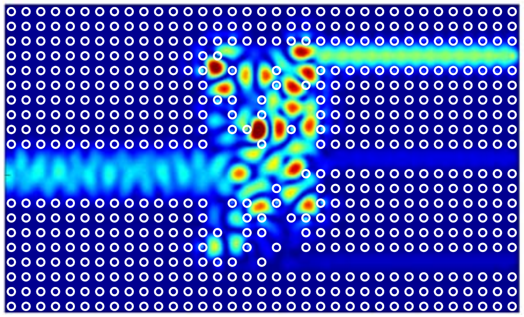
Available chip perimeter per fiber or waveguide

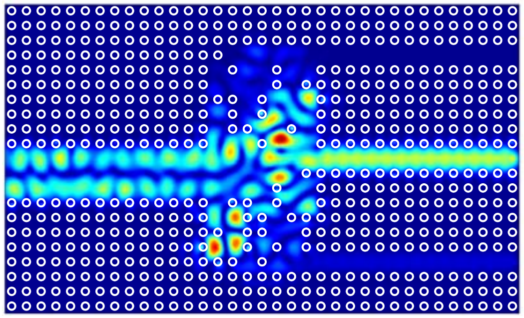


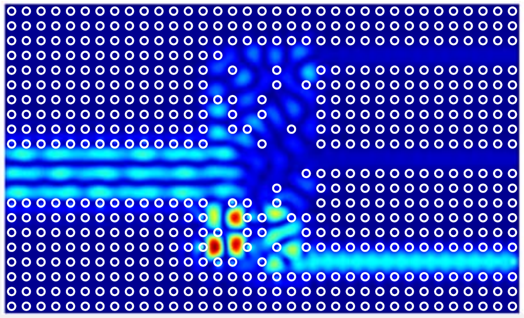
Novel optics for very large bandwidths – multiple spatial modes?

- Current interest in telecommunications for few-mode fibers, free-space quantum communications
- Need “loss-less” mode splitters
- But how to design?
- Iterative approach - randomly add and/or subtract cylinders
  - Successful after ~10000 steps (48 hrs on a Pentium III)
- We have no idea why it works!
  - Negligible intuition
- Can we design arbitrary linear optics without blind global iterations?









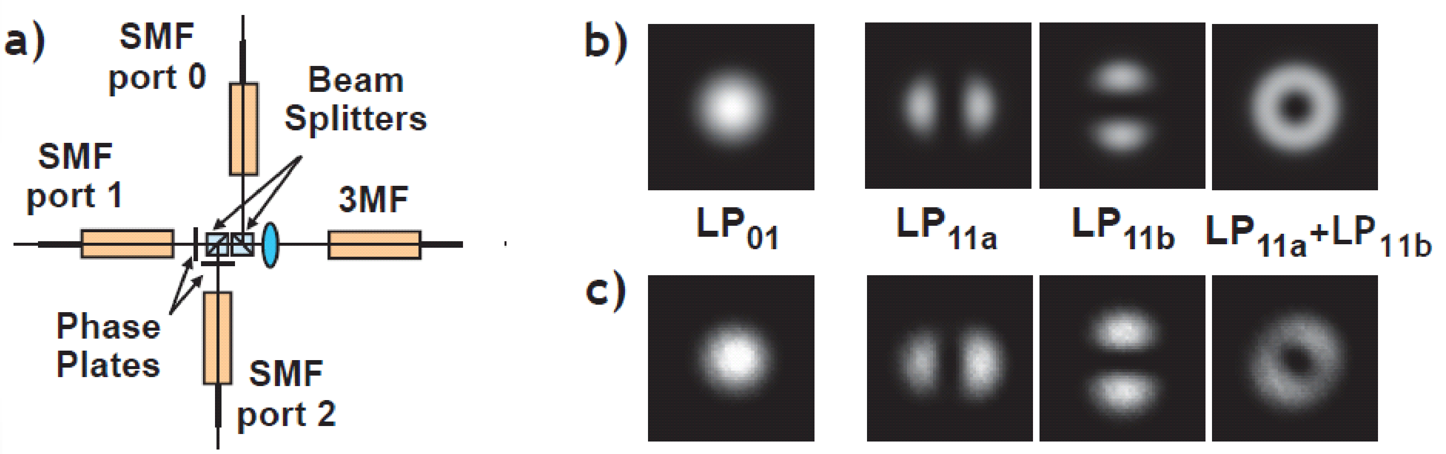
Engineer precise mode splitting with positioning of dielectric columns

Y. Jiao et al., Optics Lett. 30, 141-143 (2005)

How to split multiple modes efficiently?

We can combine three different overlapping modes into a multimode fiber (or separate them at the output)

- But this scheme has splitting loss
  - Power also is reflected or transmitted out by the beam splitters and/or dumped at the fiber inputs



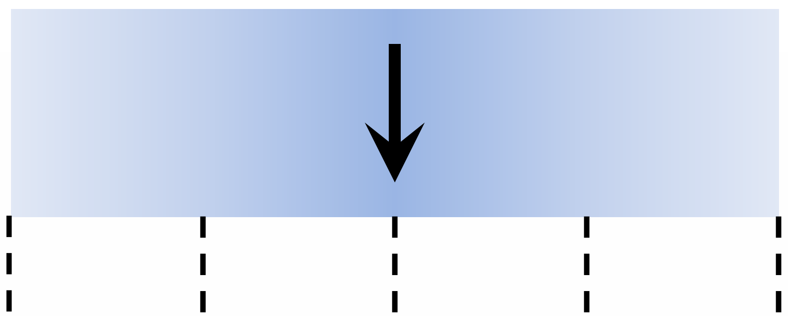
R. Ryf, C. A. Bolle, and J. von Hoyningen-Huene, ECOC 2011, paper Th.12.B.1.

How to design an arbitrary linear optical device

How to design an arbitrary linear optical device … and how to avoid it!

Coupling an arbitrary input beam

Input beam

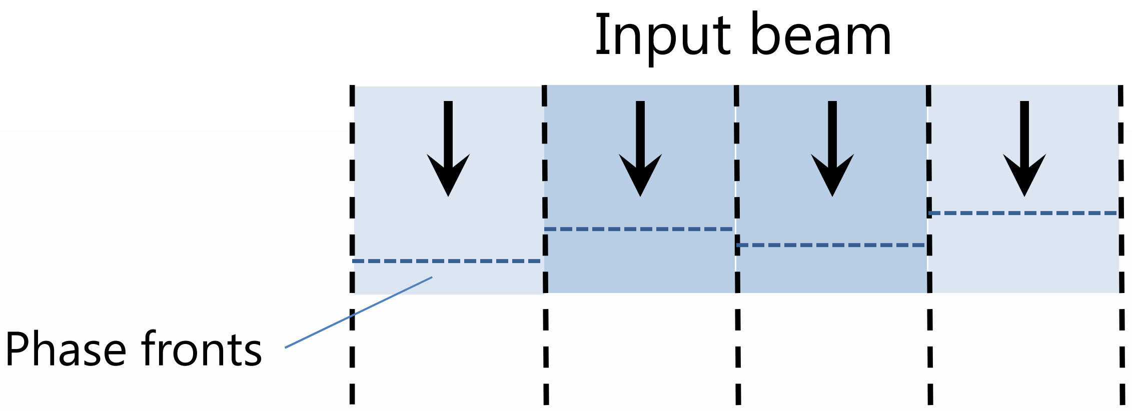


DM "Self-aligning universal beam coupler," Opt. Express **21**, 6360-6370 (2013)

Suppose, for simplicity, that

an arbitrary input beam can be adequately described by splitting it into 4 sections

Coupling an arbitrary input beam



DM "Self-aligning universal beam coupler," Opt. Express **21**, 6360-6370 (2013)

Suppose, for simplicity, that

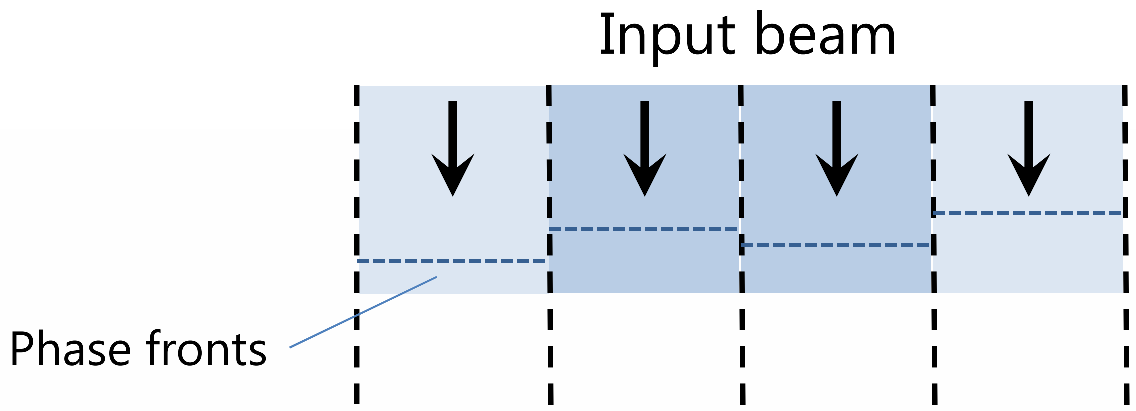
- an arbitrary input beam can be adequately described by splitting it into 4 sections,
  - each approximately uniform in intensity and "flat" in phase.

For simplicity, neglect diffraction for the moment

- assuming each of these sections will propagate as a "square" section of the beam

We are effectively "sampling" the beam in four "chunks" for the moment

Coupling an arbitrary input beam

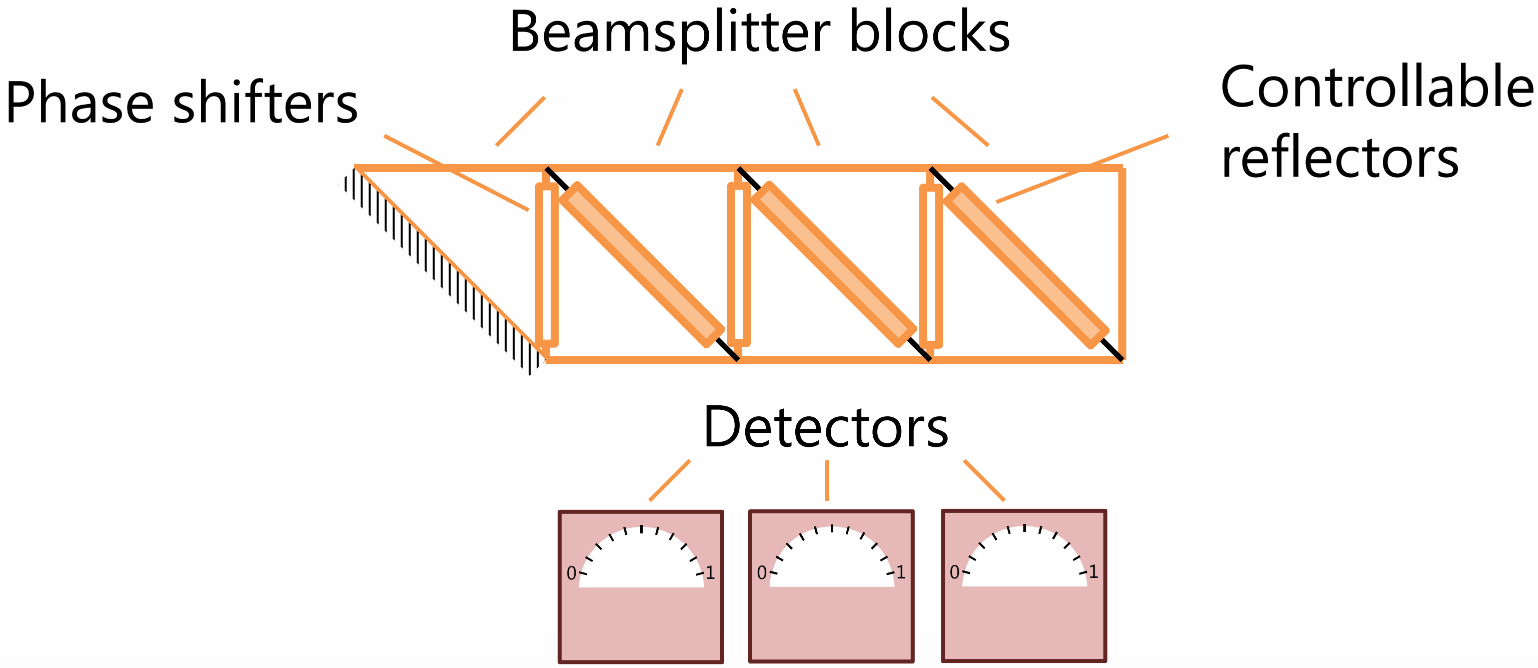


DM “Self-aligning universal beam coupler,” Opt. Express **21**, 6360-6370 (2013)

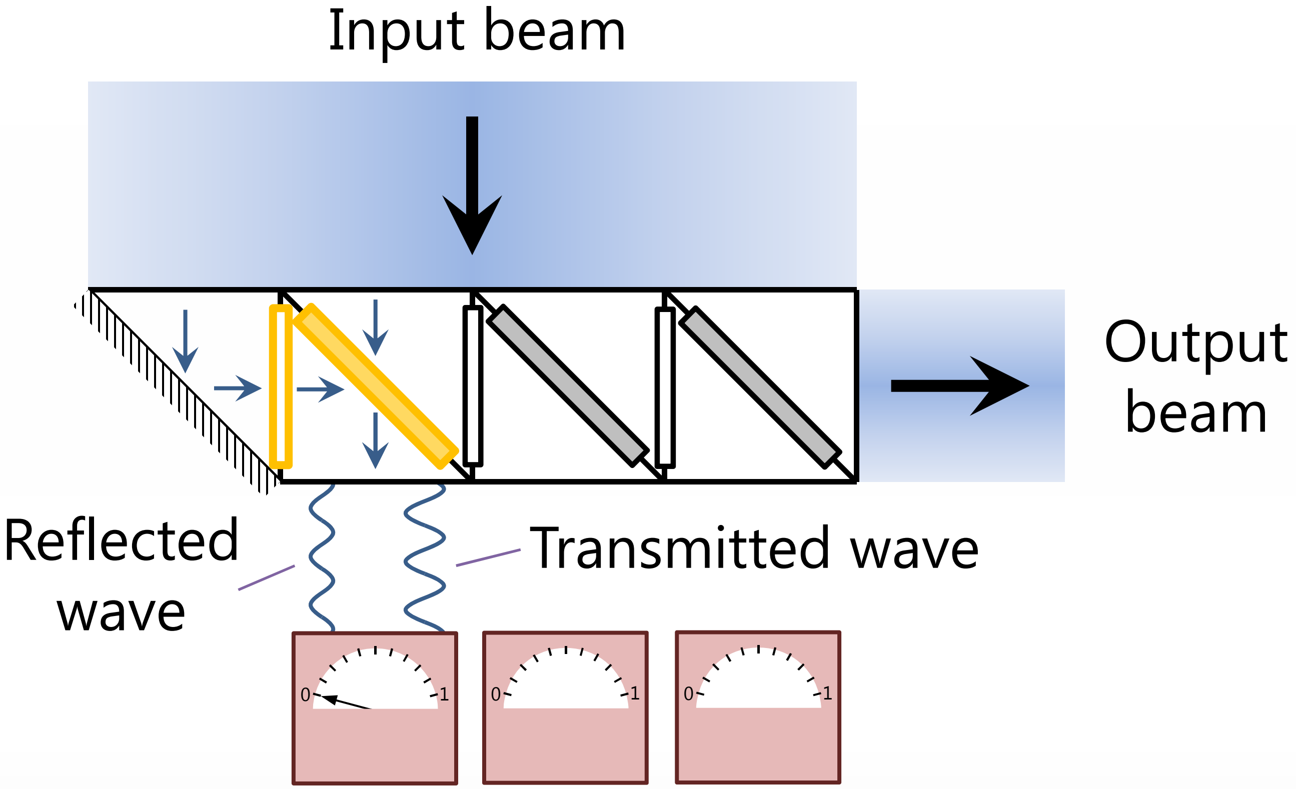
Now we build an apparatus, based on

- beamsplitters, including
  - adjustable phase shifters
  - adjustable reflectors
- successive power minimizations on detectors
  - to couple all the input power into one standard output beam

Self-aligning beam coupler



Self-aligning beam coupler



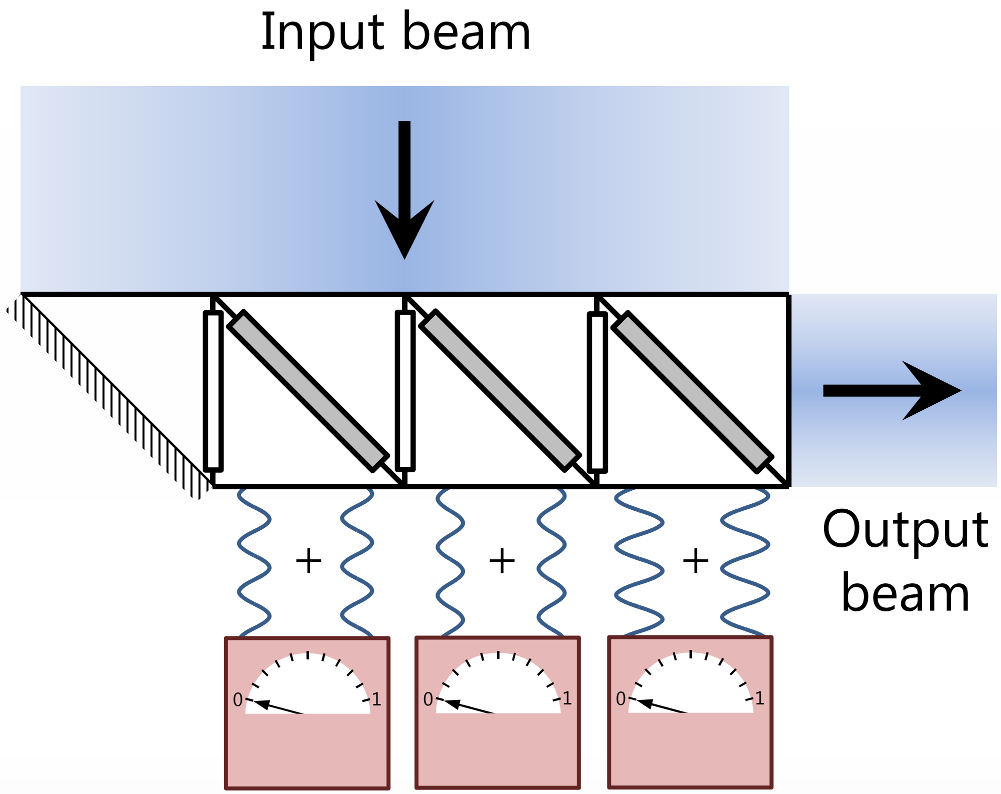
Adjust phase shifter in first block to minimize power in first detector

Adjust reflectivity in first block to minimize power again in first detector

Repeat for each block

Leaves no power in detectors, all input power in output beam

Self-aligning the beam coupler



Now all the input beam power is coupled to the output beam

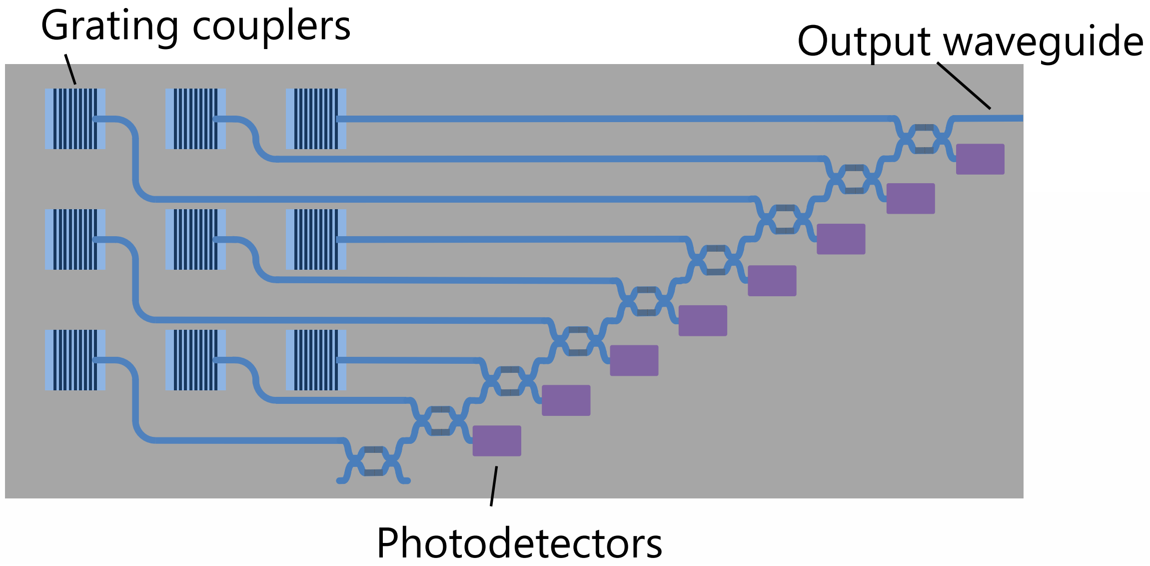
- Regardless of the form of the input beam
  - And without any calculation or detailed calibration of devices

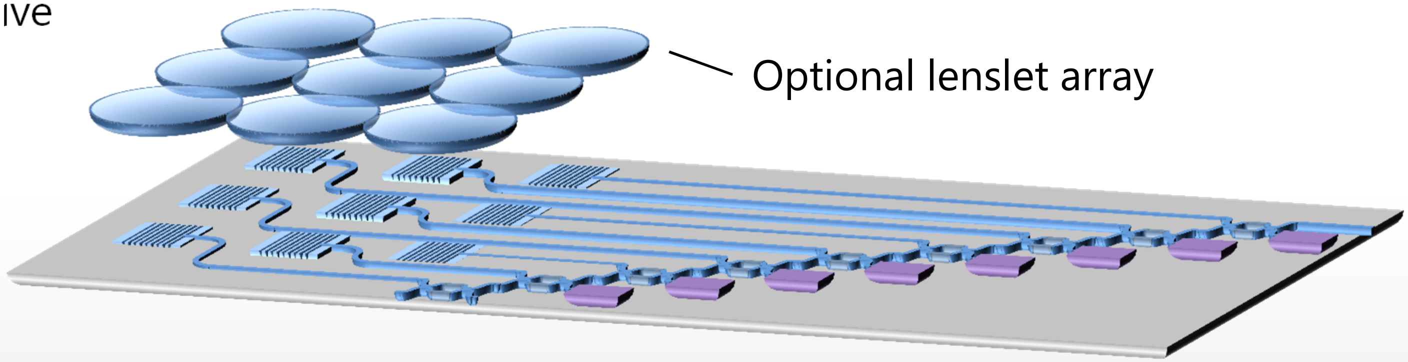
Self-aligning the beam coupler

- Sequential process
  - No overall iteration
- Only local minimization feedback loops
  - On one parameter at a time
    - No multiparameter global optimization
- **No calculations at all!**
- We can leave this process running all the time
  - Continually optimizing as devices drift
    - Or
  - Optimizing for changing inputs, e.g.,
    - Atmospheric turbulence
    - Moving sources

DM "Self-aligning universal beam coupler," Opt. Express **21**, 6360-6370 (2013)

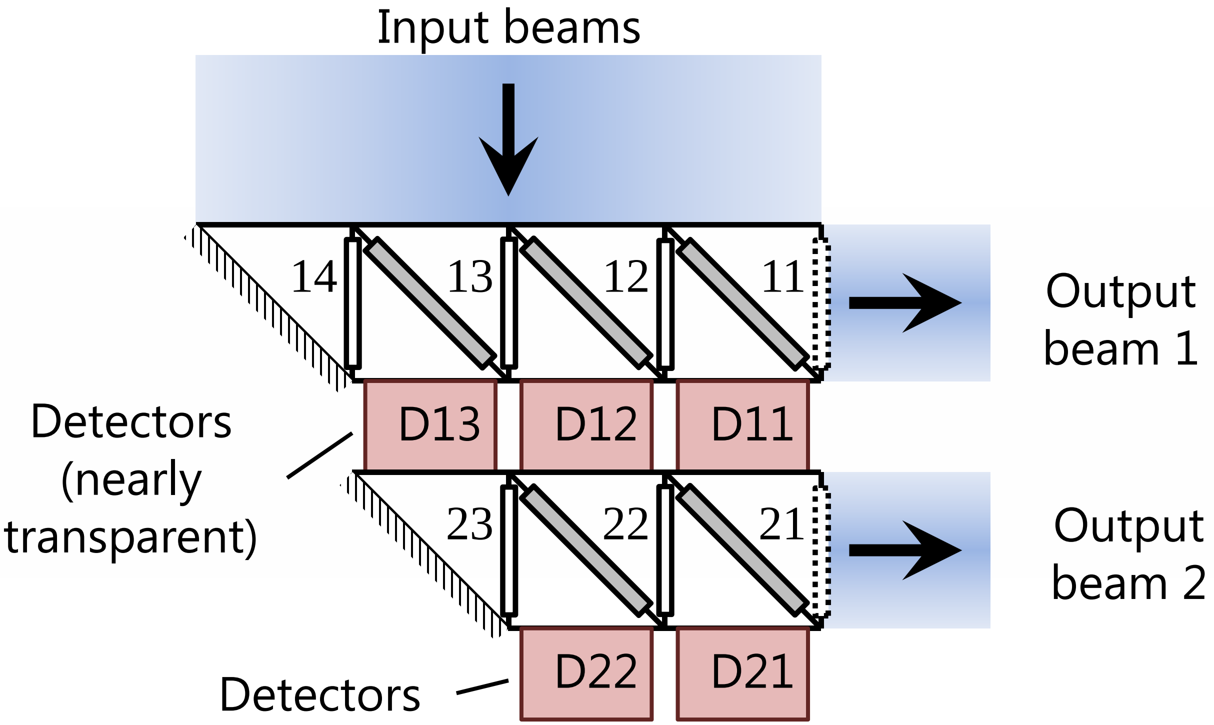
Mach-Zehnder self-aligning implementations





Grating couplers would allow us to couple a free-space beam to a Mach-Zehnder implementation of the device

Self-aligning multiple orthogonal beams

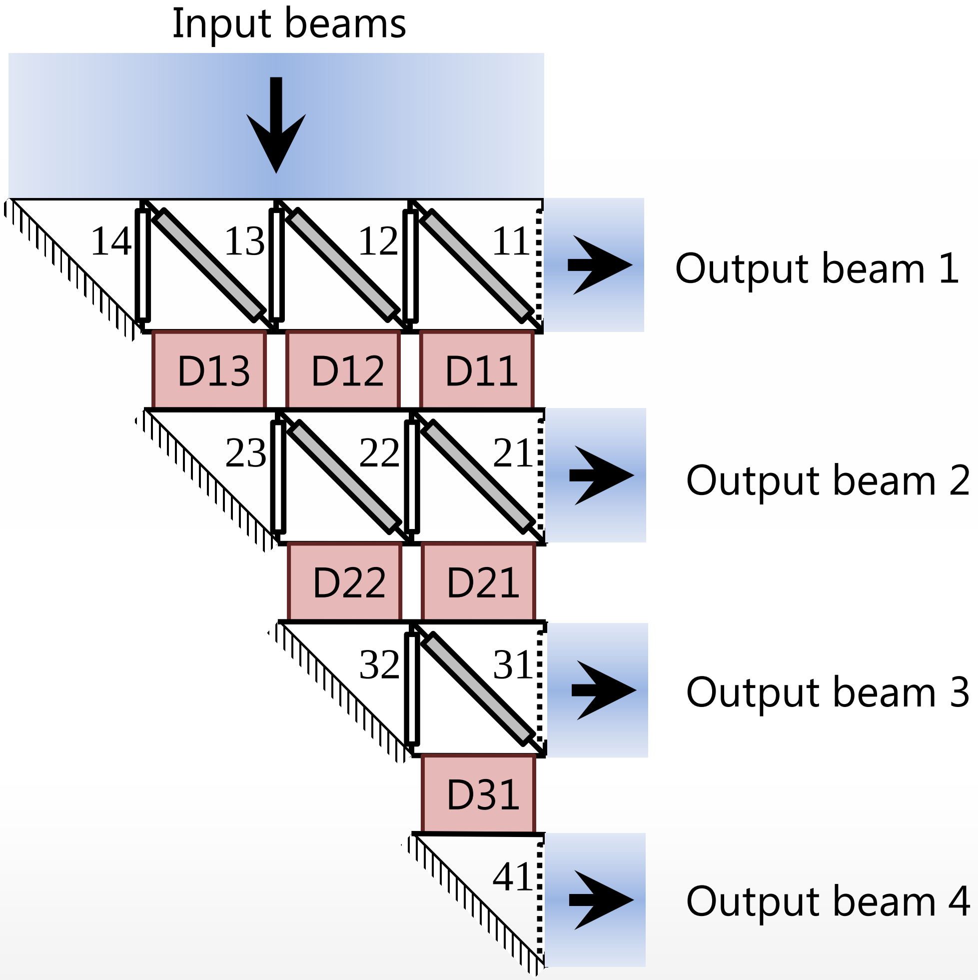


Once we have aligned beam 1 using detectors D11 – D13

- An orthogonal input beam 2 passes through the nearly transparent detectors to the second row
  - Where we can self-align it using detectors D21 – D22

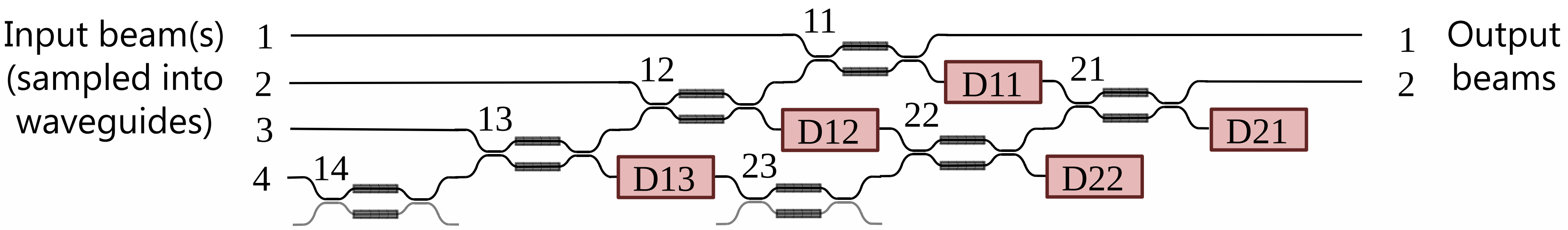
Separating two overlapping orthogonal beams to separate outputs

Self-aligning multiple orthogonal beams



- Adding more rows and self-alignments
  - Separates a number of orthogonal beams
    - equal to the number of beam “segments”
      - here, 4
- After self-aligning beam 3
  - Beam 4 is also self-aligned
    - No more degrees of freedom

Self-aligning 2 beam Mach-Zehnder implementation



D11, D12, D13 are mostly-transparent detectors

- Since alignment and re-alignment need not be performed at data-rate speeds
  - Only need small signals from the detectors
    - Many ways of making such mostly-transparent detectors

Extensions

- Make completely arbitrary spatial optical device [1]
- Extend to different polarizations [1]
- Find optimal orthogonal optical channels through any linear optical system or scatterer [2]
- Make a spatial add/drop multiplexer for arbitrary spatial modes [3]
- Prove [1] that any physically legal linear optical device can be in principle, designed, made and self-configured, including
  - multiple wavelength optics
  - time-dependent optics
  - non-reciprocal optics

[1] "Self-configuring universal linear optical component," Photon. Res. **1**, 1-15 (2013)

[2]"Establishing optimal wave communication channels automatically," J. Lightwave Technol. DOI: 10.1109/JLT.2013.2278809

[3] "Reconfigurable add-drop multiplexer for spatial modes," Opt. Express **21**, 20220-20229 (2013)

Conclusions

For a copy of these
slides, send e-mail to
dabm@ee.stanford.edu

- What is the big problem in information processing?
  - How to keep scaling to keep up with demand
    - energy and bandwidth density within machines
- Solutions
  - Nanotechnology for logic and interconnect
  - Optics for interconnect
    - Only solution for bandwidth density and interconnect energy
      - Need low energy optoelectronics and dense nanophotonics integrated with silicon
- Where are the opportunities?
  - Ultra-low energy optoelectronics?
  - Nanometallics for local light concentration?
  - Novel optical design approaches?
    - We now know how to make arbitrary linear optical devices
      - including arbitrary mode separators
      - without any calculations!

References and links

For a copy of these slides, send e-mail to dabm@ee.stanford.edu

Home page

See http://www-ee.stanford.edu/~dabm/ for links to all topics

- "Quantum mechanics for scientists and engineers" MOOC
  http://class.stanford.edu/
- Self-configuring arbitrary optics
  See overview at
  http://www-ee.stanford.edu/~dabm/Selfalign.html

"All linear optical devices are mode converters," Opt. Express **20**, 23985-23993 (2012)

"How complicated must an optical component be?" J. Opt. Soc. Am. A **30**, 238-251 (2013)

"Self-aligning universal beam coupler," Opt. Express **21**, 6360-6370 (2013)

"Self-configuring universal linear optical component," Photon. Res. **1**, 1-15 (2013)

"Establishing optimal wave communication channels automatically," J. Lightwave Technol. DOI: 10.1109/JLT.2013.2278809

"Reconfigurable add-drop multiplexer for spatial modes," Opt. Express **21**, 20220-20229 (2013)