Concepts for wave-based computing

stanford.io/4rdZDSJ

David Miller
Stanford University

What I am *not* going to talk about

What we should certainly do with optics in computing

- **optical interconnects**
  - but a new generation based on massive parallelism at moderate rates
    - Greatly reduces power dissipation in "electronic" information
      - which is mostly due to interconnect, not logic

What we should probably not do

- optical digital computing
  - and generally not "optical transistors"

"Attojoule Optoelectronics for Low-Energy Information Processing and Communications: a Tutorial Review," IEEE/OSA JLT **35** (3), 343 (2017)

stanford.io/4rdZDSJ

"Are optical transistors the next logical step?" Nature Photonics **4**, 3 (2010)

A new generation of optical interconnects

A next generation of highly parallel free-space optical interconnects to

- eliminate most of the energy of short (and longer) interconnects
  - which is most of the energy in datacenters
- and scale to increased bandwidth density

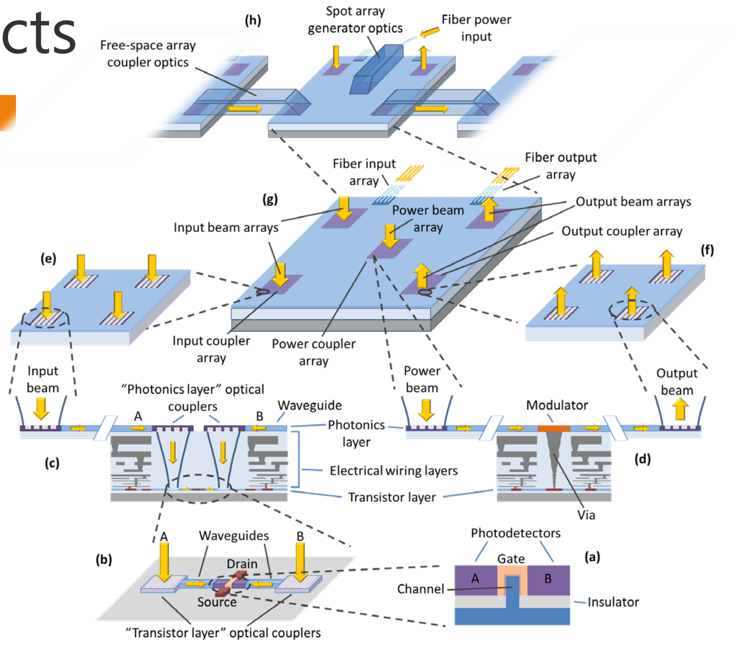
A reasonable goal – **100 - 10 fJ/bit (total system energy) up to 10 m distance**

- Note 10 fJ/bit implies only
  - **10 mW power for 1 Tb/s** interconnect bandwidth

Research on this has been completed some time ago

- This awaits investment, development and commercialization

See also this video (OFC 21, dabm.stanford.edu/videos/#OFC2021)



"Straw man" system concept exploiting

- tightly integrated optoelectronics
- efficient beam couplers
- free-space communications with 1000's to 10,000's of channels

DM, "Attojoule Optoelectronics for Low-Energy Information Processing and Communications: a Tutorial Review," IEEE/OSA JLT **35** (3), 343 (2017)

A new generation of optical interconnects

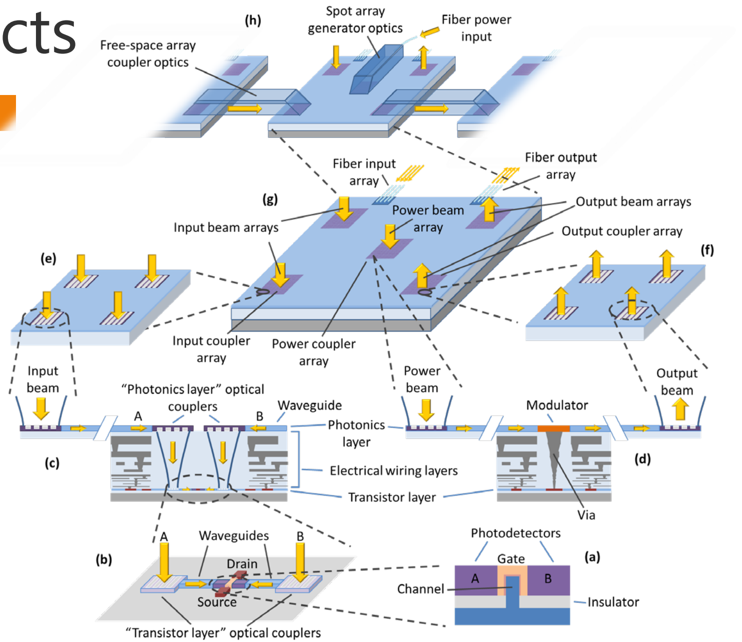
Converting from optics to electronic and back

is not fundamentally slow or inefficient

if we do it in an integrated way

with the right technology

and without unnecessary overheads



"Straw man" system concept exploiting

- tightly integrated optoelectronics
- efficient beam couplers
- free-space communications with 1000's to 10,000's of channels

See also this video (OFC 21, dabm.stanford.edu/videos/#OFC2021)

DM, "Attojoule Optoelectronics for Low-Energy Information Processing and Communications: a Tutorial Review," IEEE/OSA JLT **35** (3), 343 (2017)

What I am going to talk about

Some questions and concepts for how best to make

complex analog systems for wave-based computing

especially the kinds of

architectures and algorithms we need

to make them work

especially if we want them to be flexible

Processors for fixed problems?

Programmable vs. fixed function

- How much use there is for fixed analog processors?
  - optical computing has failed in the past in part because of its lack of general programmability
    - Do we have applications for wave-based computing in which a fixed function is useful enough?

Some fixed physical problems are certainly worth solving well

- efficient wave couplers into and out of waveguides
- converting from large continuous basis sets to discrete basis sets like waveguides or waveguide modes
- specific elements in, e.g., microscopy, for phase and angle contrast

Perhaps there are other very useful fixed transformations that must be performed on multiple different inputs

- Fourier transforms?
- fixed “front end” processing on images?
- …

Obviously, if the optical or wave system that solves some problem

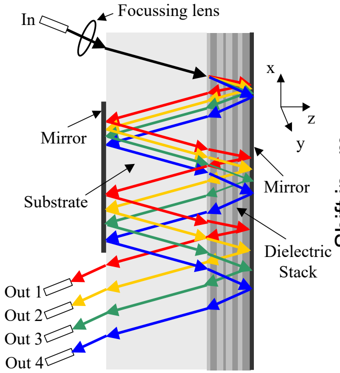
- takes greater computational effort in design
  - than actually solving the problem itself
    - it is not worth the design cost

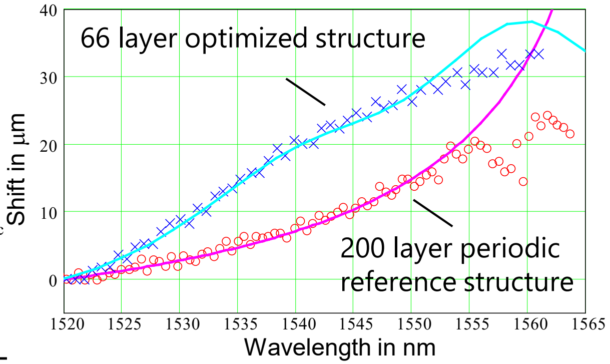
Note too that fixed complex wave systems may be hard to manufacture precisely enough

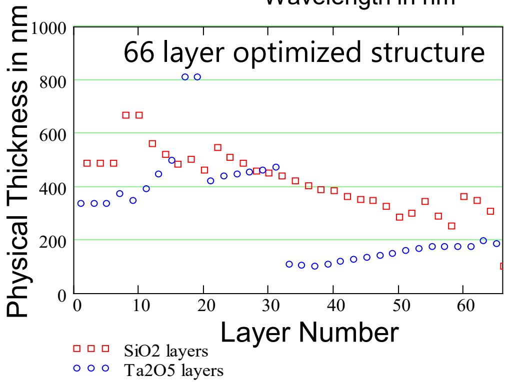
Example complex optics – custom superprism

- A multilayer dielectric stack
  - with custom layer thicknesses
    - can make a good wavelength demultiplexer by a “superprism” effect
- A 66 layer non-periodic structure worked well
  - giving ~ linear shift with wavelength
- A 200 layer periodic structure did what it should do
  - not as good as the 66 layer optimized structure
- A 200 layer non-periodic structure did not work at all
  - because there could not be any feedback control during manufacture
    - for such a complex structure

Complex fixed structures without any adjustment just may not work







"Wavelength demultiplexer using the spatial dispersion of multilayer thin-film structures," IEEE Phot. Tech. Lett. **15**, 1097 (2003); “Multilayer Thin-Film Structures with High Spatial Dispersion,” Appl. Opt. **42**, 1330 (2003)

Processors for varying problems?

For the rest of this talk, though

- I presume we are interested in making the wave-based computing system
  - easy to design and operate
    - so we can easily handle multiple different complex problems
  - possibly even programmable in real time to adapt to the problem of interest
  - possibly even self-configuring to the problem
  - ideally also self-stabilizing
    - because complex analog processors otherwise just may not continue working
      - or may not even work at all given fabrication variations

Given this desire for ease of design, stability, programmability, and self-configuration

- what does that imply for the architectures and algorithms we need?

A mathematical framework for linear wave processors

stanford.io/4rdZDSJ

A mathematical framework for linear wave processors

I presume we are working with

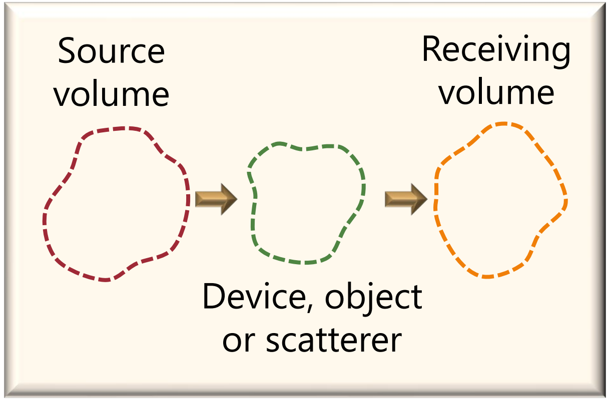
- some sources in a source volume
  - that couples through some device, object, or scatterer
    - which is effectively "processing" our waves
  - to give waves in some receiving volume

A very good way to look at such problems

- is to find the set of orthogonal source functions
  - that couple, one by one
    - to orthogonal wave functions in the receiving volume

These pairs of functions can be called

- communication modes
  - if we are thinking of these as orthogonal channels for communication
- or mode-converter basis sets
  - if we are thinking about what the object does



"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679-825 (2019)

A mathematical framework for linear wave processors

These sets of functions ***always exist*** for any linear coupling

they are found from the singular value decomposition (SVD) of the coupling operator between the spaces

This approach

- allows clear counting of channels and understanding of their coupling strengths

  including ultimately how the couplings fall of to give as practically finite number of channels

- establishes the most economical way of describing this wave system
- links directly to basic physics that applies only to these functions

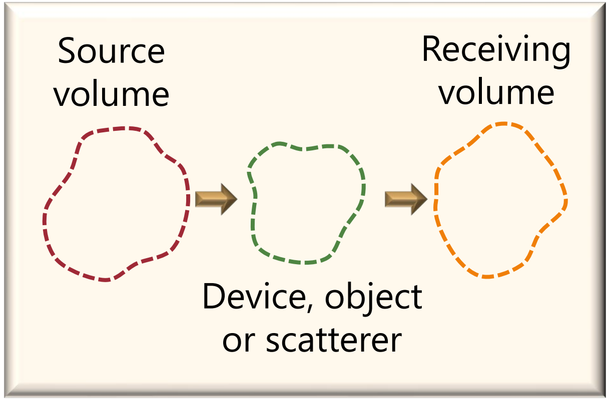
  radiation laws

  modal Einstein A&B coefficients

Note that the modes in this problem

are pairs of functions

and are ***not*** the beams between the volumes



"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679-825 (2019)

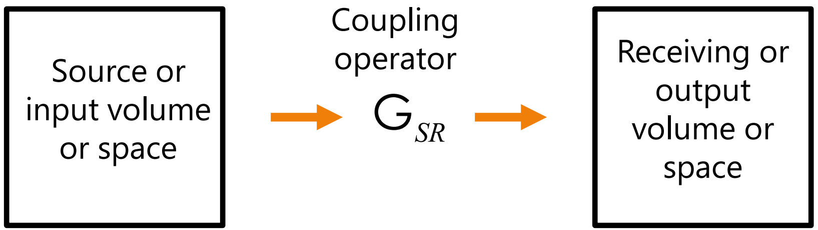
A mathematical framework for linear wave processors

These communication modes

- completely and uniquely define
  - all the orthogonal channels in the system
    - e.g., for communication or sensing
    - e.g., for understanding limits to numbers and strengths of channels and couplings

There are no better orthogonal channels

- If we can't do something using these channels
  - then we can't do it any other way with the same optics



"Waves, modes, communications, and optics," Adv. Opt. Photon. **11**, 679 (2019)

"Communicating with Waves Between Volumes ...," Appl. Opt. **39**, 1681 (2000)

Singular value decomposition (SVD)

For any linear operator D

at least as long as it is bounded, i.e., finite output for finite input

we can perform the singular value decomposition

$$\mathrm{D} = \mathrm{V}\mathrm{D}_{diag}\mathrm{U}^{\dagger} \quad \text{or equivalently} \quad \mathrm{D} = \sum_{m} s_m \left|\phi_m\right\rangle\left\langle\psi_m\right|$$

U and V are unitary operators ($\mathrm{U}^{\dagger}$ is automatically also unitary)

D_{diag} is a diagonal operator with elements s_m

which are called the singular values

$\left|\psi_m\right\rangle$ are the columns of U (and $\left\langle\psi_m\right|$ are the rows of $\mathrm{U}^{\dagger}$)

and are the orthogonal source functions

$\left|\phi_m\right\rangle$ are the columns of V

and are the orthogonal resulting wave functions

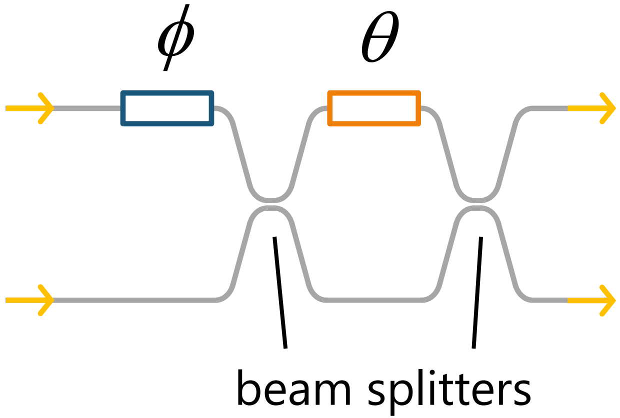
A prototypical wave processing system – Mach-Zehnder interferometer meshes

stanford.io/4rdZDSJ

Nulling a Mach-Zehnder output

Consider a waveguide Mach-Zehnder interferometer (MZI)

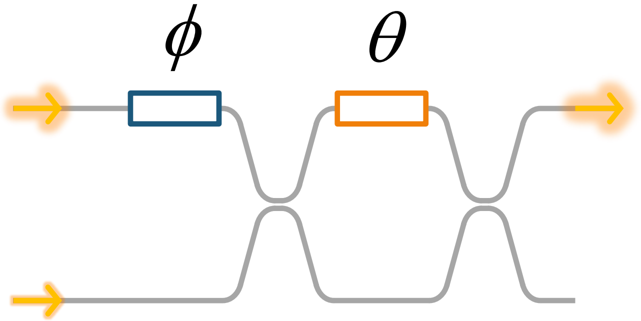
- formed from two “50:50” beam splitters
- and at least two phase shifters
  - one, ϕ, to control the relative phase of the two inputs
  - a second, θ, to control the relative phase on the interferometer “arms”



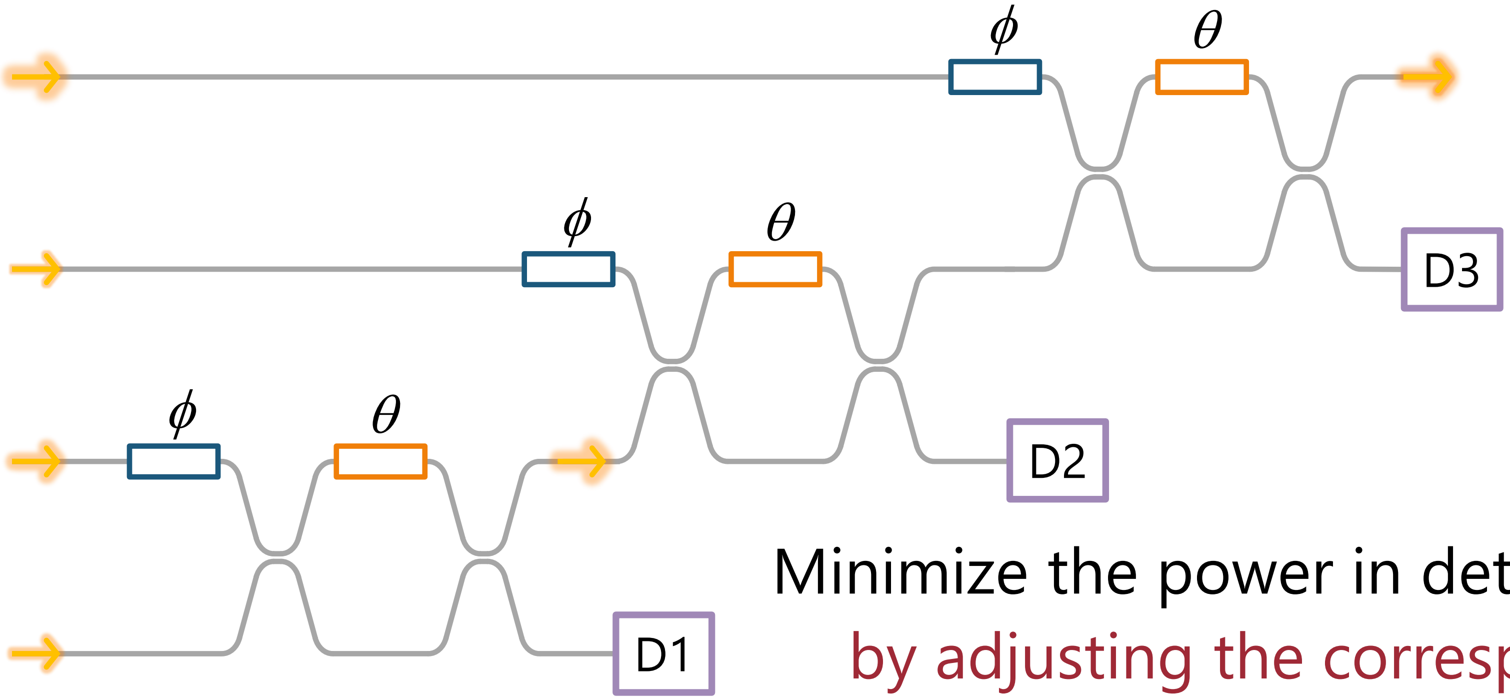
Nulling a Mach-Zehnder output

In such an MZI with 50:50 beamsplitters

- for any relative input amplitudes and phases
  - we can "null" out the power at the bottom output
    - by two successive single-parameter power minimizations
      - first, using ϕ
      - second, using θ



"Diagonal line" self-aligning coupler



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

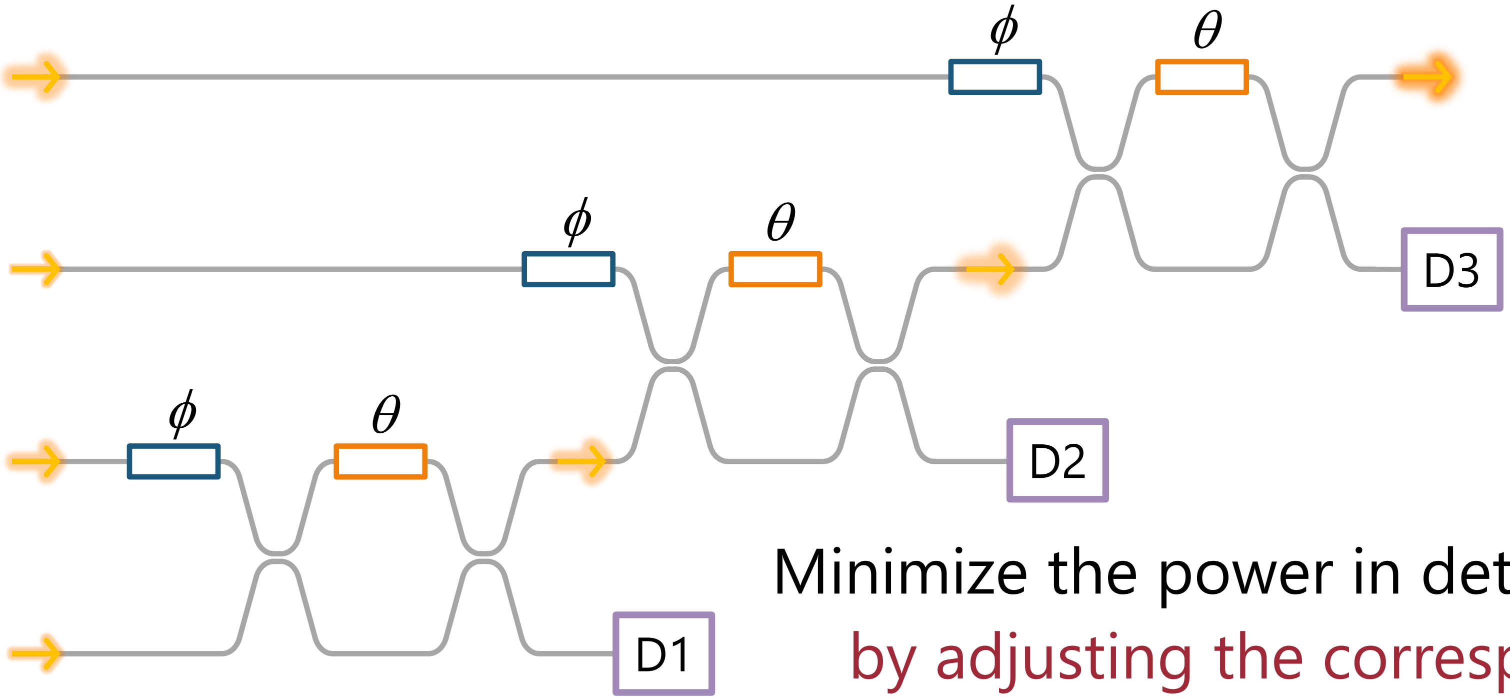
Minimize the power in detector D1

by adjusting the corresponding ϕ

and then θ

putting all power in the upper output

“Diagonal line” self-aligning coupler



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

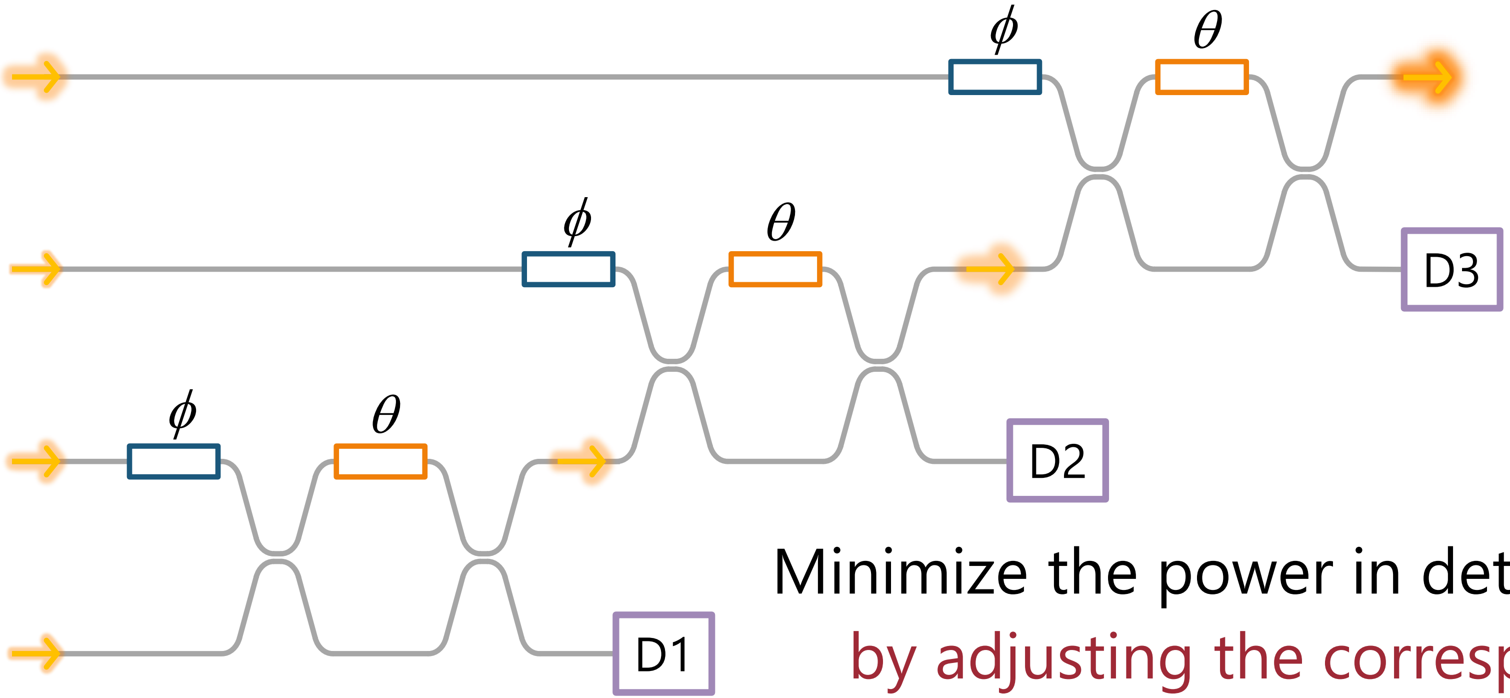
Minimize the power in detector D2

by adjusting the corresponding ϕ

and then θ

putting all power in the upper output

“Diagonal line” self-aligning coupler



Minimize the power in detector D3
by adjusting the corresponding ϕ
and then θ
putting all power in the upper output

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

Self-aligning beam coupler

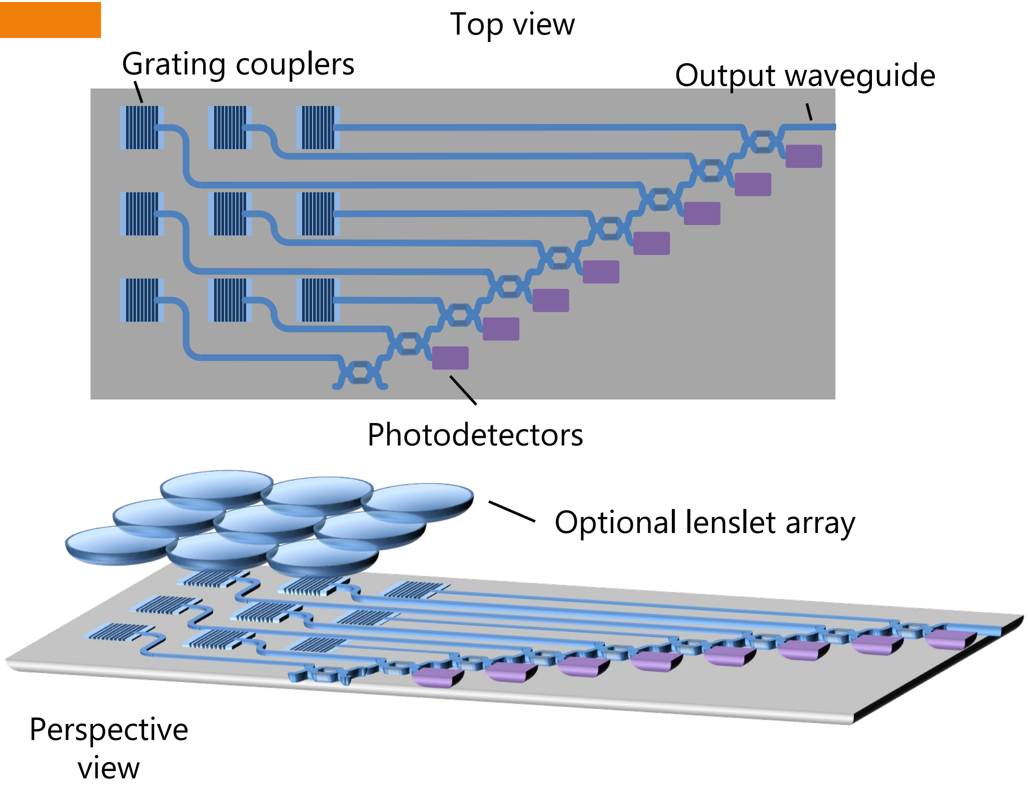
Grating couplers could couple a free-space beam to a set of waveguides

Then

we could automatically couple all the power to the one output guide

This could run continuously

tracking changes in the beam

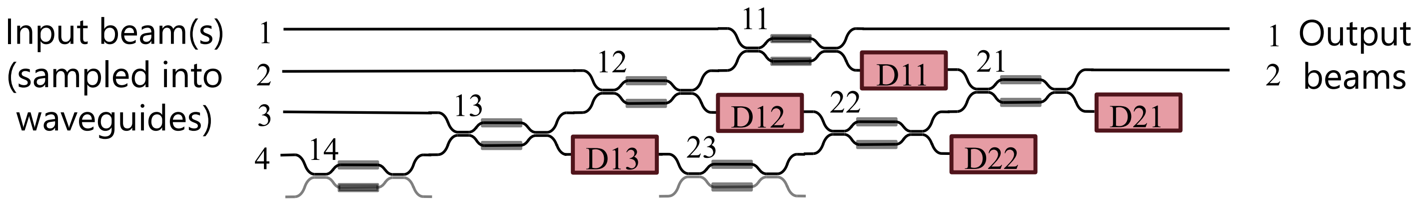


"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Separating beams with interferometer meshes

stanford.io/4rdZDSJ

Separating multiple orthogonal beams



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

Once we have aligned beam 1 to output 1 using detectors D11 – D13

an orthogonal input beam 2 would pass entirely into the detectors D11 – D13

If we make these detectors mostly transparent

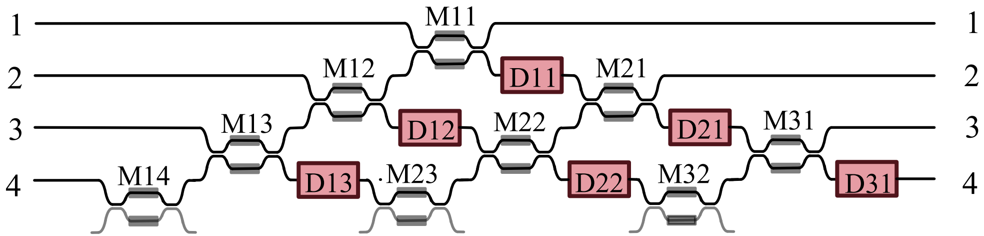
this second beam would pass into the second diagonal "row"

where we self-align it to output 2 using detectors D21 – D22

separating two overlapping orthogonal beams to separate outputs

Separating multiple orthogonal beams

Input beam(s) (sampled into waveguides)



Output beams

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

Adding more rows and self-alignments

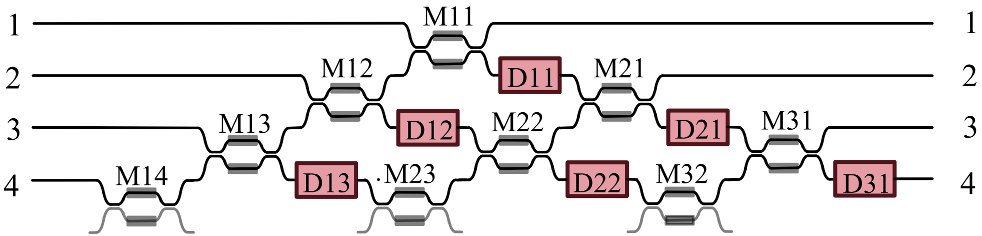
separates a number of orthogonal beams

equal to the number of beam "segments", here, 4

This makes an arbitrary 4x4 unitary processor

Separating multiple orthogonal beams

Input beam(s) (sampled into waveguides)



Output beams

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

If we put identifying "tones" on each orthogonal input "beam"

and have the corresponding diagonal row of detectors look for that tone

then the mesh can continually adapt to the orthogonal inputs

even when they are all present at the same time

and even if they change

solving the physical problem of separating overlapping light beams

Self-configuring beam separator

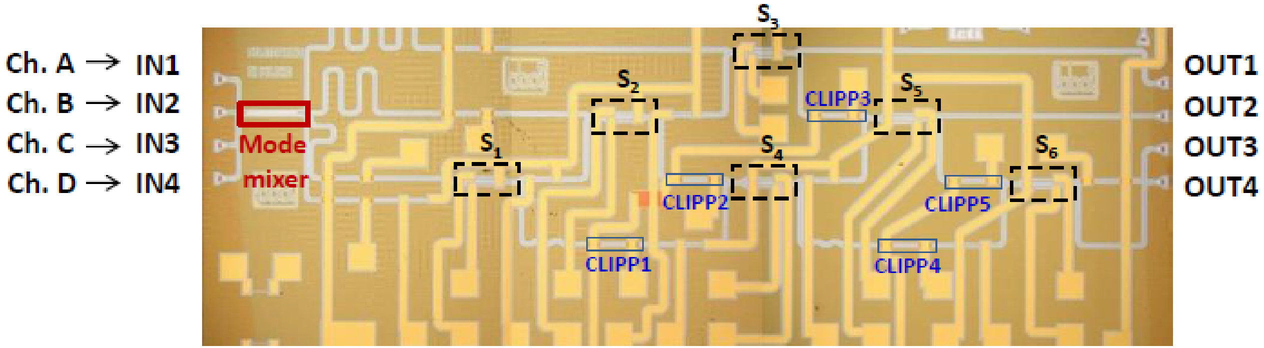
Light from four input fibers

deliberately mixed in a mode mixer

are automatically separated out again by a mesh of interferometers

by sequential power maximizations

without calculations



A. Annoni et al., "Unscrambling light – automatically undoing strong mixing between modes," Light Science & Applications 6, e17110 (2017)

See, e.g., review W. Bogaerts et al., "Programmable photonic circuits," Nature 586, 207 (2020)

Universal matrix multiplier chip

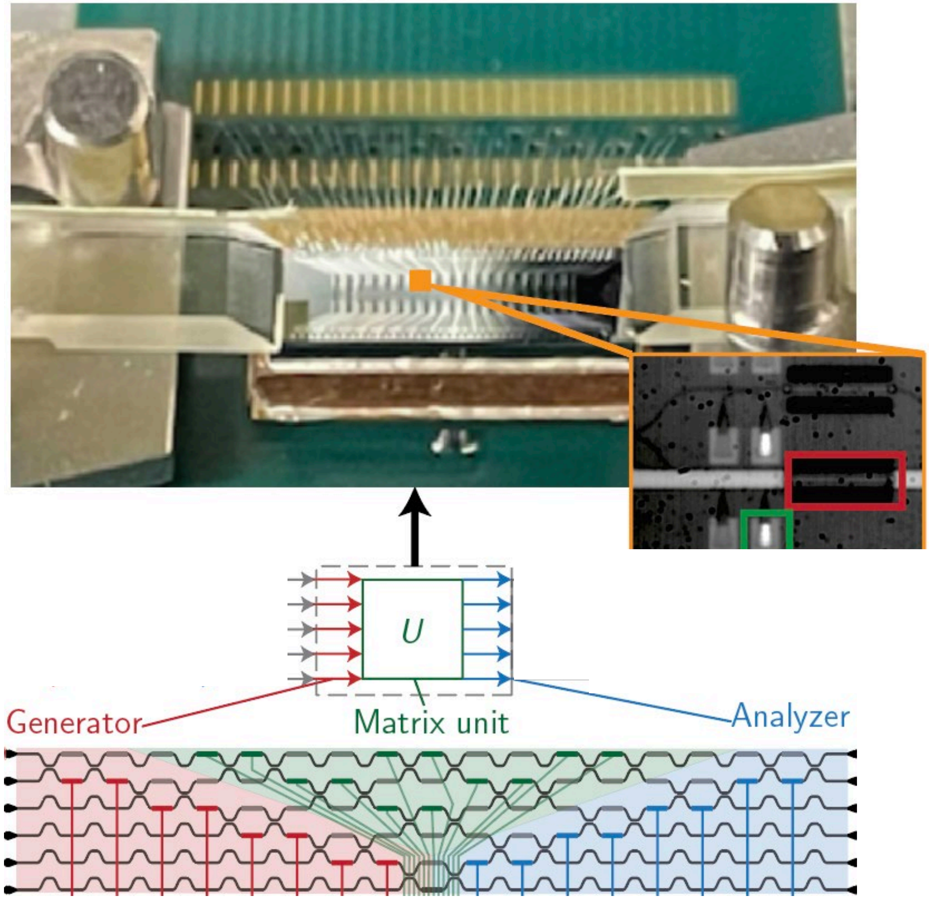
Universal matrix multiplying chip

"4x4" unitary Mach-Zehnder mesh with

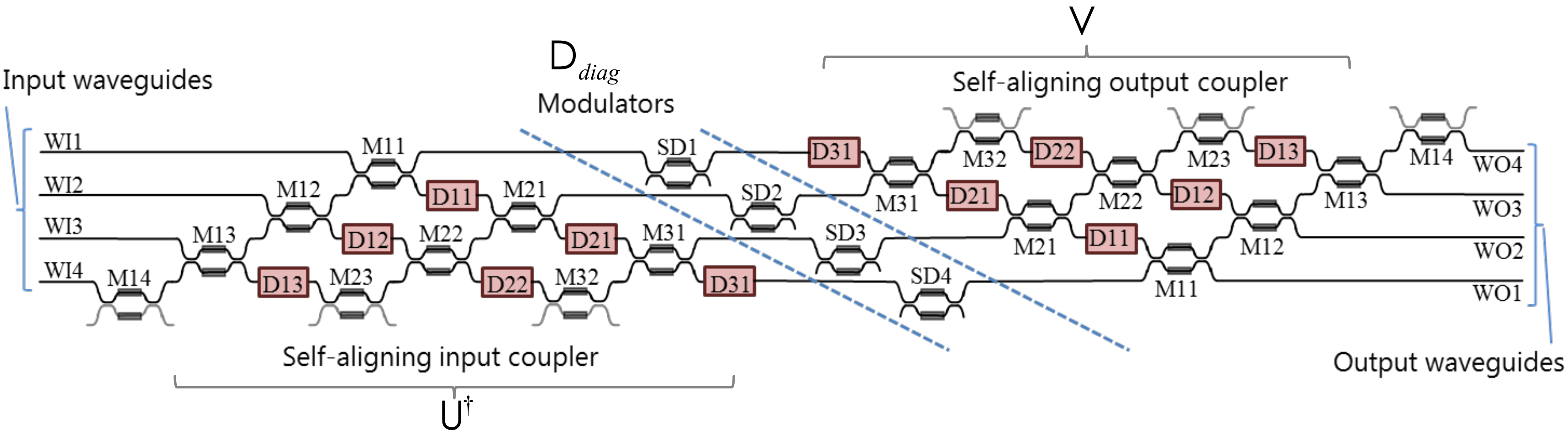
- ❑ a "generator" to create any complex input vector
- ❑ an "analyzer" to measure the complex output vector

This can be programmed to implement

any "unitary" (loss-less) transformation from the inputs to the outputs



General multiple mode converter



This mesh implements an arbitrary matrix from its the SVD $D = VD_{diag}U^\dagger$

So, for an optical system of a given dimensionality

we can emulate any linear optical system

This is the first proof of the possibility of arbitrary linear optics

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

Note we are implementing an arbitrary linear optical component

by constructing it using its communication mode or "mode converter" basis sets

Interferometer meshes as example wave processors

Interferometer meshes
- which have many working demonstrations in silicon photonic systems
- are good example architectures to help us think about linear optical processing generally
  - being able in principle to implement any linear operation at a given wavelength between inputs and outputs
    - and it is easy to design that mesh
      - and it is minimally complex
        - with just the right number of adjustable parameters
- show us that we can decompose any linear wave system
  - into a set of two-wave interferences
- give us explicit architectures and topologies for wave processing systems
  - supporting specific configuration algorithms associated with topologies
    - including self-configuration
      - which breaks down the calibration, configuration, and stabilization into a set of simple feedback loops
        - often just sets of successive, progressive single-parameter power minimizations of maximizations
      - so giving an existence proof for stabilizing and operating large analog systems

Recent extensions now let us make corresponding functionalities in the spectral domain
- e.g., universal programmable and self-configuring spectrometers

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

Programmable and self-configuring filters

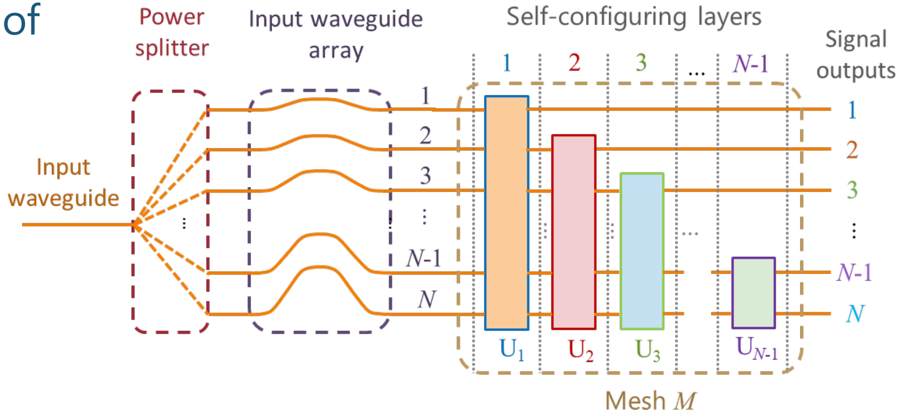
This proposed circuit can function like an arrayed waveguide grating filter

- but has a spectral response that is fully programmable
  - so it can implement any linear combination of such filter functions
  - and allows multiple different simultaneous filter functions

It can also

- self-configure to specific wavelengths
- reject N-1 arbitrary wavelengths
- measure and separate temporally partially coherent light
  - the Karhunen-Loève decomposition

D. A. B. Miller, C. Roques-Carmes, C. G. Valdez, A. R. Kroo, M. Vlk, Shanhui Fan, and O. Solgaard, "Universal programmable and self-configuring optical filter," Optica **12**, 1417-1426 (2025)



C. G. Valdez, A. R. Kroo, M. Vlk, C. Roques-Carmes, Shanhui Fan, D. A. B. Miller, and O. Solgaard, "Programmable Optical Filters Based on Feed-Forward Photonic Meshes," http://arxiv.org/abs/2509.12059

Solving a physical problem with a wave-based optical analog computer

stanford.io/4rdZDSJ

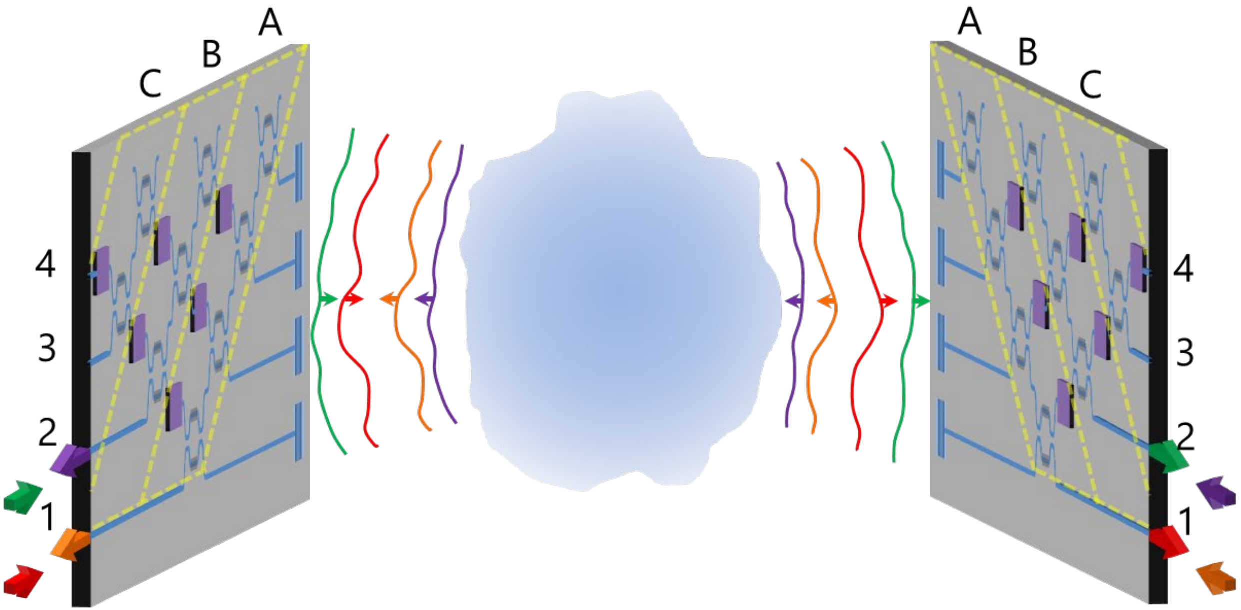
Establishing optimum orthogonal channels

In this architecture, using meshes on both sides

we proposed we could find optimal orthogonal channels through a scatterer

between waveguides on the left and waveguides on the right

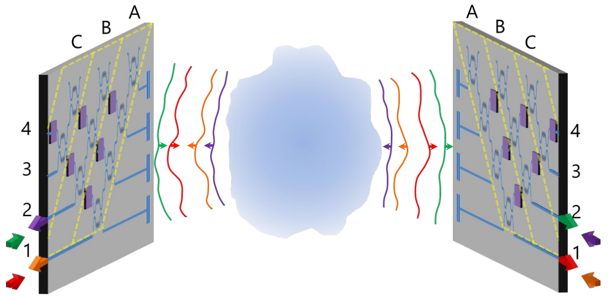
by iterating back and forward between the two sides



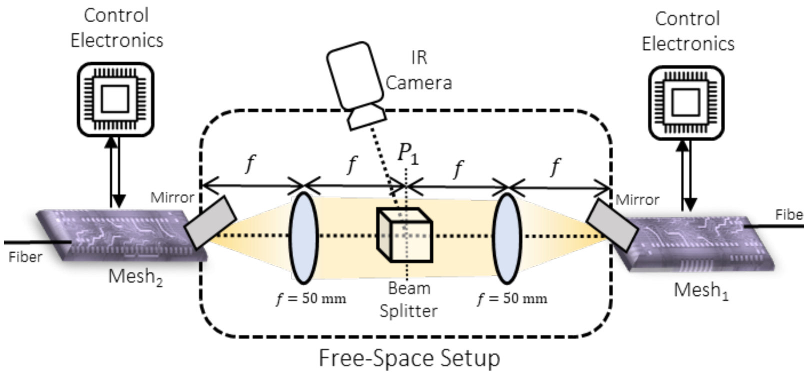
"Establishing optimal wave communication channels automatically," J. Lightwave Technol. 31, 3987 (2013)

Using optics to *perform* linear algebra

- By power maximizing on rows of the mesh at both sides
  - this circuit can automatically find the best orthogonal channels between the two sides
    - physically performing the singular-value decomposition of the optical system
- This is a true optical computer!
  - All calculations can be done in the optics
    - with only a sequence of simple single-parameter power optimizations
- If we change the optics in the middle
  - then the system automatically reconfigures itself
    - to find the best and orthogonal (low crosstalk) channels
      - from the inputs in the left to the outputs on the right
        - and *vice versa*
- Note that this processor is nonlinear
  - The nonlinear system exploits overall non-local nonlinearity
    - we change the optics (phase shifters) inside the system in response to measured optical output power through simple feedback loops
  - but the optics is linear



"Establishing optimal wave communication channels automatically," J. Lightwave Technol. 31, 3987 (2013)



S. SeyedinNavadeh et al., "Determining the optimal communication channels of arbitrary optical systems using integrated photonic processors," Nat. Photon. **18**, 149-155 (2024)

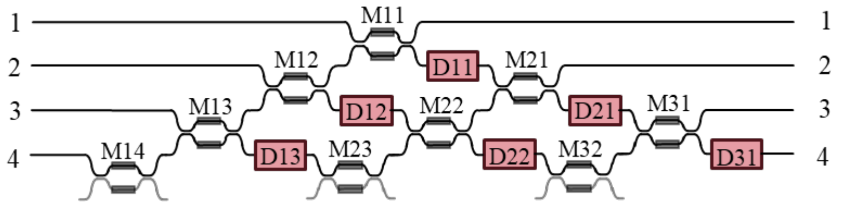
Architectural and algorithmic questions and approaches for designing wave-based computing systems

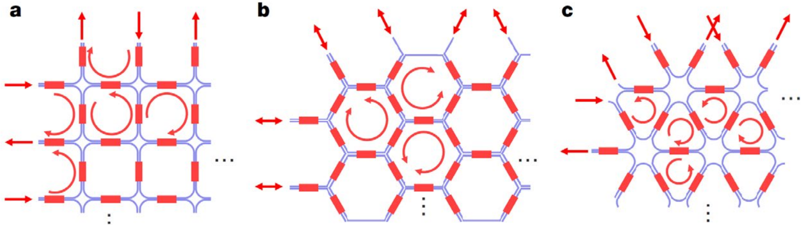
stanford.io/4rdZDSJ

Forward only vs. recirculating architectures

Wave processing architectures can be divided into two categories

- Forward-only
  - light only flows in one direction inside the processor
- Recirculating
  - light can flow backwards and forwards inside the processor
    - e.g., by scattering or reflections



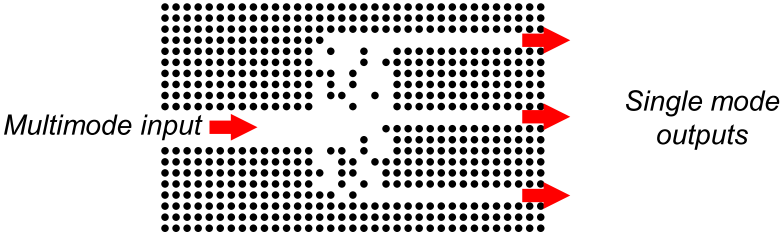


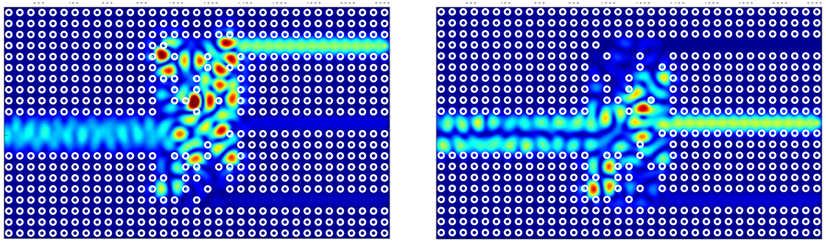
W. Bogaerts et al. "Programmable photonic circuits," Nature **586**, 207 (2020)

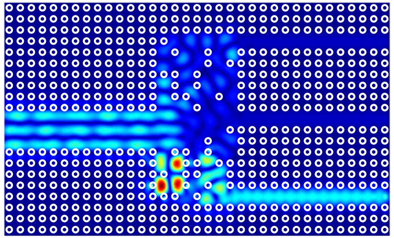
Recirculating architectures

Recirculating architectures

- such as many inverse-designed structures
  - can be very compact
    - but are generally much harder to design
      - because the design cannot be factorized into successive "blocks"
        - e.g., because of coherent back-reflections and resonances
  - and can be very difficult to program
    - because of the interaction between all parts of the structure, forward and backward
  - they cannot generally be factorized into successive linear operations





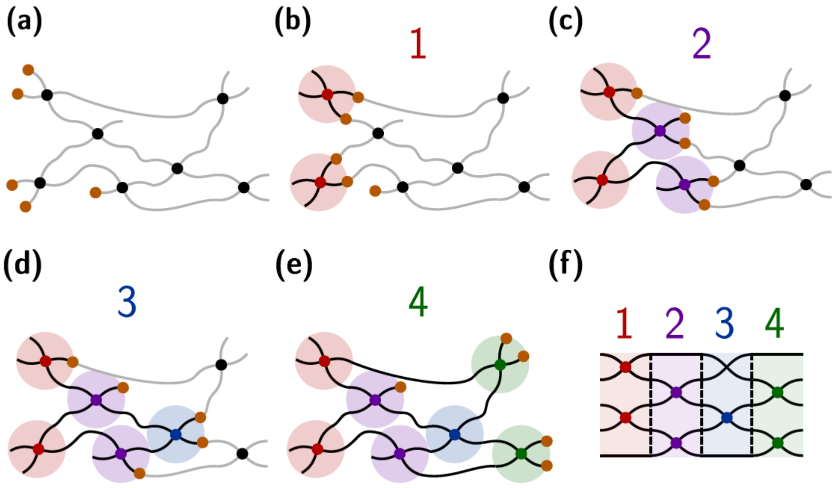


Engineer precise mode splitting with positioning of dielectric columns

"Demonstration of Systematic Photonic Crystal Device Design and Optimization By Low Rank Adjustments: an Extremely Compact Mode Separator," Optics Letters **30**, 141 (2005)

Forward only architectures

Forward-only architectures
- can be physically factorized into successive blocks
  - each with unitary operations and matrices
- are topologically directed acyclic graphs
  - defined by “column” topologies
    - each with functions that may be easier to understand physically
- are universal, capable of implementing any linear mapping between inputs and outputs at a given frequency
  - e.g., “SVD” interferometer mesh architecture
    - so recirculating architectures are not required just to implement functions at a given frequency

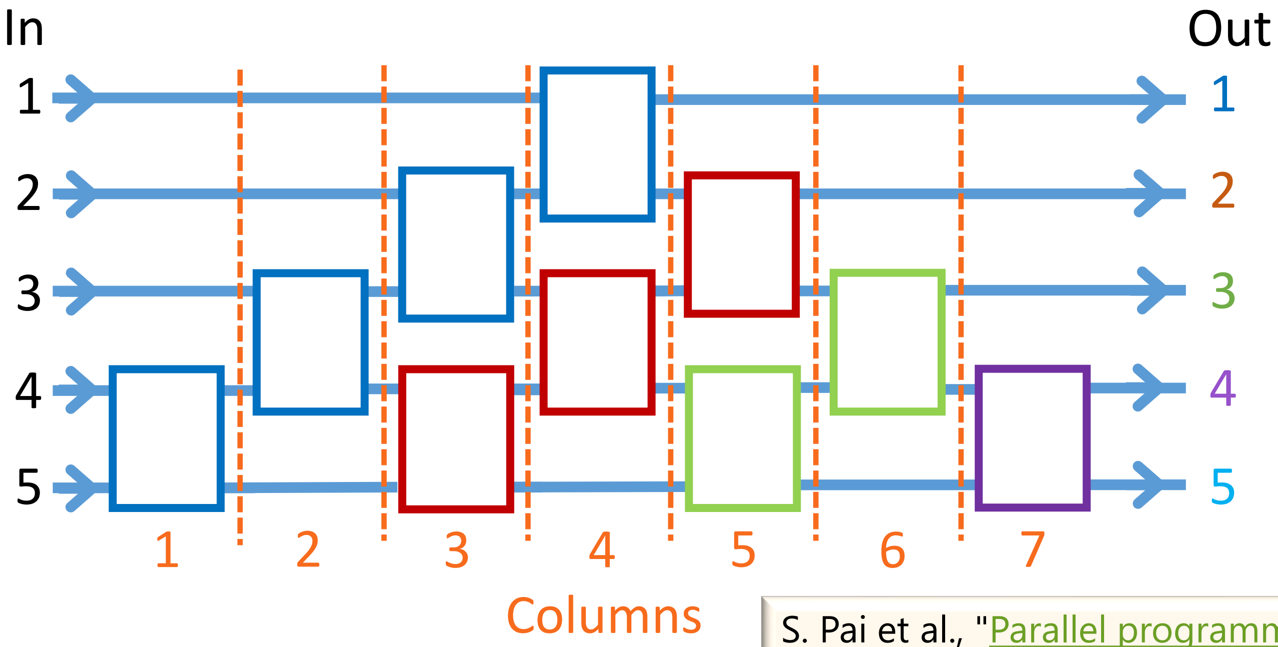


Topological sorting of an optical network into columns for parallel configuration

S. Pai et al., "Parallel programming of an arbitrary feedforward photonic network," IEEE J. Sel. Top. Quantum Electron. 25, 6100813 (2020)

Column topology

"Columns" can be identified with a simple topological algorithm and configured or calibrated in parallel



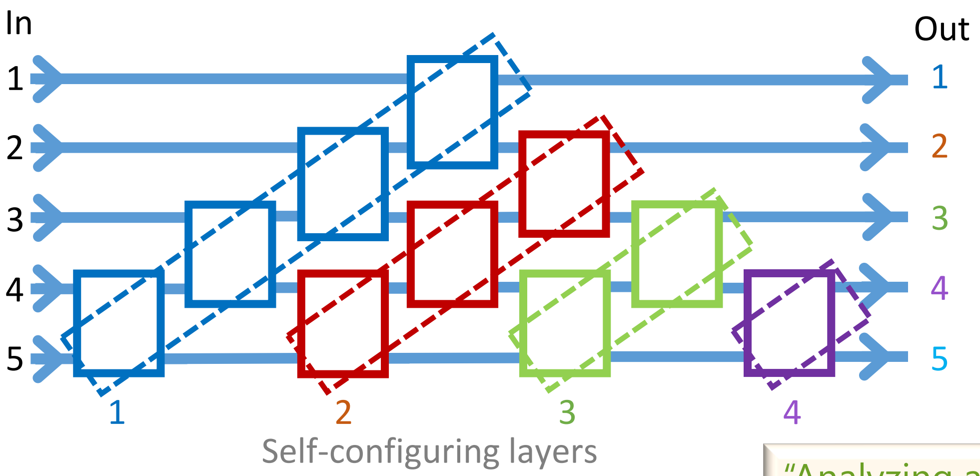
S. Pai et al., "Parallel programming of an arbitrary feedforward photonic network," IEEE J. Sel. Top. Quantum Electron. 25, 6100813 (2020)

Self-configuring layer topology

"Self-configuring layers" can also be defined topologically:

- they have one (and only one) connection path through 2x2 blocks from their output to each of their inputs

For example, a complete "triangular" mesh can be viewed as being built from successive "diagonal line" self-configuring layers



Not all mesh topologies support self-configuring layers

- e.g., a "rectangular" mesh does not

"Analyzing and generating multimode optical fields using self-configuring networks," Optica **7**, 794 (2020)

Algorithmically global vs factorizable architectures

Do we have to design the entire structure as a global optimization

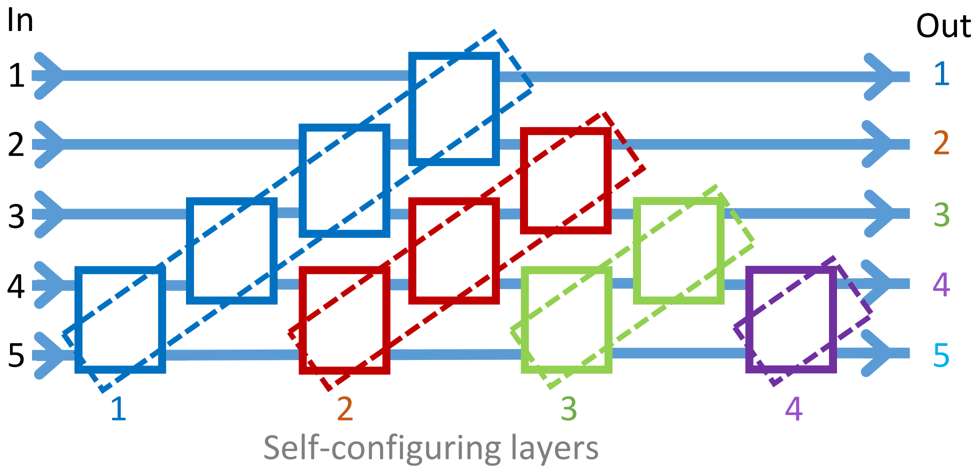
or can we factorize it into successive simpler designs?

e.g., can we "peel off" one problem at a time in an architectural "layer"

leading to a progressively simpler design for each subsequent layer

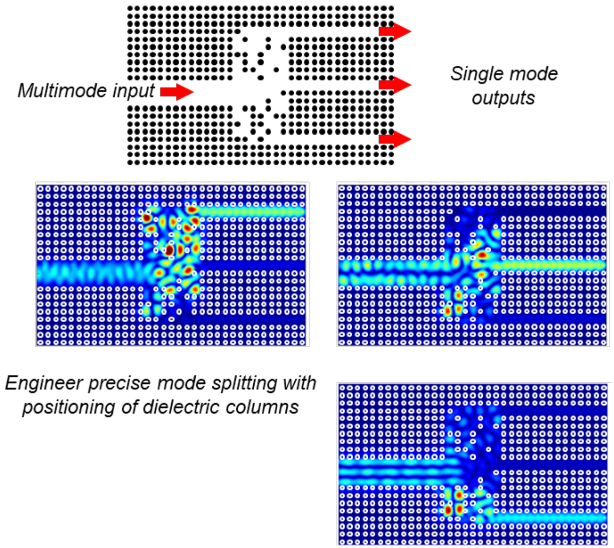
and separating those designs?

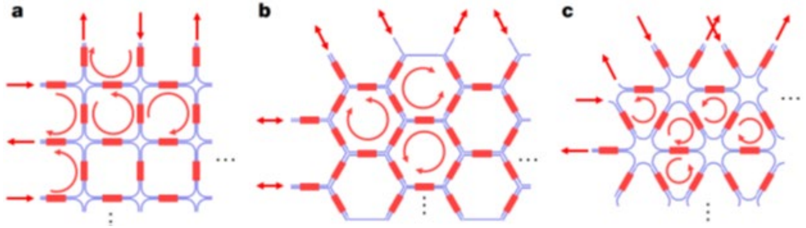
If so, we can call such a structure "algorithmically factorizable"



Algorithmically global vs factorizable architectures

- Algorithmically non-factorizable architectures generally require global optimization or design
  - optimizing ~ NxN variables at once
    - which makes them harder to program or self-configure
    - and makes real-time reconfiguration to different problems particularly hard
- Apparently, physically recirculating architectures
  - which are physically non-factorizable
    - are generally also algorithmically non-factorizable





Algorithmic factorizability

Algorithmically factorizable architectures are a subset of the physically factorizable architectures

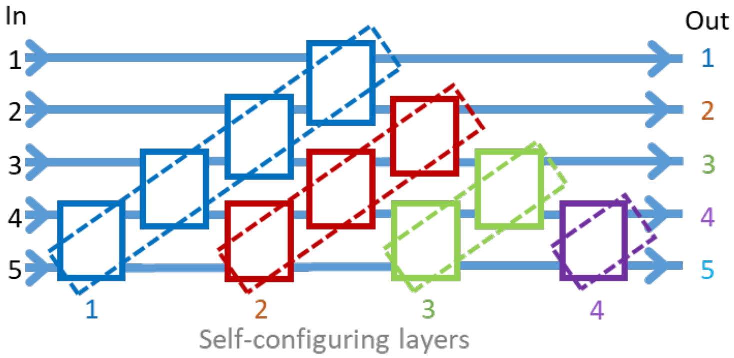
- and allow the algorithm to be "factorized" into progressive and successive operations

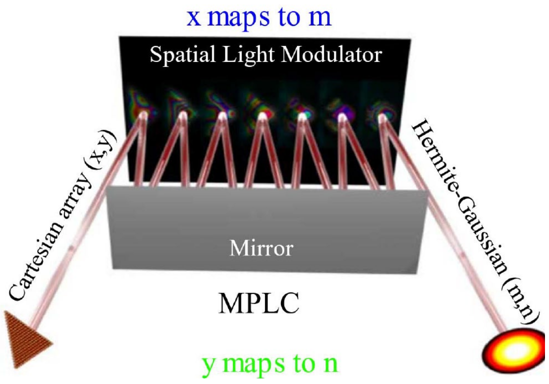
So, forward-only can be factorizable from a design point of view

- multiple successive "self-configuring" layers
  - each defined topologically

though non-(algorithmically) factorizable forward-only architectures also exist

- rectangular meshes
- multiplane light converters (MPLC)
  - which generally require global optimization to design them





Fontaine et al. "Photonic Lanterns, 3-D Waveguides, Multiplane Light Conversion, and Other Components That Enable Space-Division Multiplexing," Proc. IEEE **110**, 1821 (2022)

Algorithmic factorizability

Algorithmically factorizable architectures
- allow us to reduce to a succession of simpler designs or programmings
  - from global optimization, optimizing over ~NxN variables at once
    - which makes real-time programmability or adaptation hard
  - to, e.g., N successive designs, each of order ~ N,
    - or even NxN successive designs, each of order 1
      - that is, completely progressive single-parameter designs or configurations
- forward-only architectures may be a necessary condition for algorithmic factorizability
  - though there are forward-only architectures that do not factorize algorithmically

Apparently, physically recirculating architectures
- which are physically non-factorizable
  - are generally also algorithmically non-factorizable

Self-configuring architectures

Self-configuring architectures

- are factorizable both physically and algorithmically
- are defined topologically

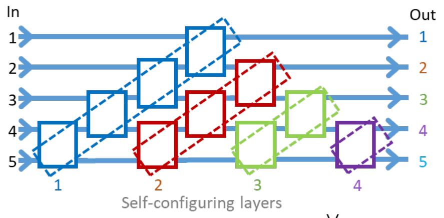
and are discoverable by topological algorithms

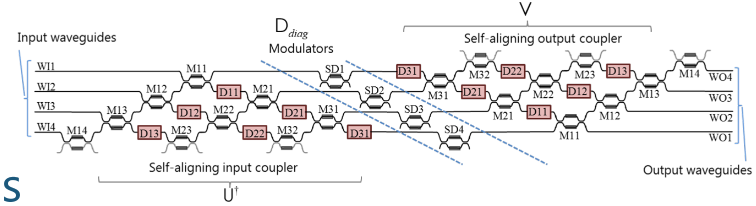
- can be completely universal linear optical systems at a given wavelength
- and so can be progressively designed and/or configured layer by layer

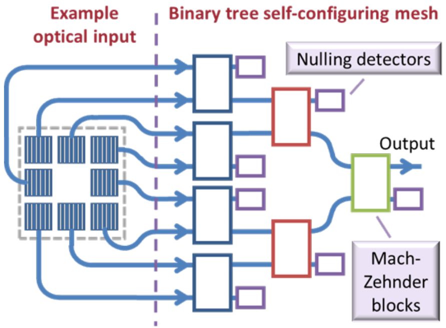
When working with coherent light

each layer can be configured progressively, device by device, with no calculations

giving a completely progressive, device by device, configuration for the entire network







"Analyzing and generating multimode optical fields using self-configuring networks," Optica 7, 794 (2020)

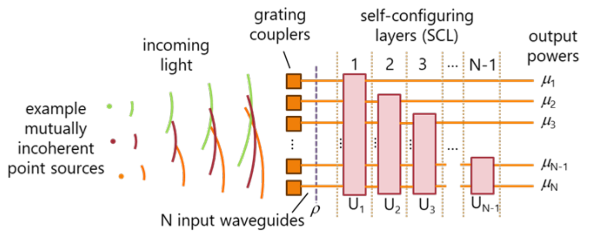
Self-configuring architectures

When working with incoherent light

- these can self-configure with global optimizations just within a layer
  - performing operations previously not apparently possible in optics
    - separating partially coherent light into its mutually orthogonal, mutually incoherent components

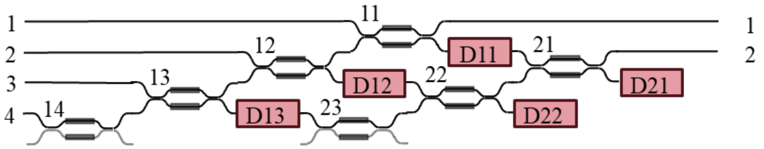
We can also have efficient architectures that find the first (and best) M vectors out of N, e.g., NxM mesh

- e.g., a self-configuring network with N inputs and M layers
  - which may map well onto real "sparse" problems
    - which is an example of a processor performing dimensionality reduction
      - that is, in an N dimensional space
        - only some relatively small number ~ M of orthogonal vectors may be important



Roques-Carmes et al., "Measuring, processing, and generating partially coherent light ..." LSA **13**, 260 (2024)

Roques-Carmes et al., "Automated Modal Analysis of Entanglement with Bipartite Self-Configuring Optics," ACS Photonics (2025) https://doi.org/10.1021/acsphotonics.5c00813



a "4x2" mesh separating 2 orthogonal beams from a 4-dimensional input

"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013); "Self-configuring universal linear optical component," Photon. Res. **1**, 1 (2013)

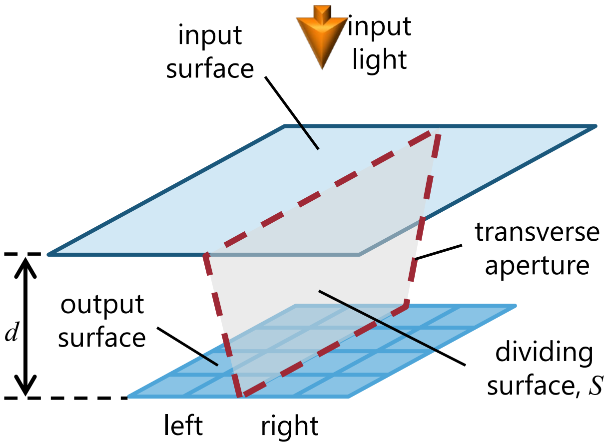
Why optics needs thickness

For metasurfaces and metastructures

- and for compact optics generally
  - we need to understand whether they need thickness
    - Can we make a given optical device in just one “layer”, for example?

Generally, no

- But why?



David Miller, "Why optics needs thickness," Science 379, 41 (2023)

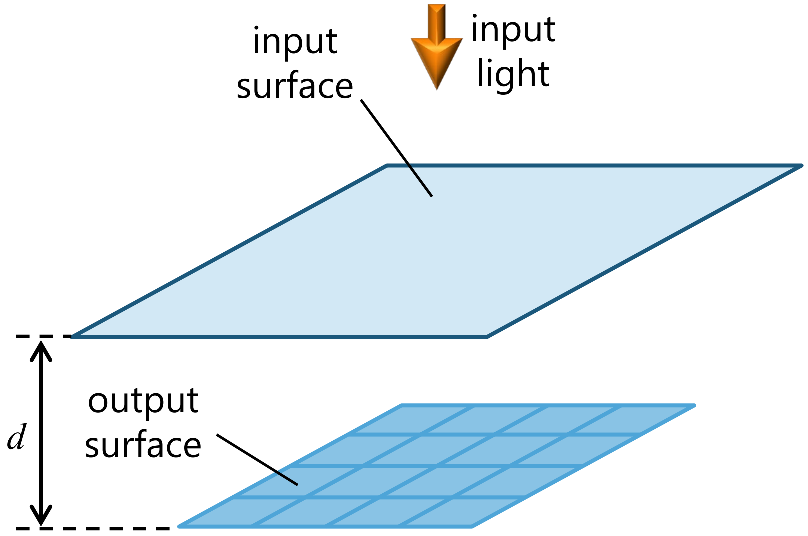
Why optics needs thickness

Think of an optical system with

- an input surface
  - such as a lens surface or metasurface
- an output surface
  - such as an image sensor plane
- with a distance *d* between them

Note we are not yet specifying what is between these two surfaces

- and we will not need to do so



"Why optics needs thickness," Science 379, 41 (2023)

The key idea – channels through a transverse aperture

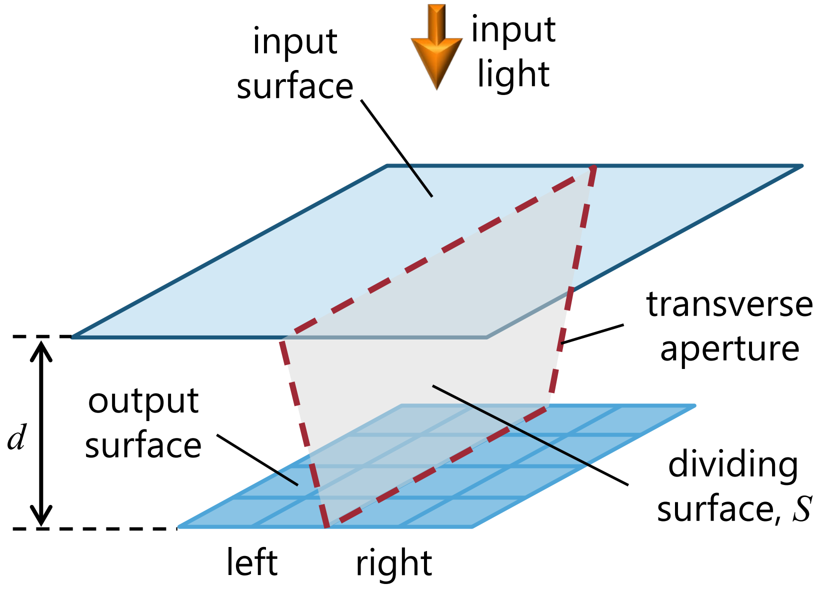
Now imagine we divide each surface in two parts
- left and right
  - by passing an imaginary mathematical dividing surface *S* through them

This defines a "**transverse aperture**"

Because of what we want the system to do
- some number *C* of channels must pass
  - from right to left (or left to right)
    - through this aperture

We call *C* the "**overlapping nonlocality**"

The transverse aperture must be large enough
- for these channels to propagate through it
  - which requires minimum area and/or thickness
    - e.g., half a wavelength thickness for each channel (in 1D problems)



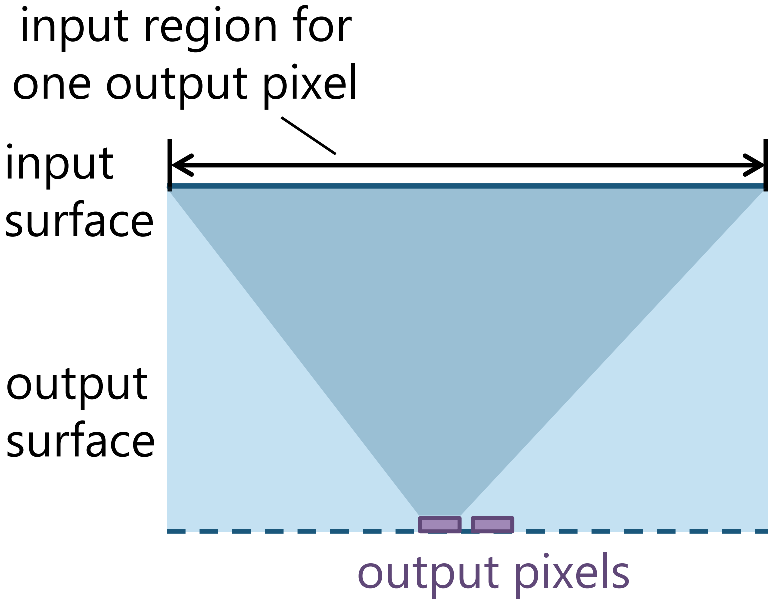
"Why optics needs thickness," Science 379, 41 (2023)

Nonlocality in optics

nonlocality

the output at one point depends on the input at many points

Imager example



For a general discussion of nonlocality, see Monticone et al., "Nonlocality in photonic materials and metamaterials: roadmap," Opt. Mater. Express **15**, 1544-1709 (2025)

Nonlocality in optics

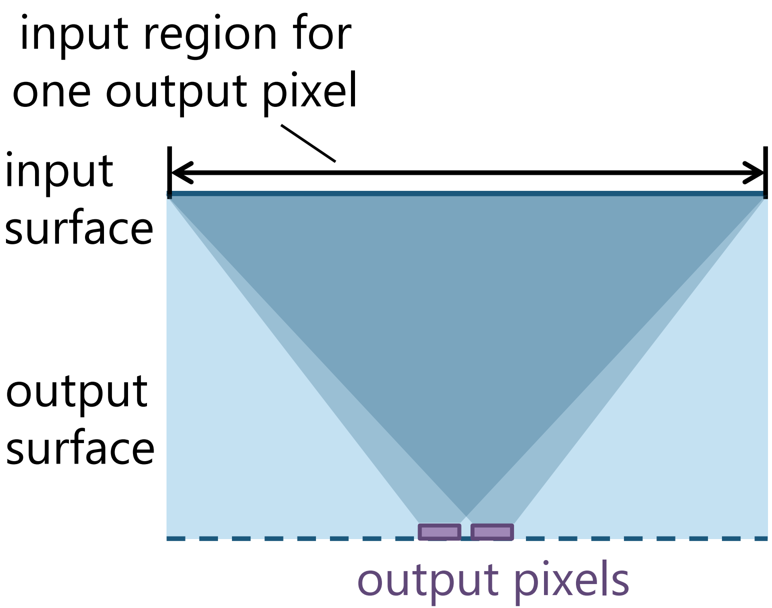
nonlocality

the output at one point depends on the input at many points

overlapping nonlocality

the input regions for different output points overlap with one another

Imager example



For a general discussion of nonlocality, see Monticone et al., "Nonlocality in photonic materials and metamaterials: roadmap," Opt. Mater. Express **15**, 1544-1709 (2025)

Nonlocality in optics

nonlocality

- the output at one point depends on the input at many points

overlapping nonlocality

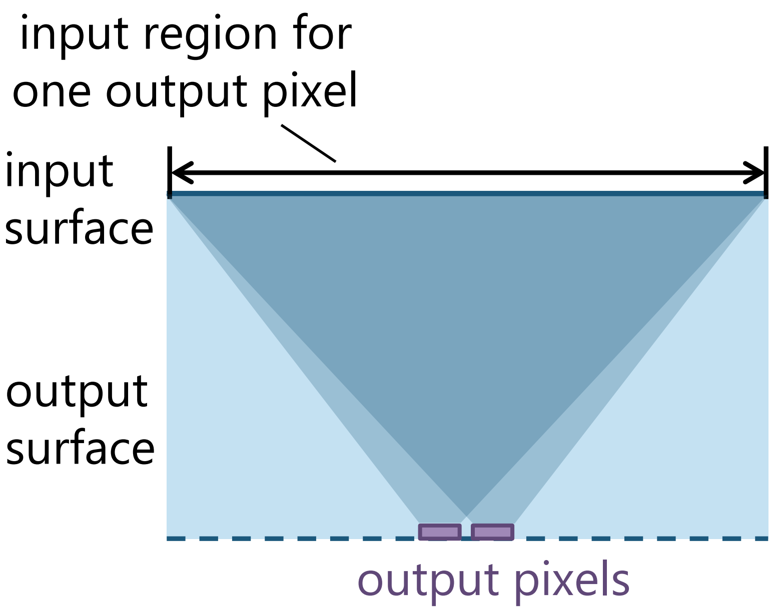
- the input regions for different output points overlap with one another

overlapping nonlocality C

- loosely, the number of such overlapping "channels"

For an imager, C ends up being half the number of pixels

Imager example



For a general discussion of nonlocality, see Monticone et al., "Nonlocality in photonic materials and metamaterials: roadmap," Opt. Mater. Express **15**, 1544-1709 (2025)

Nonlocality in optics

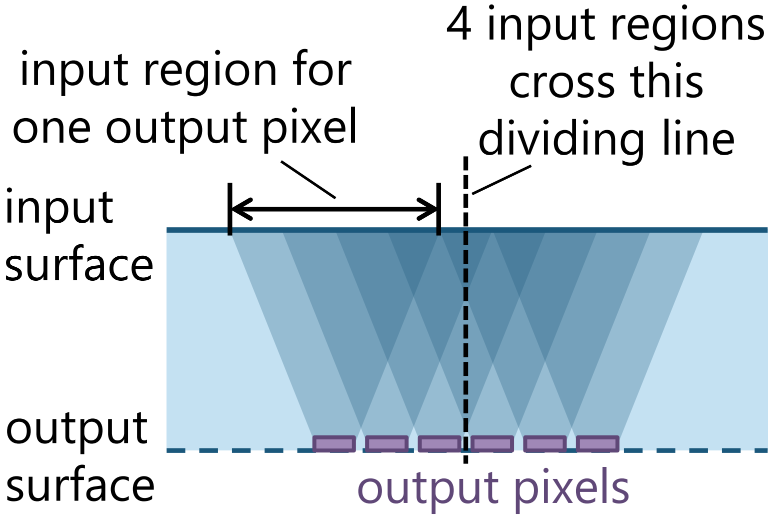
nonlocality
- the output at one point depends on the input at many points

overlapping nonlocality
- the input regions for different output points overlap with one another

overlapping nonlocality C
- loosely, the number of such overlapping "channels"

For this example, C is 4

Space-invariant example
e.g., image differentiator



$C = 4$

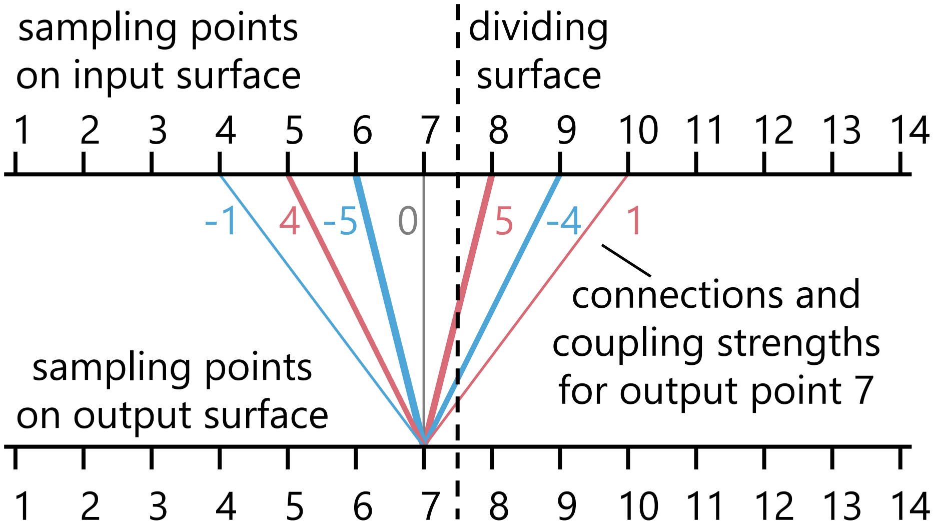
A pixelated differentiator

Consider a 5^{th} order finite difference derivative kernel

- formed from a
    - -1, 4, -5, 0, 5, -4, 1
- weighting of adjacent input points

In this case, we can set up a matrix D

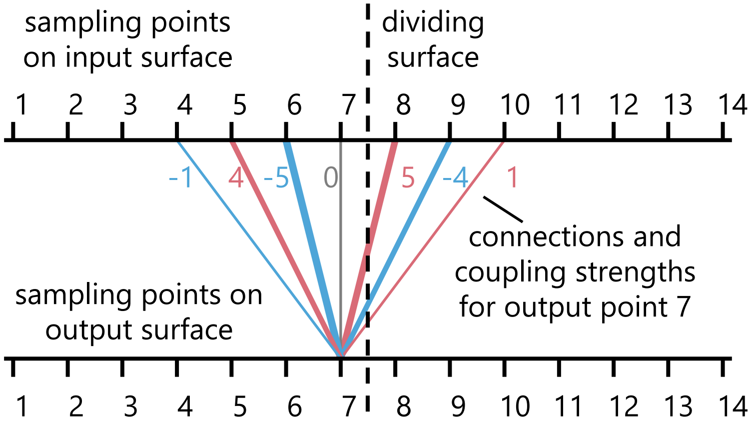
- which gives all the connection strengths between inputs and outputs
- for the full "space-invariant" kernel

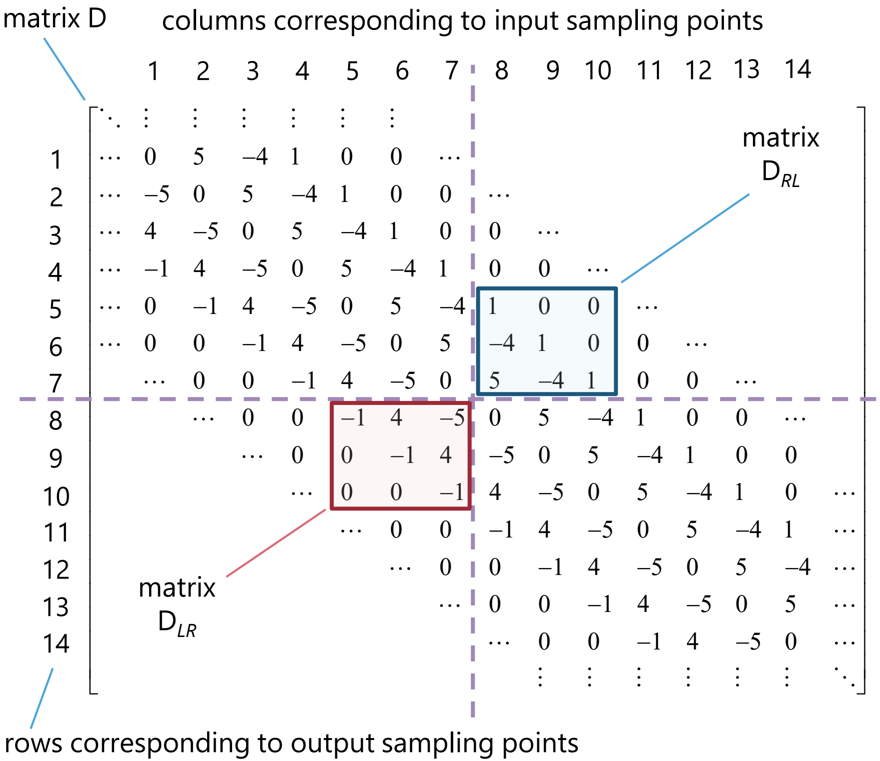


A pixelated differentiator

We can construct the full matrix D of the full “space-invariant” kernel

- arbitrarily choosing one vertical position for the dividing surface
  - between pixels 7 and 8

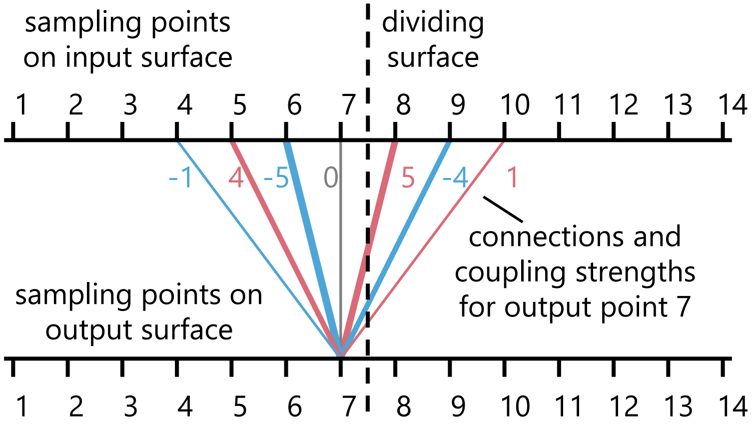


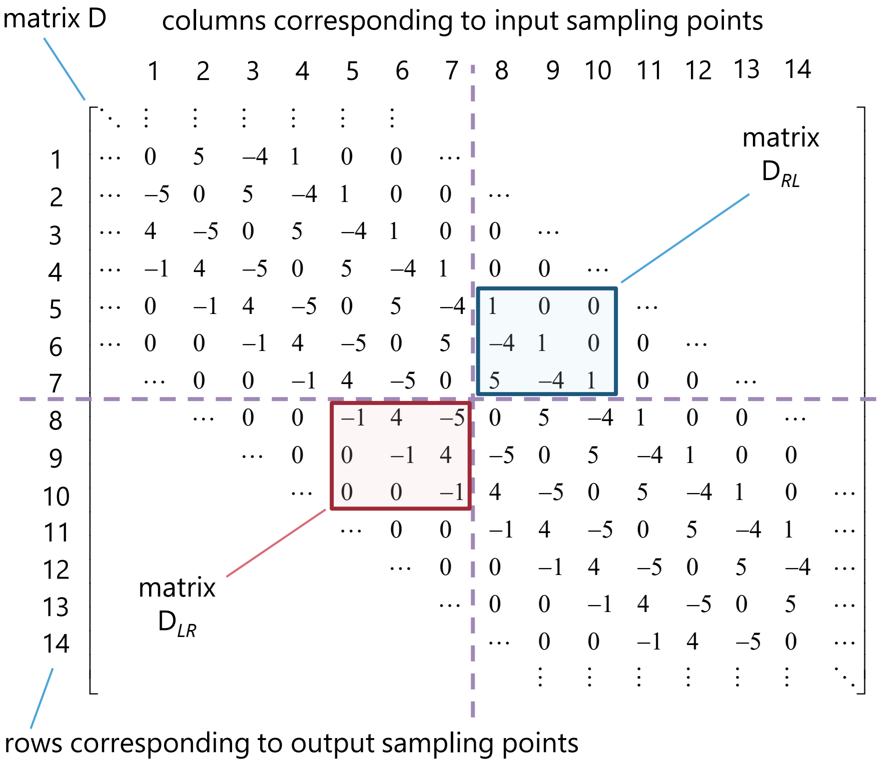


A pixelated differentiator

Sub-matrix D_{RL} gives all the connections
from the right inputs to the left outputs

Sub-matrix D_{LR} gives all the connections
from the left inputs to the right outputs





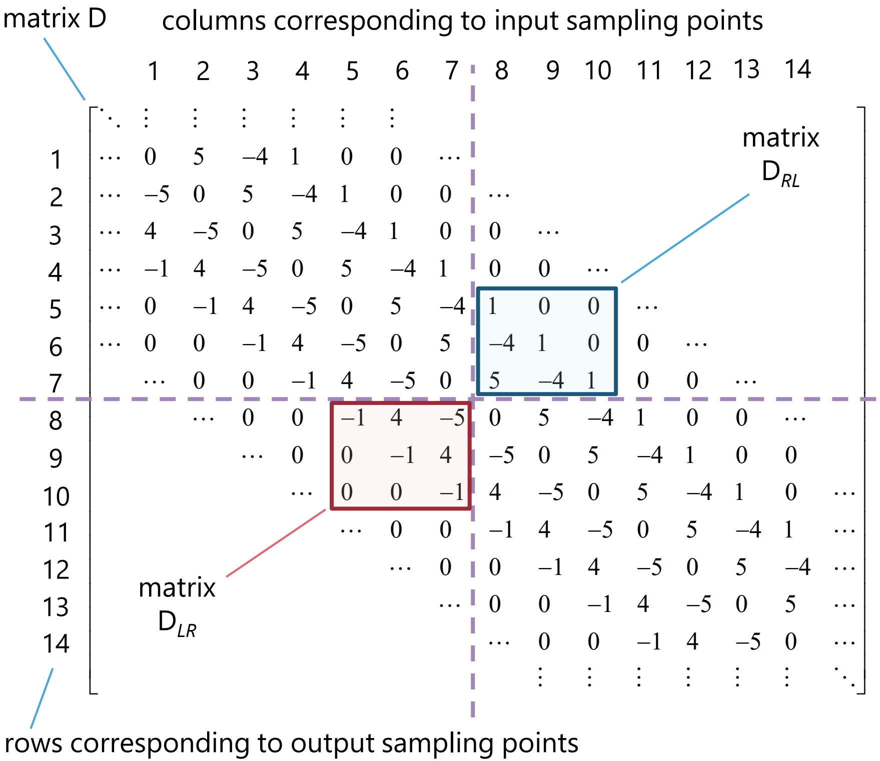
Singular-value decomposition approach

We can count directly as before
- deducing $C = 6$

But with these matrices
- we can take another formal approach -
  - singular value decomposition (SVD) of the matrices D_{RL} and D_{LR}

which gives C_{RL} and C_{LR} as the numbers of singular values of these matrices

Though we don't need this approach here
- we can use this approach for other problems where counting is not so clear



See "Waves, modes, communications, and optics: a tutorial," Adv. Opt. Photon. **11**, 679 (2019) for the SVD approach to optics

Local vs. non-local (overlapping nonlocality)

For the linear problem to be solved by our wave-based computer

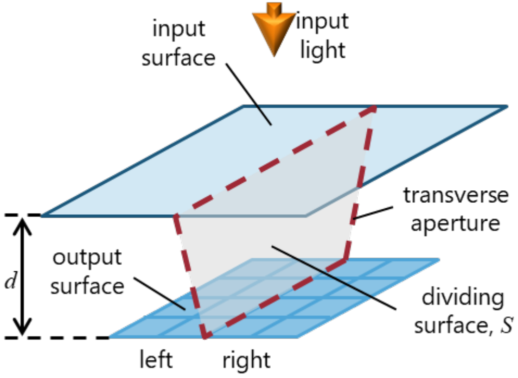
- we can deduce directly, by SVD
  - before starting design
  - what the overlapping non-locality of the problem is
    - which can tell us a minimum thickness for our wave-based computer
      - and tell us something about how we must construct it
        - including whether it may need multiple "layers"

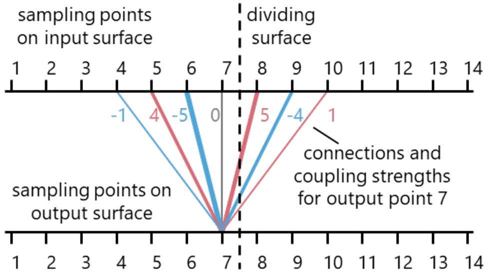
Note we need to go into real space to understand overlapping nonlocality

- we cannot stay in a k-space view of the problem
  - because we need to put the transverse aperture at some specific ("worst") point in space to understand the number of transverse channels we need
    - and hence thickness

So a key question is what is the overlapping non-locality of our problem

- because that sets the thickness we need in the wave processor

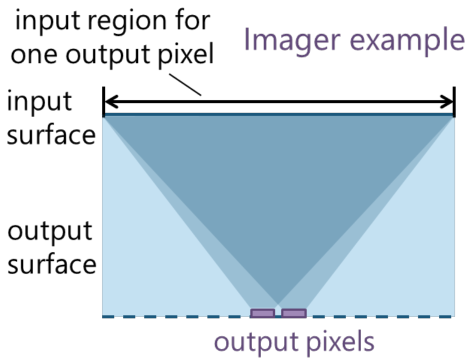


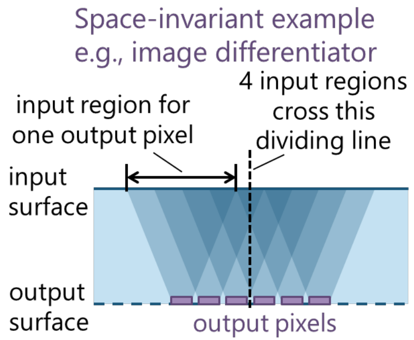


Space-variant vs. space-invariant

A related question is whether our problem is space-variant vs. space-invariant

- Space-invariant architectures are algorithmically much simpler than general space-variant
  - though there are some very simple space-variant architectures that are useful
    - e.g., lens





Circuits vs devices

Use circuits to make the system work

despite variations or imperfections in components

This is standard in electronics

Can we do this in optics?

Perfect optics from imperfect components

The beamsplitters in a Mach-Zehnder interferometer

- may not have a perfect 50:50 ratio when fabricated

Circuit solution

- use Mach-Zehnder interferometers as the beam splitters
  - and have an algorithm to set them automatically to function as 50:50 splitters
    - which is possible even if the fabricated beamsplitters are as bad as 85:15

Hence we can “perfect” the device automatically

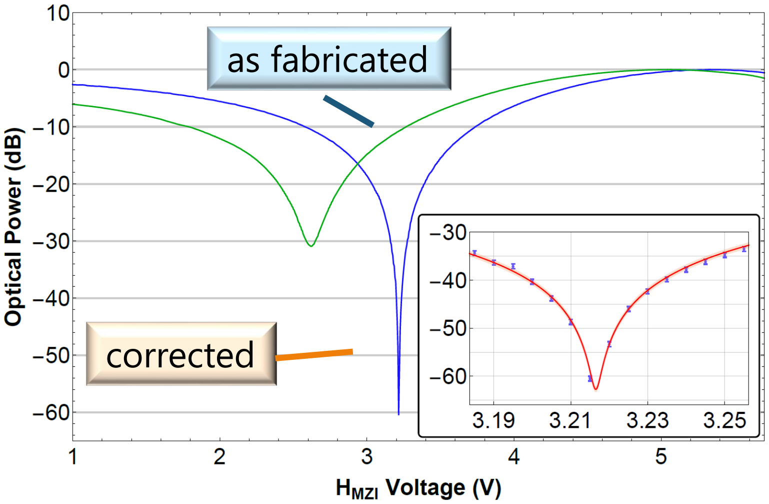
- e.g., improving the rejection ratio from -30 dB to -60 dB

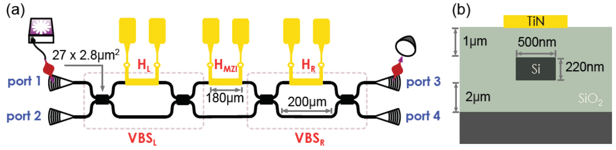
No calibration required

No calculations

Based only on

- power minimization or maximization
  - in an output detector





"Perfect optics with imperfect components," Optica **2**, 747-750 (2015); Wilkes et al., "60 dB high-extinction auto-configured Mach–Zehnder interferometer," Opt. Lett. **41**, 5318-5321 (2016)

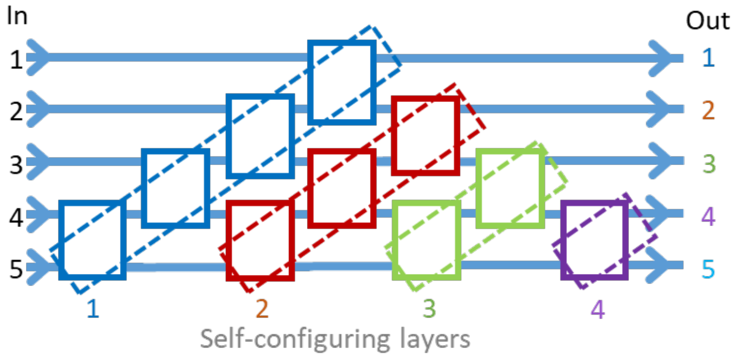
Standard blocks vs. custom designs

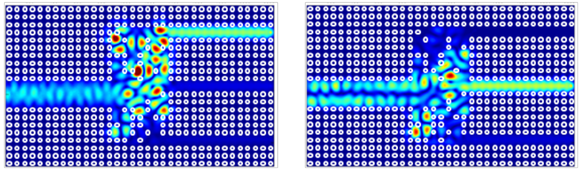
Standardized designs or design blocks vs. full custom design

- Electronics uses standardized design blocks extensively - PDKs
  - Allows design of much more complex systems because it supports abstractions.
  - Allows manufacture of much more complex systems because we standardize manufacture
  - Allows portability of designs to different manufacturers

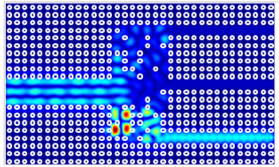
So, should we move to standardized designs based on blocks for wave-based computing processing?

- Again, the interferometer meshes give us an example of universal processors built from standardized “2x2” blocks
- The topologies and algorithms stay the same even if we change the physical implementation of the 2x2 blocks



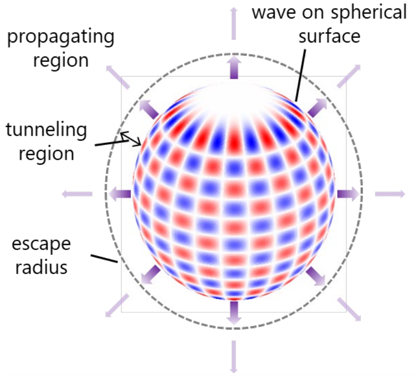


Engineer precise mode splitting with positioning of dielectric columns



Bound on the number of wave channels in or out of a volume

- Why it's so hard to beat the diffraction limit
  - Complicated waves must tunnel to get in or out of small volumes
- Rigorous theory of spherical waves
  - shows a previously "hidden" radial tunneling phenomenon
- If the wave is too complicated
  - i.e., relies on spherical wave components with too many "bumps"
    - it has to tunnel to get in or out
- Explains why we have relatively strong wave couplings in or out of a volume
  - up to a relatively sharp cutoff
    - corresponding to the onset of tunneling
- Helps understand overall size requirements in wave-based computing



D. A. B. Miller, Z. Kuang, O. D. Miller, "Tunneling escape of waves," Nat. Photon. **19**, 284–290 (2025)

Wave-based computing

What are the technical barriers to wave-based computing?

- what are the 4 or 5 most important
  - An approach to viable design and manufacture of complex structures
    - e.g., layered, complex, forward-only metasurfaces
    - e.g., miniaturization of interferometric mesh components and structures
      - e.g., inverse designed beamsplitters, couplers,
      - wavelength-independent designs of couplers and phase shifters
      - micromechanical adjustable components for very low power control of components
  - Factorizable design approaches of complex structures for designability, programmability, stabilization, self-configuration
    - e.g., forward-only complex structures
  - Understanding of applications of wave-based computing
    - many or even most of those may lie beyond fixed processors

stanford.io/4rdZDSJ

Wave-based computing

- It's not sufficient just to have an idea for an ideal wave-based computing system
  - Since it is a complicated analog system
    - we have to have a strategy for how we are actually going to get it to work and to continue to work
      - including imperfections in manufacture and design and variations in the problem being solved
    - That strategy may have to involve some of the ideas presented here
      - such as physical and algorithmic factorizability, programmability, self-configuration, self-stabilization, circuits to counter imperfect components, standard design blocks, and capability of working with a variety of problems
  - Equivalently, how are we going to control it
    - because we are going to have to control it
      - Note in particular the challenge of controlling any resonant effects we use in our system
- No designs based on materials and processes that don't exist or that rely on extreme physics
  - The system and application ideas should show promise
    - with materials and technologies that already exist
      - or that could be created with finite development and operating in realistic conditions

stanford.io/4rdZDSJ

Wave-based computing

What would it take to achieve them?

Significant research work towards understanding how to make viable large analog systems

that could solve problems people would care about

which includes understanding what real problem areas will ultimately require

How do we measure success or progress?

People beating a path to our door!

Equivalently, someone outside our research community has to care!

stanford.io/4rdZDSJ

Supported by Air Force Office of Scientific Research
FA9550-17-1-0002 and FA9550-21-1-0312

Supplementary slides

stanford.io/4rdZDSJ

Why not to make optical transistors

- The speed of electronics is not limited by transistors
  - It is limited by interconnect power and density
    - and avoiding melting the chip by running too many gates too fast
      - even with the low energies of CMOS logic
- Main reasons against optical transistors
  - surprisingly difficult to satisfy the necessary criteria for cascadable logic gates
  - they essentially all take way too much energy
- Biggest failure of my professional career
  - The paper written to persuade mostly not to work on optical transistors

    "Are optical transistors the next logical step?"
    Nature Photonics **4**, 3 (2010)

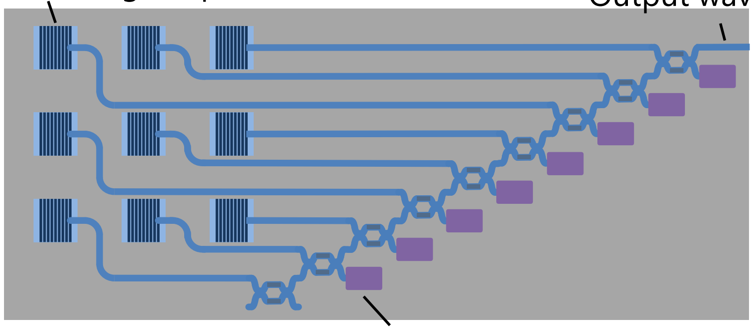
    - has been cited over 600 times
      - probably mostly by people trying to make optical transistors

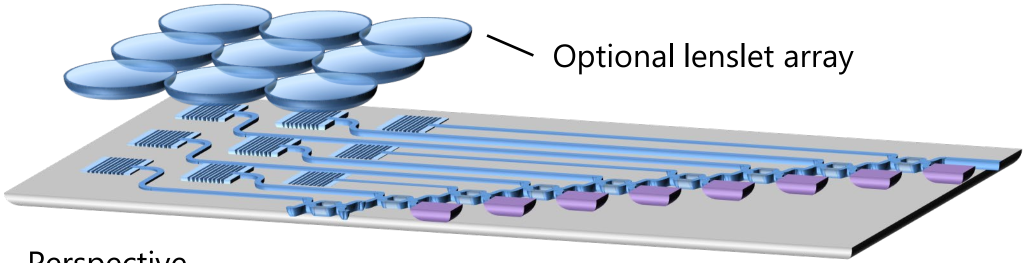
stanford.io/4rdZDSJ

Self-aligning beam coupler

This has several different uses

- ❑ tracking an input source
  both in angle and focusing
- ❑ correcting for aberrations
- ❑ analyzing amplitude and phase of the components of a beam
  - ❑ …

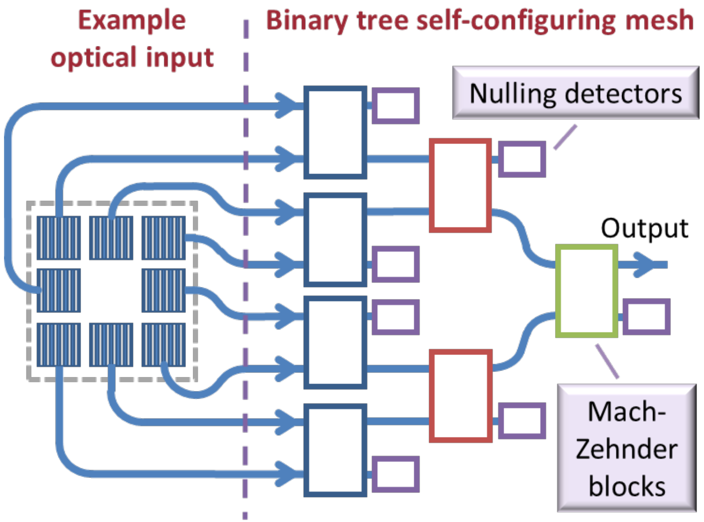




"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Measuring and generating arbitrary beams

- Self-configuring this “binary tree” layer to route all power to the output
  - automatically measures the relative amplitudes and phases of the input light
    - with the values deduced from the resulting mesh settings.
- Run backwards, it can generate any beam emerging from the “inputs”
  - generation of arbitrary beams
  - reference-free measurement of arbitrary beams



“Analyzing and generating multimode optical fields using self-configuring networks,” Optica 7, 794 (2020)

See also J. Bütow et al. "Spatially resolving amplitude and phase of light with a reconfigurable photonic integrated circuit," Optica 9, 939 (2022)

Optically separating exoplanets

Finding exoplanets around distant stars is optically very challenging

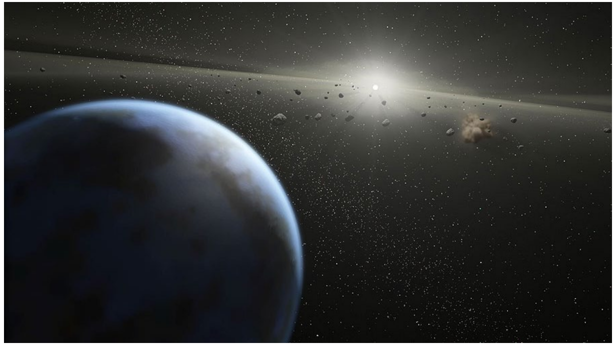
- the star may be 10^{10} times brighter than the planet
  - and the planet may lie in the weak wings of the star's diffraction pattern in the telescope

Interferometer meshes may allow

- optimized modal filtering
  - to suppress the star "modes"
    - to improve the rejection of the star light

Preliminary experiments with meshes are already showing ~ 90dB rejection

Dan Sirbu et al., "AstroPIC: near-infrared photonic integrated circuit coronagraph architecture for the Habitable Worlds Observatory," Proc. SPIE 13092, 130921T (2024)



Use a programmable photonic mesh to provide optimal modal filtering to reject star light and pass possible exoplanet light

Separating partially coherent light

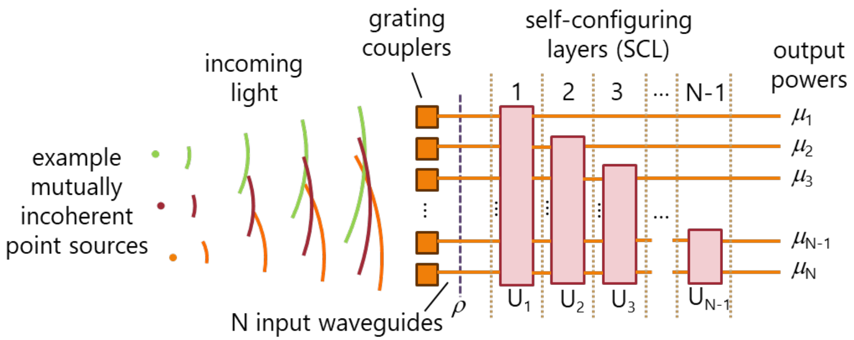
With partially coherent input light

- by power maximizing on the successive self-configuring layers
  - this circuit can measure the coherency matrix of that light
    - simultaneously separating it into its
      - mutually incoherent and mutually orthogonal components

No other known apparatus can apparently perform this separation

This concept can also be extended to

- measure the single photon density matrix
- automatically perform a modal analysis of entanglement with two-mesh bipartite self-configuring optics



Roques-Carmes et al., "Measuring, processing, and generating partially coherent light …" LSA **13**, 260 (2024)

C. Roques-Carmes, A. Karnieli, D. A. B. Miller, and S. Fan, "Automated Modal Analysis of Entanglement with Bipartite Self-Configuring Optics," ACS Photonics (2025) https://doi.org/10.1021/acsphotonics.5c00813

Example – metastructure for smoothed derivative

Wang et al. designed a "thick" 2D photonic crystal to perform a smoothed ("Gaussian") derivative with kernel

$$D(u;x) = \frac{(x-u)}{\beta} \exp\left(-\frac{(x-u)^2}{\beta^2 \Delta_t^2} \right)$$

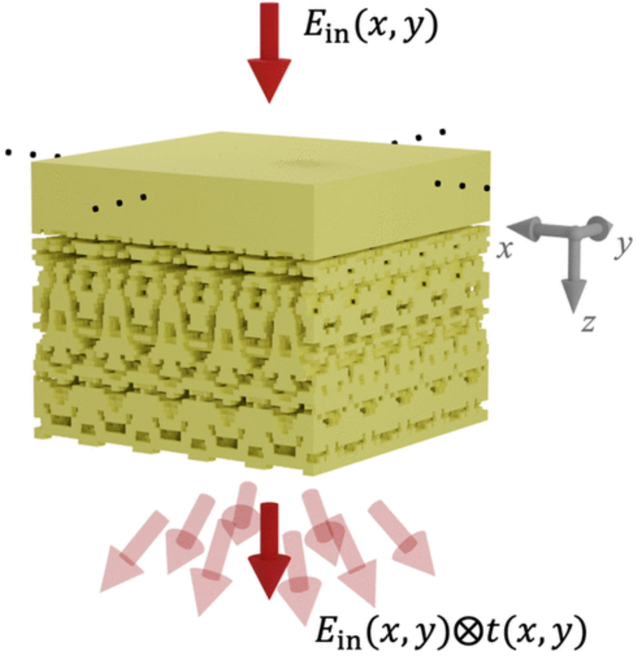
The "divided" kernel has ~ 6 significant singular values

so we should need ~ 6 physical channels through the "transverse aperture"

The thickness of the actual designed structure is ~ 6 wavelengths thick

so more than thick enough at half a wavelength thickness per channel

obeying the proposed (1D) limit here



H. Wang, W. Jin, C. Guo, N. Zhao, S. P. Rodrigues, S. Fan, ACS Photonics 9, 1358–1365 (2022)

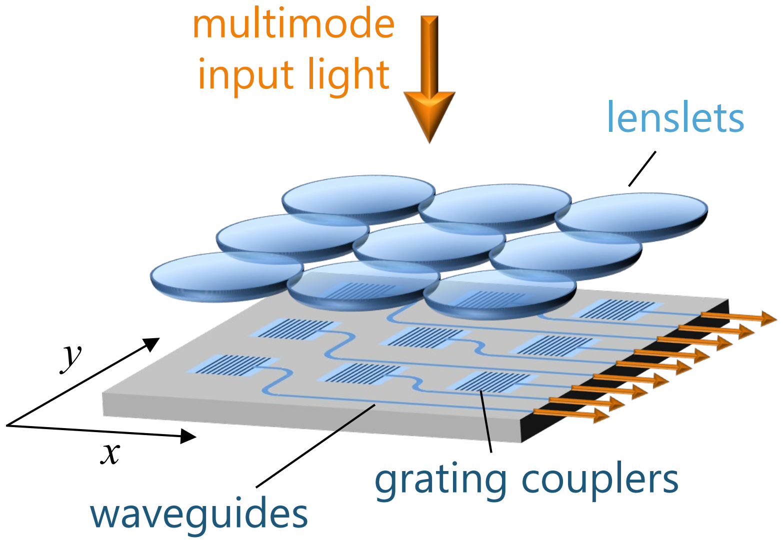
Converting from 2D to 1D -dimensional interleaving

Can we just "interleave" the channels

taking degrees of freedom that were in x

and interleave them into y?

In principle, yes – the "supercoupler" does this



"supercoupler"

converts 2D input modes

to output modes in a 1D line

e.g., in waveguides

Converting from 2D to 1D -dimensional interleaving

Can we just "interleave" the channels

taking degrees of freedom that were in x

and interleave them into y?

In practice, this "dimensional interleaving" is much harder

None of the following appear to support dimensional interleaving

- free-space propagation
- conventional imaging systems
- simple dielectric stack structures
- 2-D photonic crystals

Question: is dimensional interleaving possible with continuous optics?

Universal matrix multiplier chip

Full complex matrix multiplication

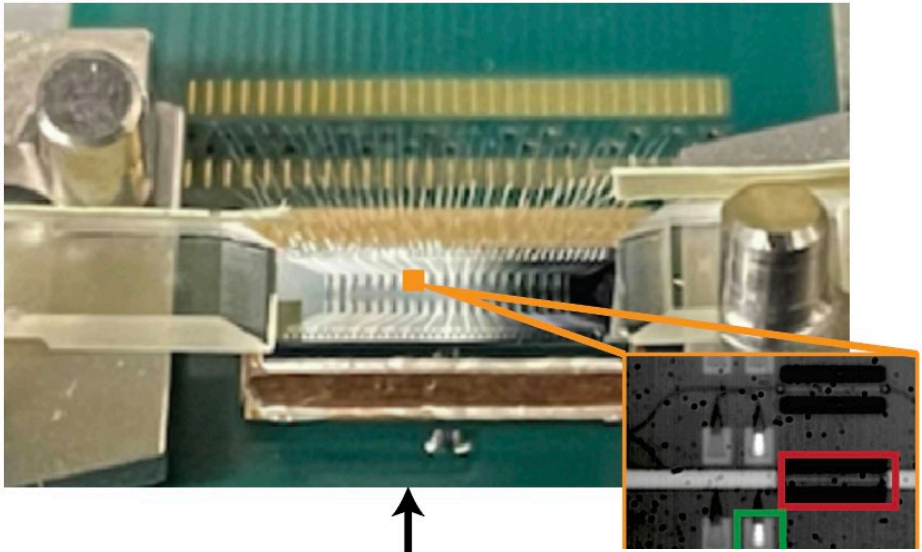
with vector generation and vector analysis

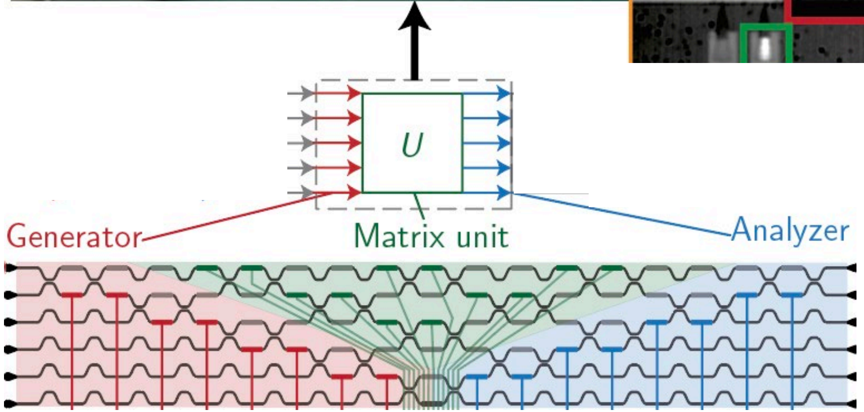
Photonic back-propagation neural net training

S. Pai, Z. Sun, T. W. Hughes, T. Park, B. Bartlett, I. A. D. Williamson, M. Minkov, M. Milanizadeh, N. Abebe, F. Morichetti, A Melloni, S. Fan, O. Solgaard, D. A. B. Miller, "Experimentally realized in situ backpropagation for deep learning in photonic neural networks," **Science** 380, 398-404 (2023)

Digital matrix multiplication for cryptography

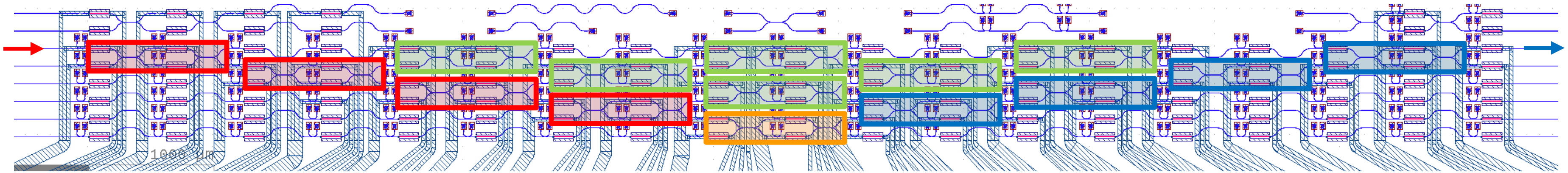
S. Pai, T. Park, M. Ball, B. Penkovsky, M. Dubrovsky, N. Abebe, M. Milanizadeh, F. Morichetti, A. Melloni, S. Fan, O. Solgaard, and D. A. B. Miller, "Experimental evaluation of digitally verifiable photonic computing for blockchain and cryptocurrency," **Optica** 10, 552-560 (2023)





Mask layout and block diagram

Matrix unit



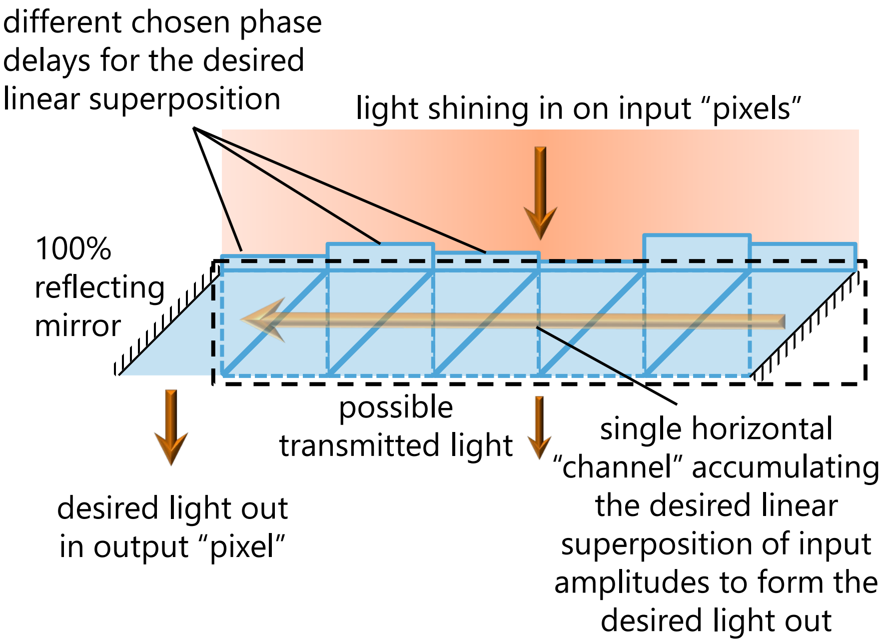
Vector generator

Vector analyzer

Nonlocality in optics

A system of beamsplitters collects possibly all the light from 6 different input regions

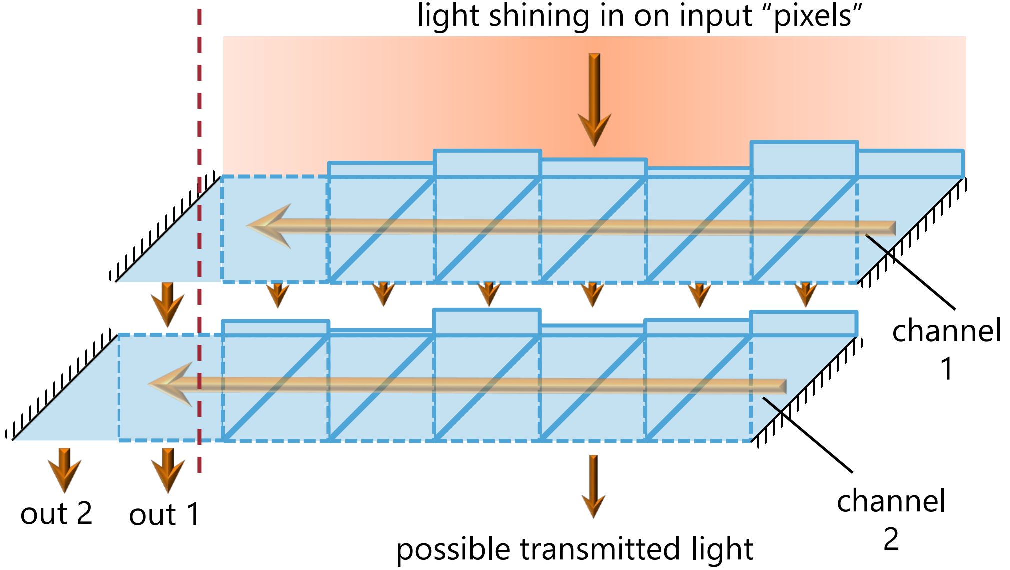
- so, with a “nonlocality” of 6
- to only one output “pixel”
  - at the extreme left
    - so, with no overlap in the nonlocality
      - i.e., $C = 1$ “channels”



Nonlocality in optics

Two rows of beamsplitter blocks

- collect two orthogonal 6-element light beams
- into two separate outputs
- with an **overlapping nonlocality** $C = 2$



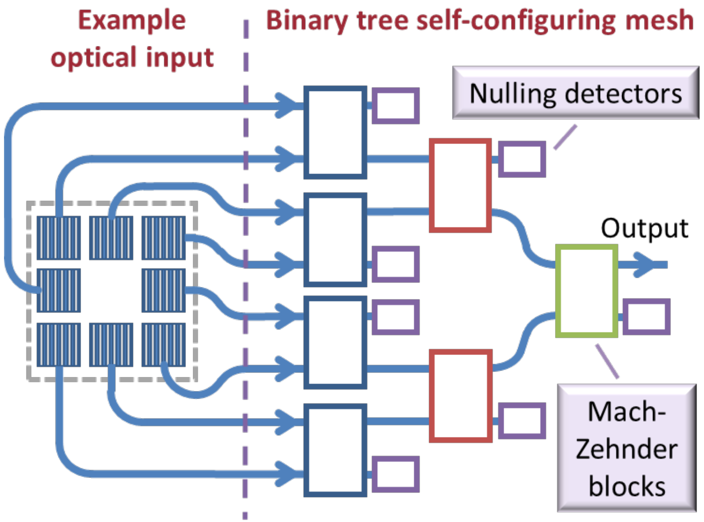
Measuring and generating arbitrary beams

Self-configuring this "binary tree" layer to route all power to the output

- automatically measures the relative amplitudes and phases of the input light
  - with the values deduced from the resulting mesh settings.

Run backwards, it can generate any beam emerging from the "inputs"

- generation of arbitrary beams
- reference-free measurement of arbitrary beams



"Analyzing and generating multimode optical fields using self-configuring networks," Optica 7, 794 (2020)

See also J. Bütow et al. "Spatially resolving amplitude and phase of light with a reconfigurable photonic integrated circuit," Optica 9, 939 (2022)