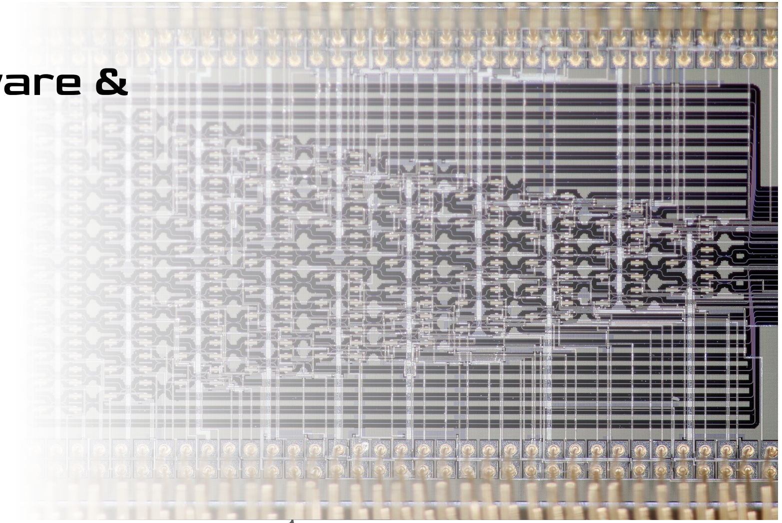
MIT AI Hardware Program

QP & AI
http://qp.mit.edu



Co-design of machine intelligence and hardware & the need for verifiable AI

MIT AI Hardware Program Symposium | 2025.11.12

Dirk Englund ∈ {RLE, MTL, EECS}@MIT

Quantum Computing & Networking

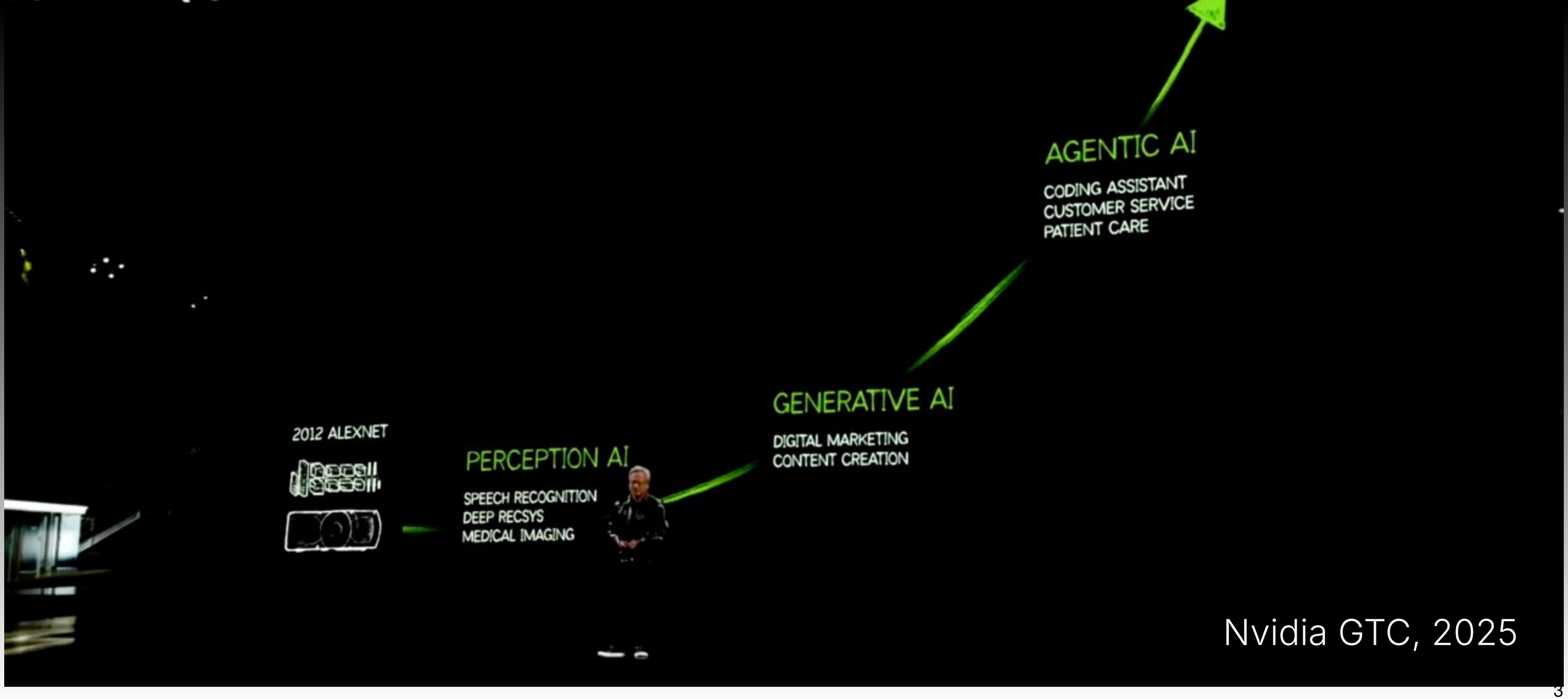
ML signal processing

A Renaissance in Specialized Computing

1. **Quantum Computing**

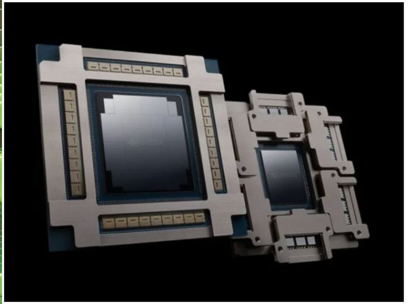
2. **AI: perception, generative, agentic**

3. **Physical AI**



Nvidia GTC, 2025

Photonics at the forefront



Nvidia looks to silicon photonics to cut datacentre AI power

Technology News | March 21, 2025

By Nick Flaherty

OPTOELECTRONICS AI

POWER MANAGEMENT

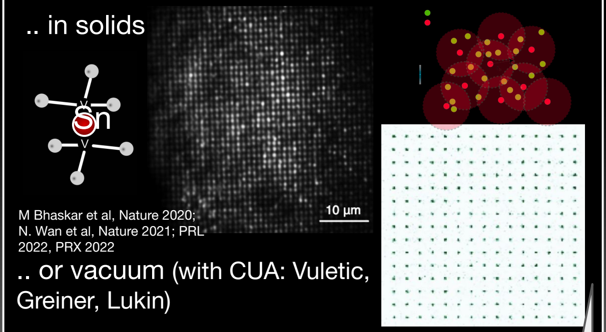
Nvidia GTC, 2025

MIT research themes

MIT QP & AI http://qp.mit.edu

Scalable qubit systems (10^3-10^6)

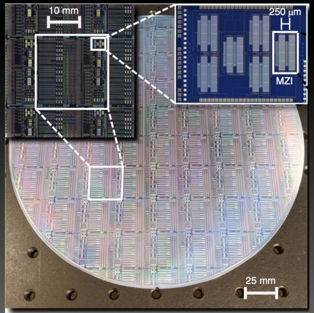
.. in solids



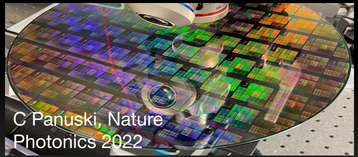
M Bhaskar et al, Nature 2020; N. Wan et al, Nature 2021; PRL 2022, PRX 2022

.. or vacuum (with CUA: Vuletic, Greiner, Lukin)

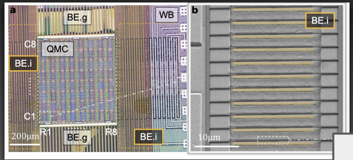
Semiconductor devices



M Dong et al, Nature Photonics 2021

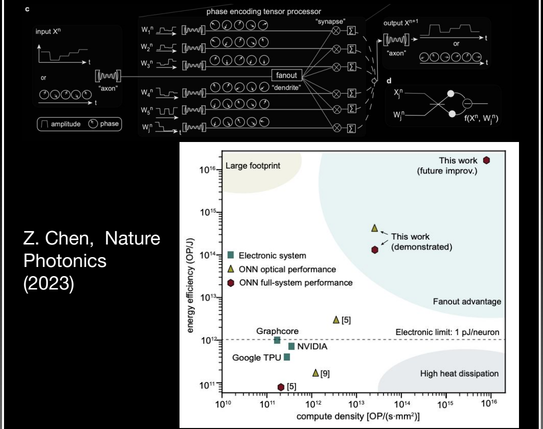


C Panuski, Nature Photonics 2022



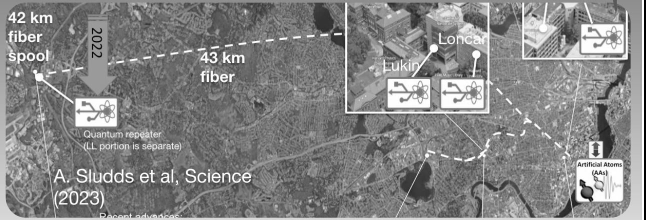
L Li et al (with Prof Ruonan Han)

AI / machine learning



Z. Chen, Nature Photonics (2023)

Full systems: quantum networks, AI, quantum sensors



A. Sludds et al, Science (2023)

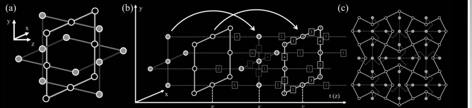
with Manya Ghobadi (CSAIL), MIT LL Grp 67, and Karl Berggren (EECS)

Theory & algorithms: physics, control theory, CS

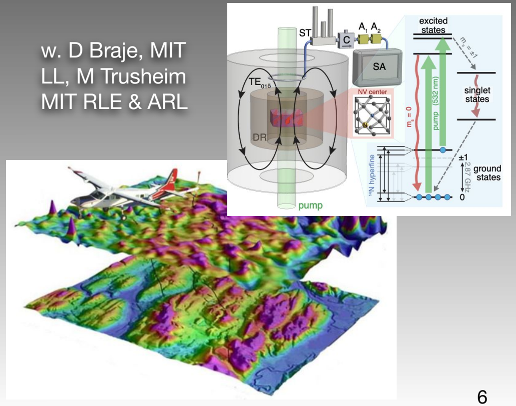
QPG-MIT

qp-slm

$$H = \frac{\hbar\Omega_{\mu w}}{2}\sum_i \sigma_x^i - \frac{\hbar\delta_{\mu w}}{2}\sum_i \sigma_z^i + \sum_{i\neq j}\frac{C_3}{R_{ij}^3}\left(\sigma_+^i\sigma_-^j + \sigma_-^i\sigma_+^j\right)$$



w. D Braje, MIT LL, M Trusheim MIT RLE & ARL



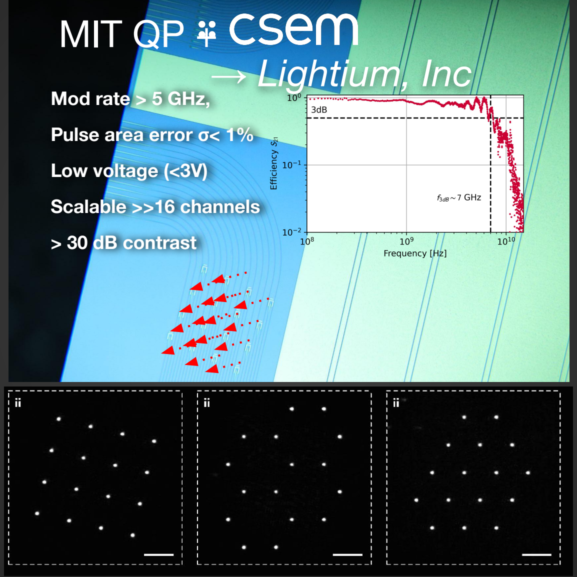
PIC technology

M Saha, Y. Henry Wen, et al, arXiv https://arxiv.org/abs/2406.17662 (2025)

See also
M Dong et. al, Nature Photonics (2021)
A Menssen et. al. Optica (2023)

- MIT-Sandia National Laboratory - MITRE - Harvard
Separately: QuEra - SNL
- Process pioneered by Matt Eichenfield et al at Sandia NL
Spin-off company in the works on PIC displays & sensors - Interested ?

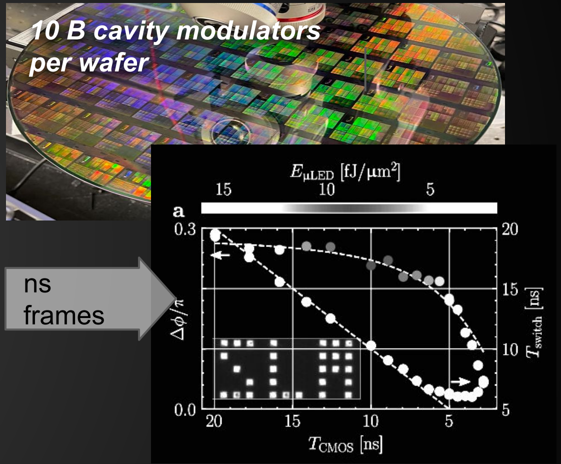
16-CH APIC lithium niobate on insulator[3]



16 programmable beams, 10 GHz bandwidth

[3]I. Christen et al, Nature Comm (2025)
H. Larocque, D. Vitullo, et al,

2D Spatial Light Modulator in Si demonstrator (v_{mod}~1GHz) [4]



[4] C. Panuski et al, Nature Photonics (2022)
[CP awarded the 2022 Carl E. Anderson Division of Laser Science Dissertation Award]

48-ch EOMs (v_{mod}>30 GHz):
A. Sludds et al, *Science 378* (2022).

Nanophotonic waveguide chip-to-world beam scanning at 68M Spots/s·mm²

Matt Saha[*,1], Y. Henry Wen,[*,1,2,α] Andrew S. Greenspon[*,1,2], Matthew Zimmermann[*,1], Kevin J. Palm[1,2], Alex Witte[1], Yin Min Goh[2], Chao Li[2], Jonathan Bumstead[3], Kevin Schädler[4], Mark Dong[1,2], Andrew J. Leenheer[6], Genevieve Clark[1,2], Gerald Gilbert[5,‡], Matt Eichenfield[6,7,&], and Dirk Englund[2,4†]

M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al,
https://arxiv.org/pdf/2406.17662

The Scanning Waveguide Micro-Display for Augmented Reality

MITRE team,

PoC: Dr Henry Wen, MITRE

Henry Wen <hwen@mitre.org>

The challenge:

3000 nits

12hr batt

~1 cc
~1 gram

4K ~8 MP

A bright, efficient, dime-sized, high-res display for each eye, in the frame of light-weight glasses

Case study: Meta Orion

Most advanced research prototype

Result of >$10B R&D

Not mass producible

>$10,000 per unit cost

Large Field-of-View (70°)👍

Low resolution (13 PPD→0.39 MPix)👎

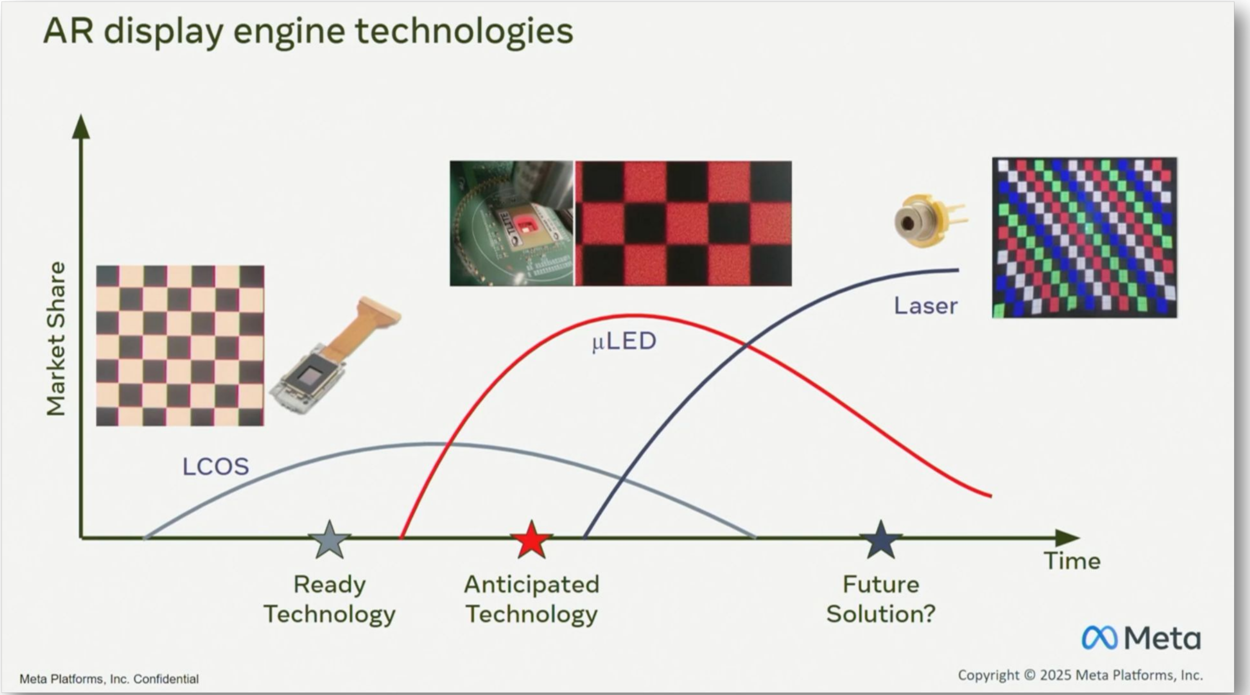
Low Brightness (400 nits→indoor only)👎

Short battery (2-3hrs)👎

Requires: →8 MP for immersive

→3000 nits for outdoor

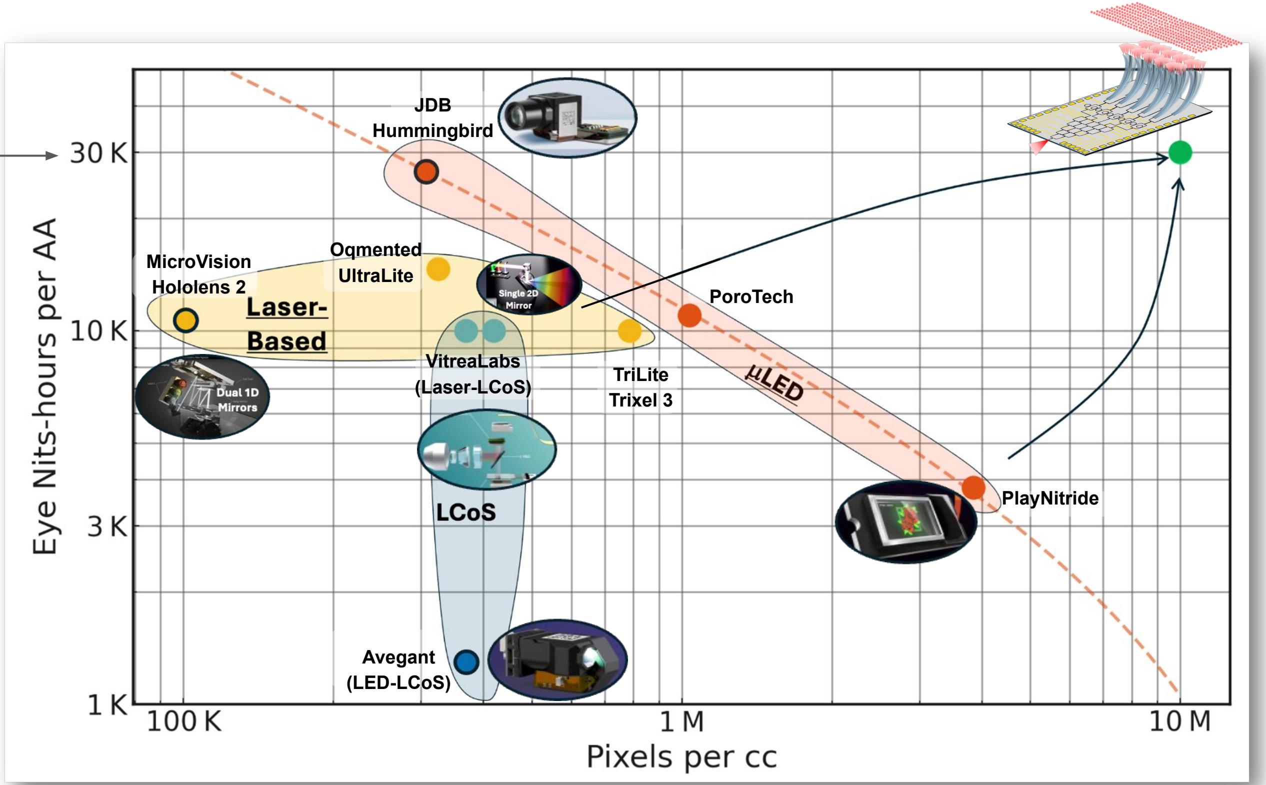
Bright + efficient requires lasers



State of art

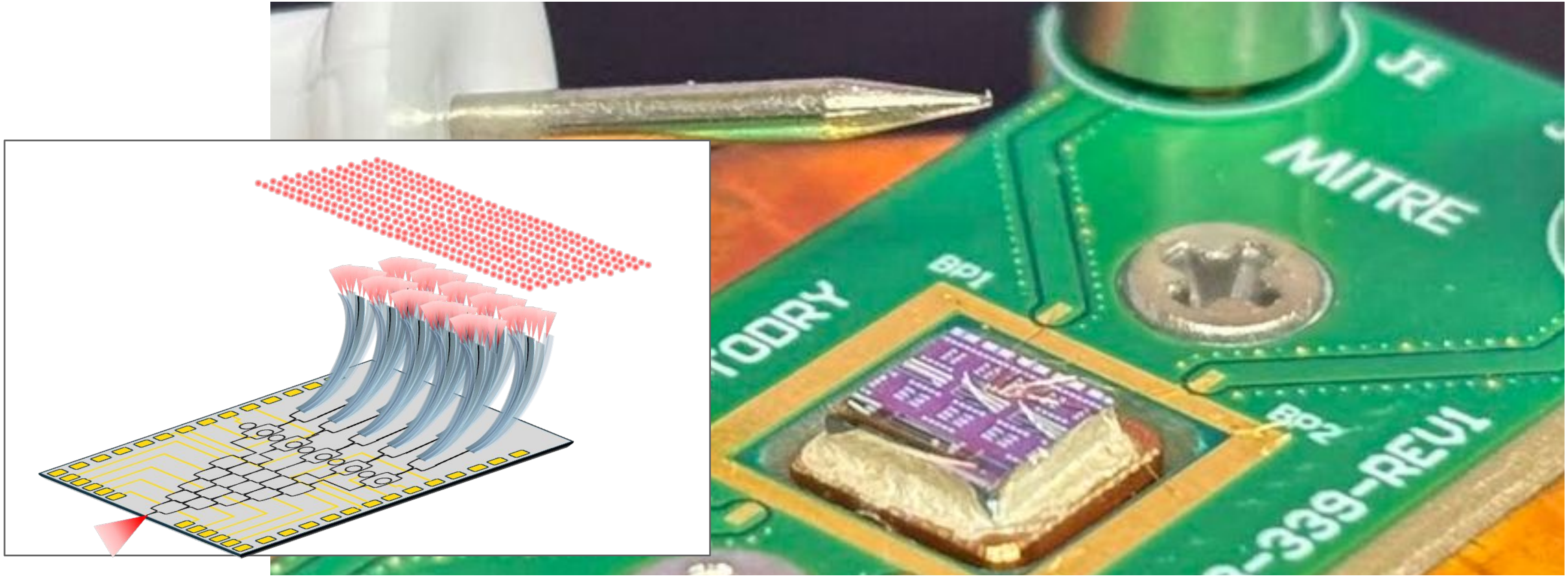
10hrs @ 3000 nits

How bright and how long it can last



How many pixels for a given volume

Nanoelectromechanical PIC display



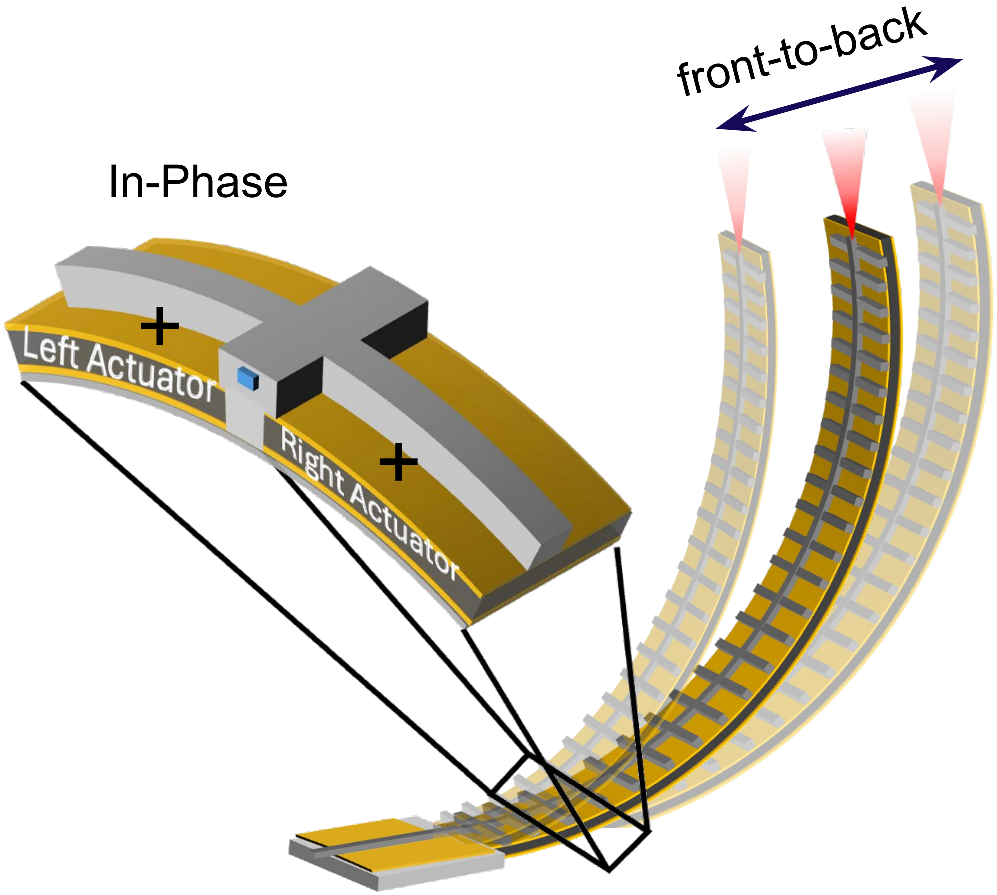
Scan

Low voltage (<5V) requirement enables CMOS compatibility

Scan frequency > 4 KHz

<1 um pixel pitch

6.8M pixels per second per scanner



M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al, https://arxiv.org/pdf/2406.17662

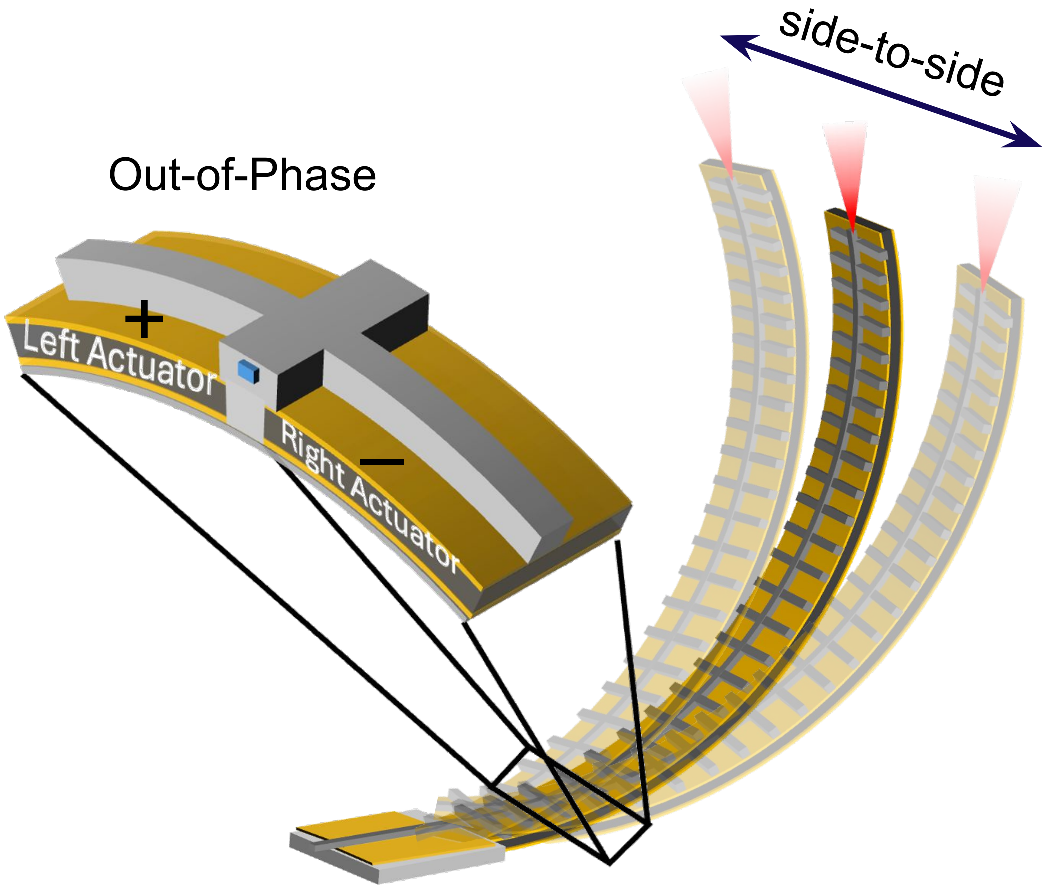
Scan

Low voltage (<5V) requirement enables CMOS compatibility

Scan frequency > 8 KHz

<1 um pixel pitch

6.8M pixels per second per scanner



M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al, https://arxiv.org/pdf/2406.17662

TECH DEMONSTRATION

1 FRONT-BACK

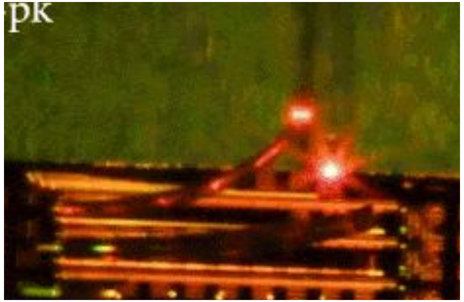
Front-to-back movement

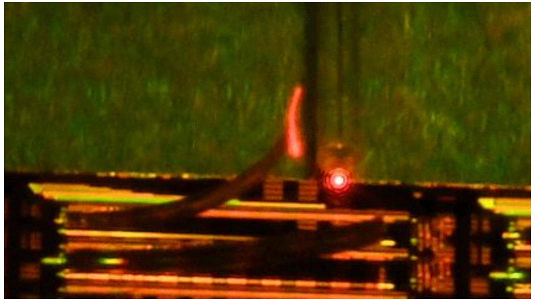
2 SIDE-TO-SIDE

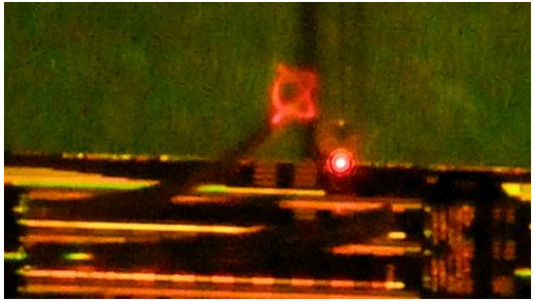
Side-to-side movement

3 COMBINED

Arbitrary control

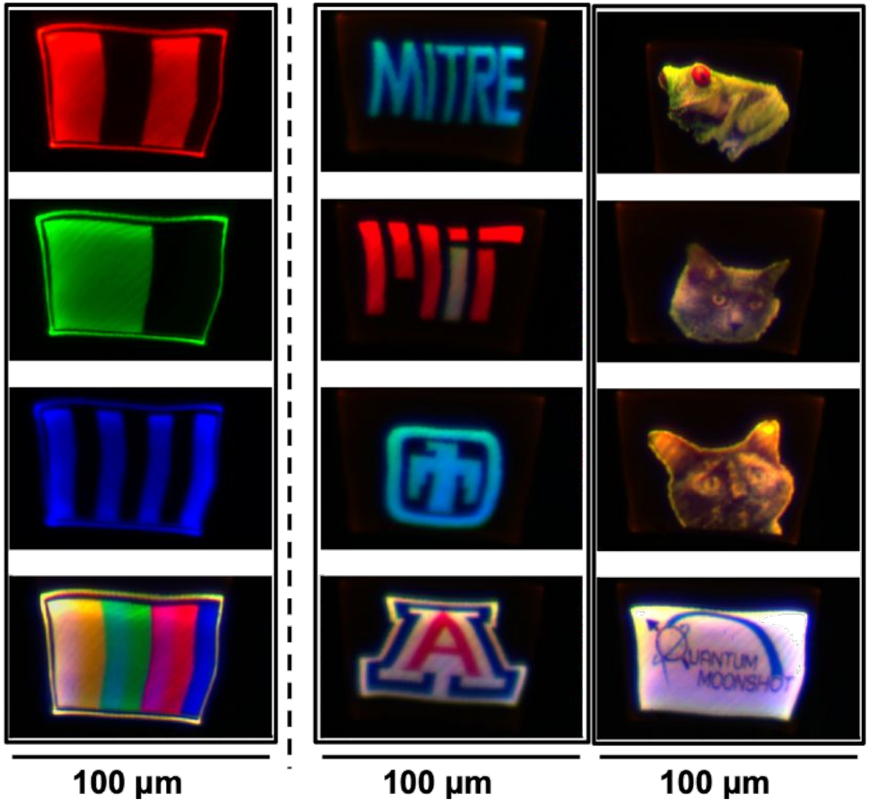






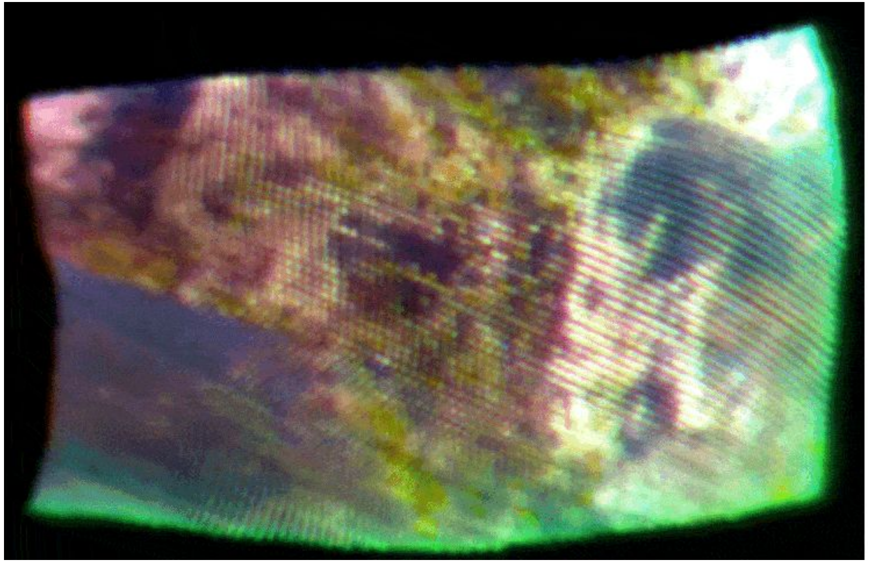
M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al, https://arxiv.org/pdf/2406.17662

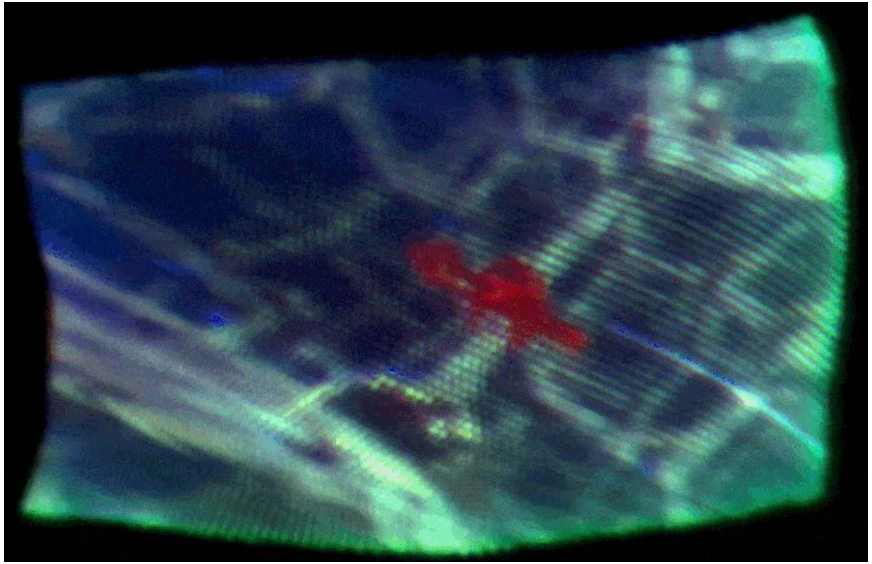
FULL COLOR IMAGES

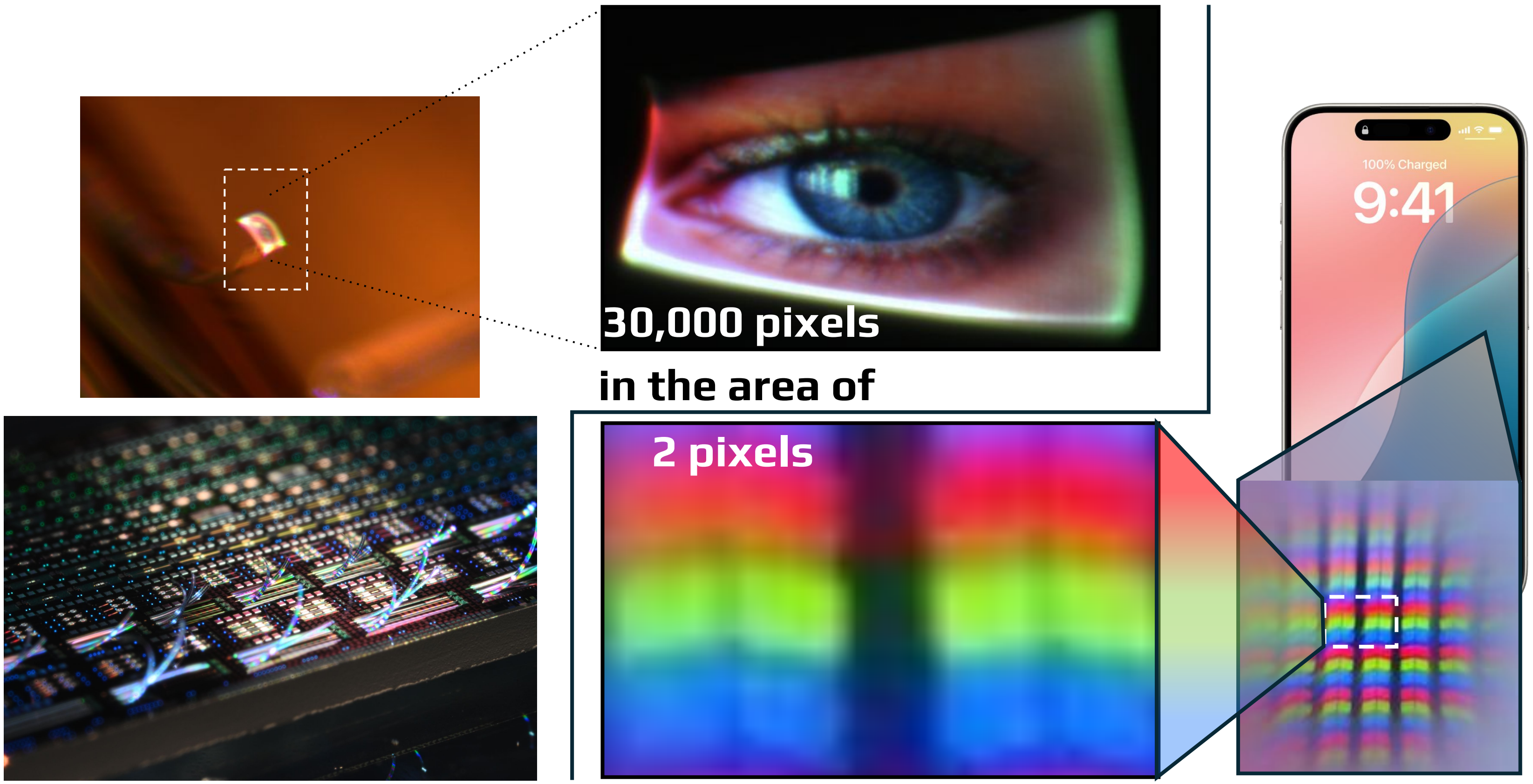


M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al, https://arxiv.org/pdf/2406.17662

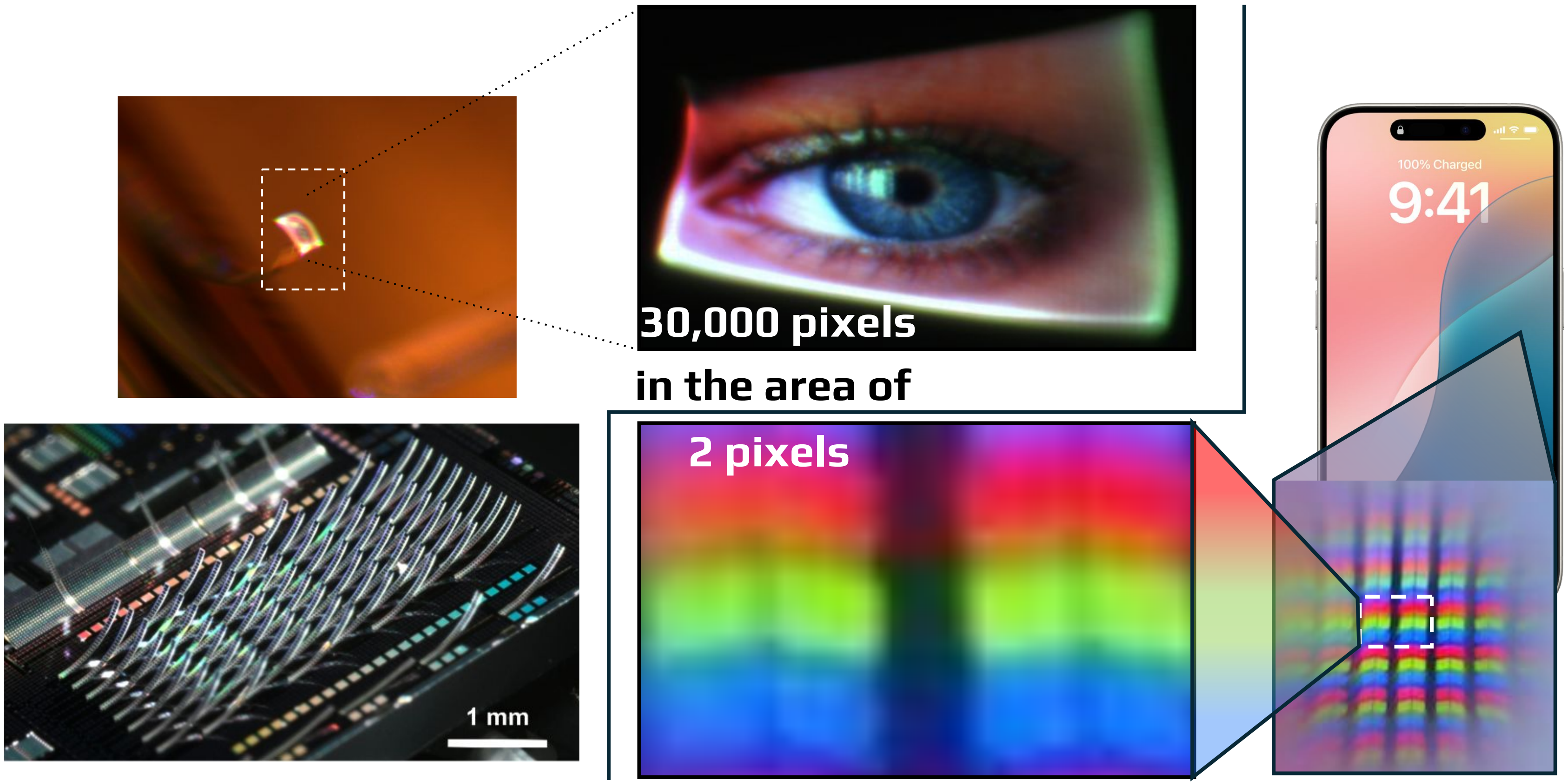
FULL COLOR VIDEO





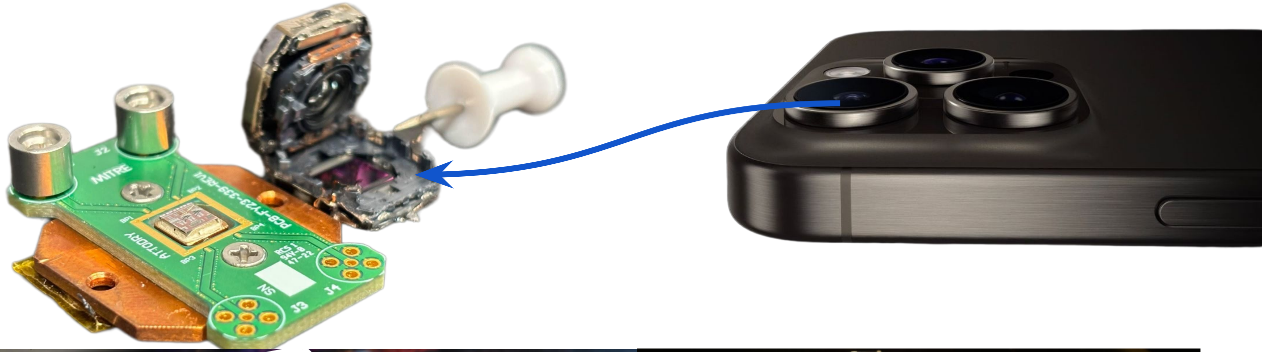


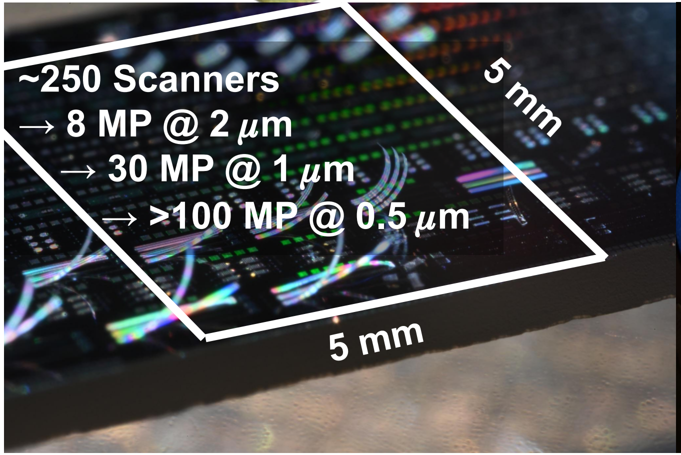
M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al, https://arxiv.org/pdf/2406.17662

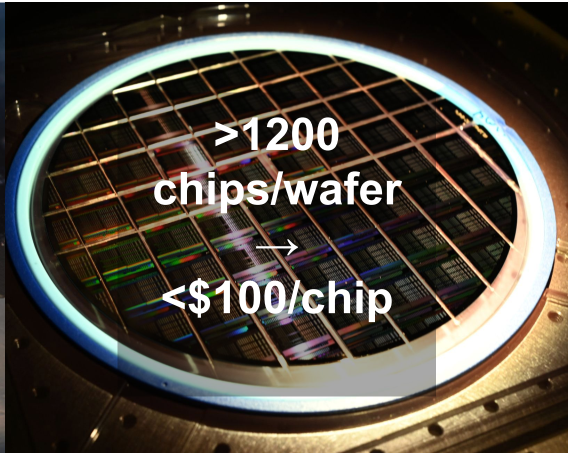


M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al, https://arxiv.org/pdf/2406.17662

SWARM: Scanning waveguide AR microdisplay







M Saha, Y. H. Wen, A. Greenspon, M. Zimmerman, et al, https://arxiv.org/pdf/2406.17662

WHAT'S NEXT: LAB TO DEVICE

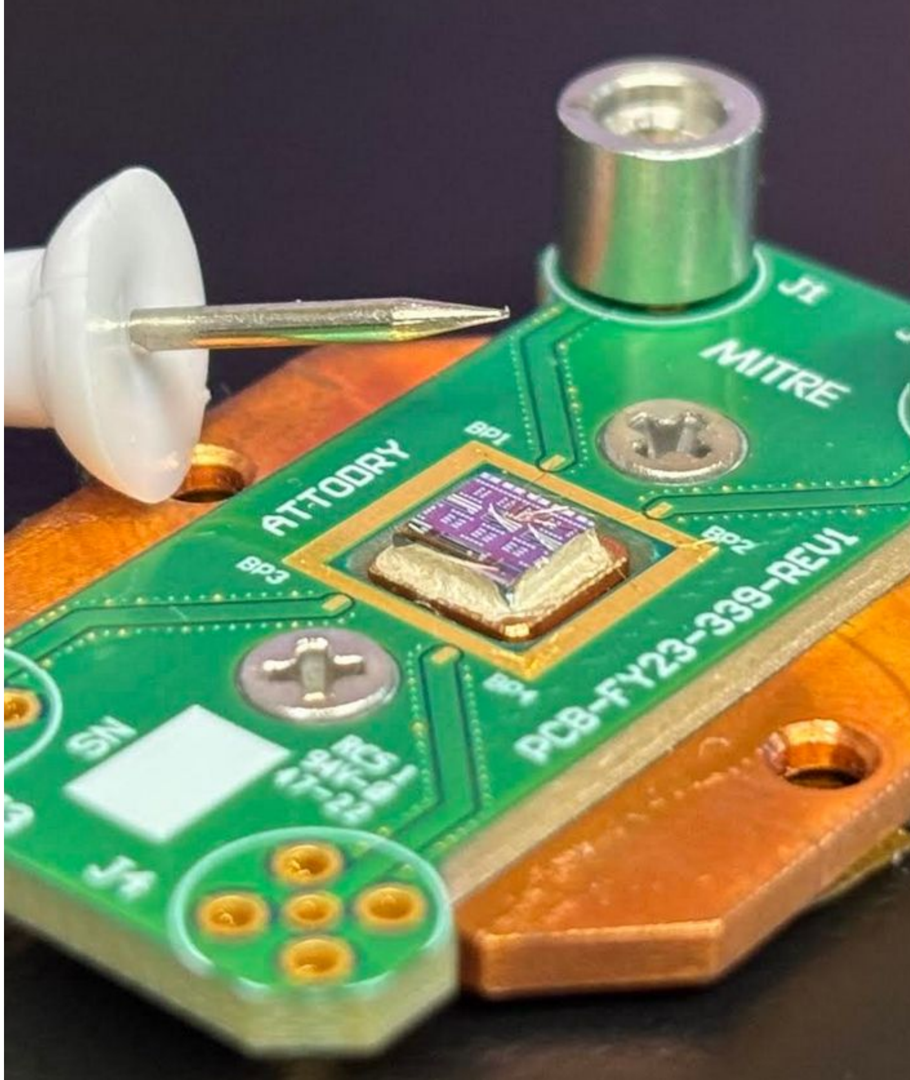
FABRICATION UNIFORMITY

MICROLENS ARRAY DESIGN FOR ABERRATION COMPENSATION

COUPLING OPTICS INTEGRATION

ELECTROMECHANICAL INTEGRATION INTO SMALL FORM FACTOR

CONTROL & SOFTWARE INTERFACE



Deep-Dive #2:

Verifiable Measurements

Bidirectional Nonlinear Optical Tomography: Unbiased Characterization of Off- and On-Chip Coupling Efficiencies

Bo-Han Wu,[1,2,*] Mahmoud Jalali Mehrabad,[1] and Dirk Englund[1]

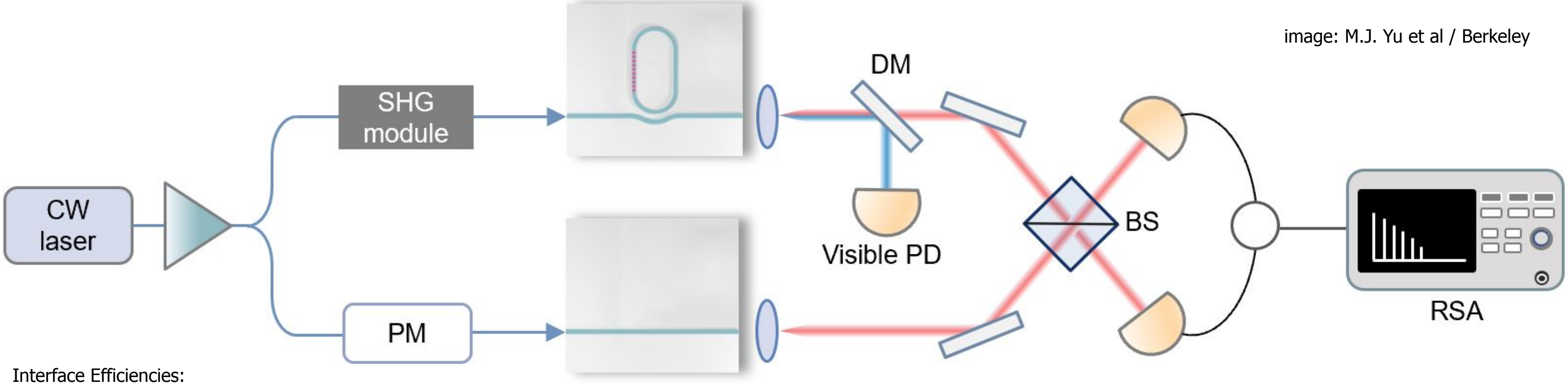
[1]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Electrical and Computer Engineering, University of Hawai'i at Mānoa, Honolulu, Hawai'i 96822, USA*

(Dated: October 16, 2025)

https://arxiv.org/pdf/2510.13110

Squeezed light generation and detection

image: M.J. Yu et al / Berkeley



Interface Efficiencies:

- $\eta 1(2\omega)$: Input coupling efficiency at pump wavelength ($2\omega \approx 775$ nm)
- $\eta 2(\omega)$: Output coupling efficiency at signal wavelength ($\omega \approx 1550$ nm)

Cavity Parameters:

- $\kappa c(i)$: Coupling decay rate for ring i (Hz)
- $\kappa i(i)$: Intrinsic loss rate for ring i (Hz)
- $\mu(2\omega)$: Inter-ring coupling rate at 2ω (Hz)
- $\omega 0$: Cavity resonance frequency (rad/s)

Material Properties:

- $n\omega = 2.18$: Refractive index at 1550 nm
- $n2\omega = 2.14$: Refractive index at 775 nm
- deff = 19.5 pm/V: Effective nonlinear coefficient

A1: Single-Mode Operation

- All fields propagate in the fundamental TE mode.
- Scattering to higher-order modes is negligible.
- Justification: Mode overlap calculations indicate greater than 99% fundamental mode content.

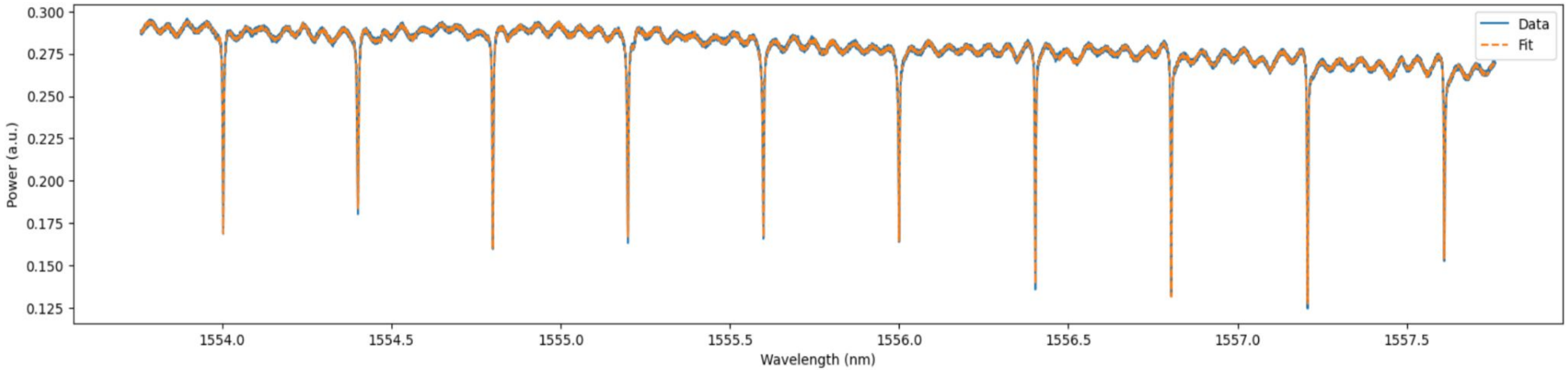
A2: Lorentzian Resonance

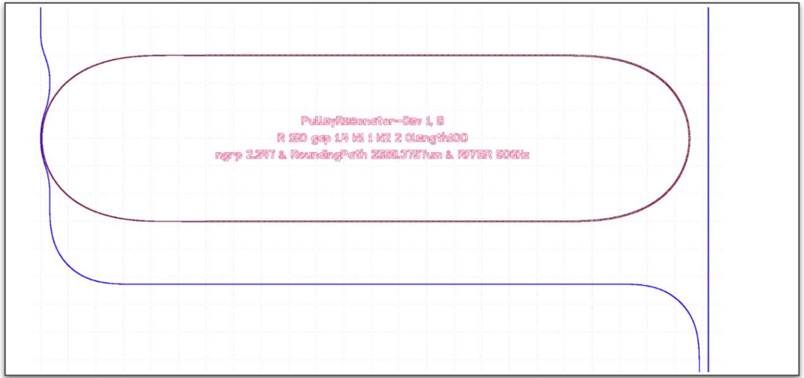
- The cavity transmission exhibits a Lorentzian lineshape, valid for Q > 10^4 (typical for TFLN rings).
- Mathematical Representation: $T(\omega) = \kappa c^2 / ((\omega - \omega 0)^2 + (\kappa tot/2)^2)$

A3: Time-Invariant Device (not currently the case → experiment upgrades + model upgrades). We want, but can't have: Device parameters (κ, μ, ω0) remain constant during measurements;

Digital twin: fitting the data with ambiguity

- Refined fitting routine give is expressive enough -- but it's not unique.
-





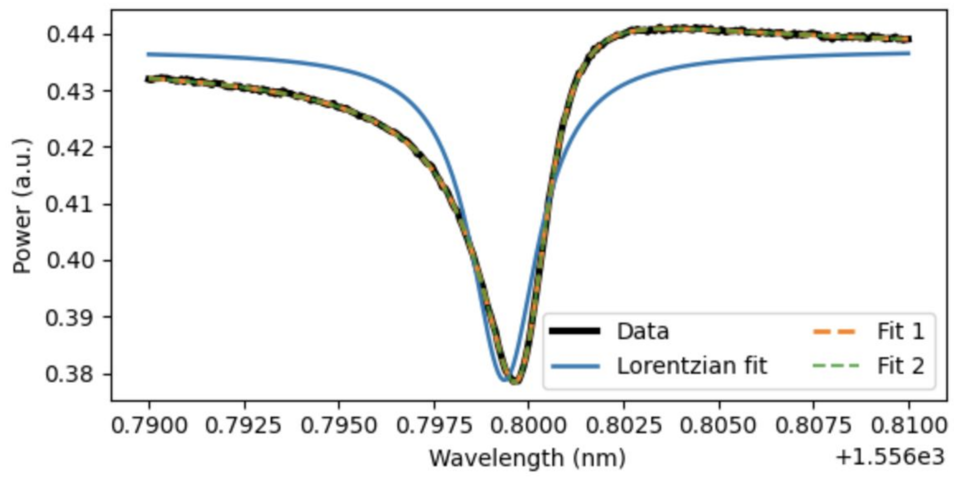
<u>Ring</u>
Bending radius = 150um
Gap = 1.4, 1.2, 1.0, 0.9, 0.8, 0.7um
Bus wg width = 1 um
Ring wg width = 2 um
Coupling length = 100 um
Ring path length = 2668 um
FSR = 50 GHz
Simulated neff = 1.939147
Simulated ng = 2.242544

Estimated bus waveguide length (facet to facet) = 9.1 mm

S. Y. Ma / MIT
Berkeley team
(measurements)

Digital twin: fitting the data with ambiguity

- Different parameters can lead to the same fitting curve: ambiguity



fitting includes laser model + detector model + linear optics

$$E_{\text{det}}(\omega) = E_{\text{chip}}(\omega) + E_{\text{bg}}(\omega),$$
$$P(\omega) = |E_{\text{det}}(\omega)|^2 .$$

Dispersive coupler

$$\kappa(\omega) = \kappa_0 + \kappa_1(\omega - \omega_{\text{ref}})$$

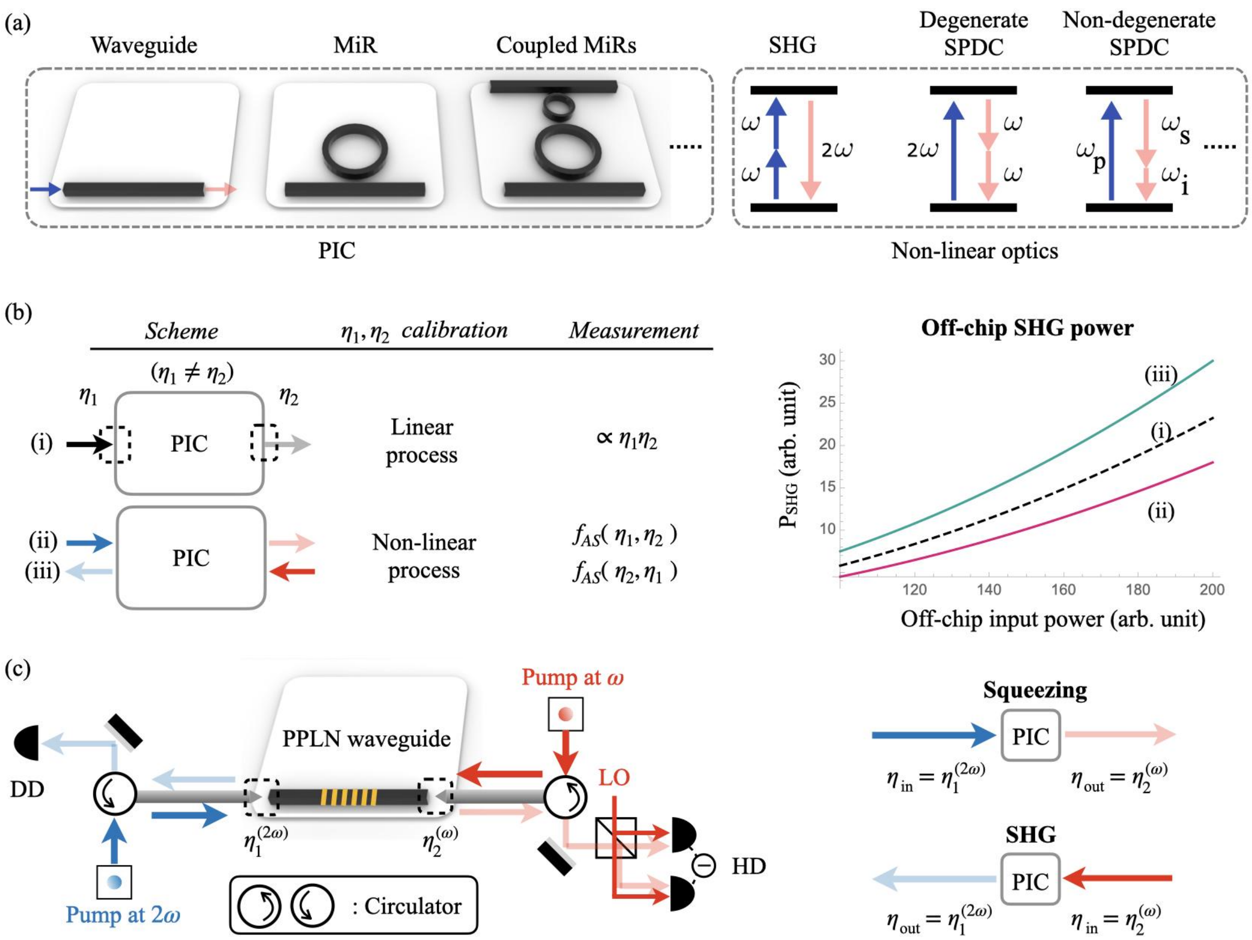
Lorentzian fit: default symmetric fitting.

Fit 1: backscattering c = 0.0173, minor model $|E_{bg}|<0.001$.

Fit 2: backscattering c = 0.001, minor model $|E_{bg}|\sim 0.008$, altered κ ~ 2%.

Cannot uncover ground truth in **one measurement** → Need to lift collinearity with

Bidirectional nonlinear optical tomography (BNOT)



→ A general calibration framework for nonlinear integrated optics.

→ Conv. approach, (i), uses linear process to "**degenerately"** estimate interface efficiencies.

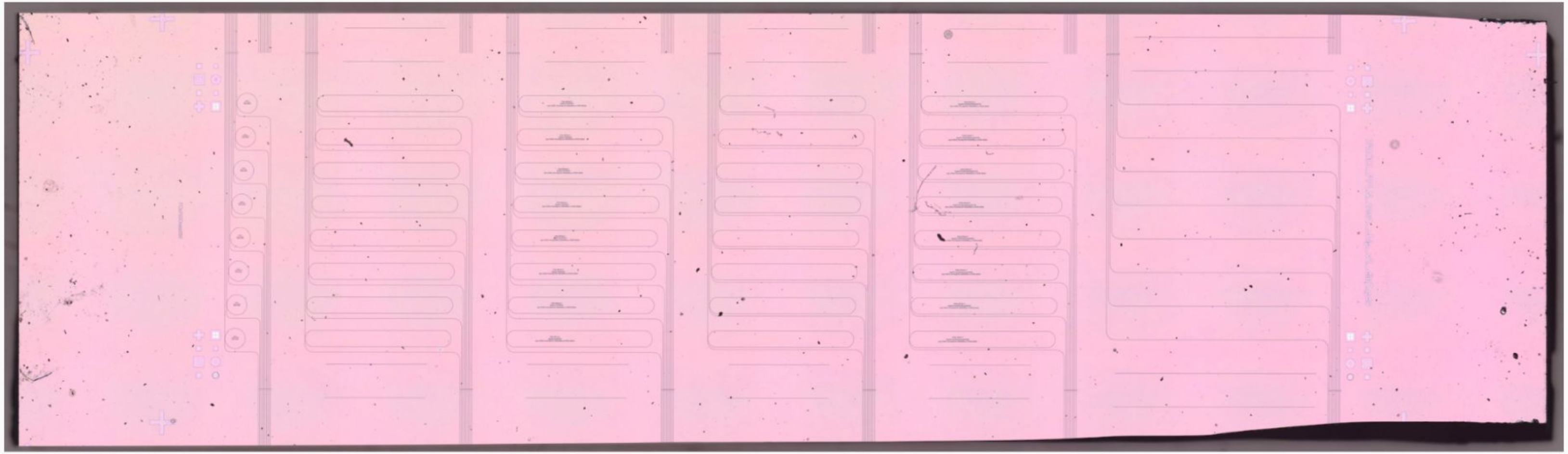
→ BNOT approach, (ii) and (iii), uses nonlinear process to "**non-degenerately"** estimate interface efficiencies.

→ Particularly important for strongly squeezed light experiments (i.e., >16 dB).

arXiv:2510.13110

Closed-Loop Architecture

Workflow: (1) Ingest lab data → (2) Run BNOT → (3) Update twin → (4) Predict performance



M.J. Yu group, Berkeley

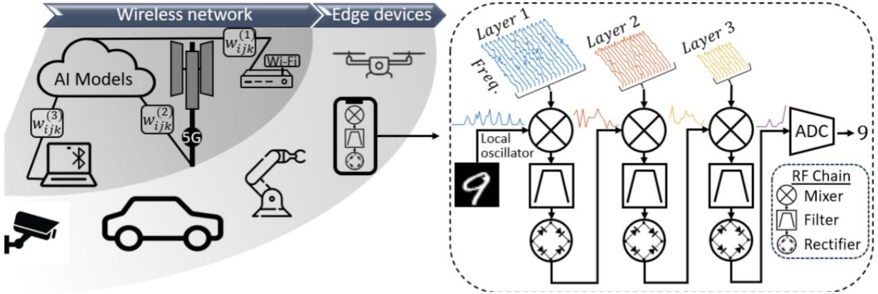
Toolchain: Python-based simulator with pytest validation suite and continuous integration.

H. Wu et al, arXiv:2510.13110

Deep Learning on Wireless Networks:

Matching the Accuracy of Von Neumann Machines below Their Thermodynamic Limits

Sri Krishna Vadlamani, Kfir Sulimany, Zhihui Gao, Tingjun Chen, and Dirk Englund - **arXiv:2504.1775**



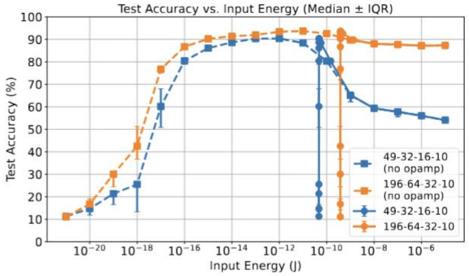


FIG. 4. **MNIST classification accuracy of RF analog and digital deep learning** MNIST classification accuracy is presented as a function of the total energy per MAC for digital and analog systems.

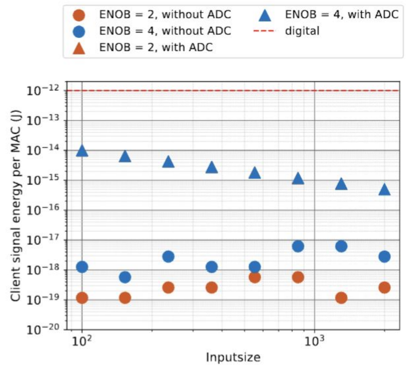
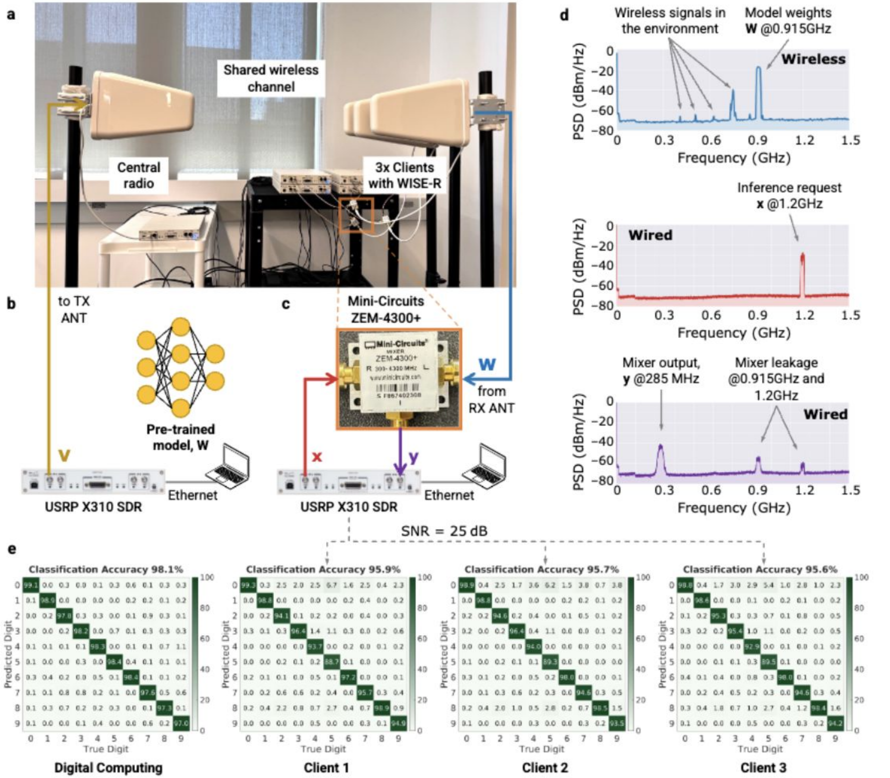


FIG. 3. **Energy per MAC versus inputsize of the inner product vectors**

Disaggregated Deep Learning via In-Physics Computing at Radio Frequency

Zhihui Gao, Sri Krishna Vadlamani, Kfir Sulimany, Dirk Englund, and Tingjun Chen - **arXiv:2505.09267**



Basic research → **Technology transition** → Industry, gov't partnerships / **Startups**

LIGHTMATTER

https://lightmatter.co/

matrix PARTNERS · Hewlett Packard Enterprise · Viking · SPARK CAPITAL · SIP Global Partners · Google ventures

QuNETT

Quantum Network Technologies, Inc

Quantum communications, interconnects, memories

qunett.com

|QuEra>

.. with CUA team members Lukin, Greiner, Vuletic

Compute the Impossible™

AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

CELEBRATING 10 YEARS OF RE:INVENT

COMPUTING

Mathematicians are deploying algorithms to stop gerrymandering

THE 22 STARTUPS TO WATCH IN BOSTON IN 2022

MIT Technology Review

COMPUTING

This new startup has built a record-breaking 256-qubit quantum computer

DUST IDENTITY

physical objects · secure digital

SAP · KLEINER PERKINS · UPSTART100 · LOCKHEED MARTIN · AIRBUS VENTURES · NEW SCIENCE VENTURES · CASTLE ISLAND VENTURES · Angular Ventures

LIGHTIUM

CRAFTING LIGHT, SHAPING THE FUTURE

PRODUCTION-GRADE PHOTONIC FOUNDRY FOR THIN-FILM LITHIUM NIOBATE

Axiomatic AI

First-Principles Engineering

Boston (HQ) Toronto · Barcelona

https://axiomatic-ai.com

OptiCore

Home · Team · News

Zaijun Chen
Co-founder

Mengjie Yu
Co-founder

Ryan Hamerly
Co-founder

Startup hiring (https://netpreme.com/)

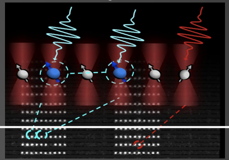
Netpreme

Dr. Zhizhen Zhong

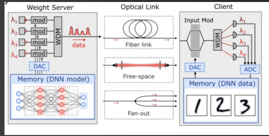
From question to ~~solution~~ question to…

requirement / question **q**

q = "I need photonic control system"



q ="Develop reconfigurable photonic link for AI acceleration"



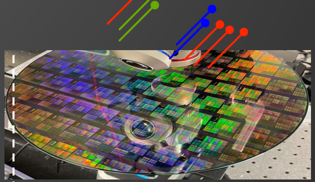
$f_{KG}(.)$

Design system $\mathbf{s}_\Theta$ → target **s(x)** :

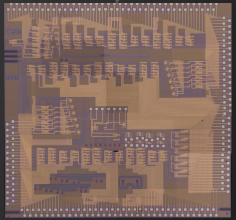
$\min_\Theta C = ||\mathbf{s}_\Theta(\mathbf{x}) - \mathbf{s}(\mathbf{x})||$

→ optimum **a**

$\mathbf{s}_\Theta$ = high-speed beam array projector



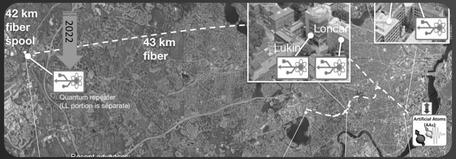
$\mathbf{s}_\Theta$ =silicon photonic AI system PIC



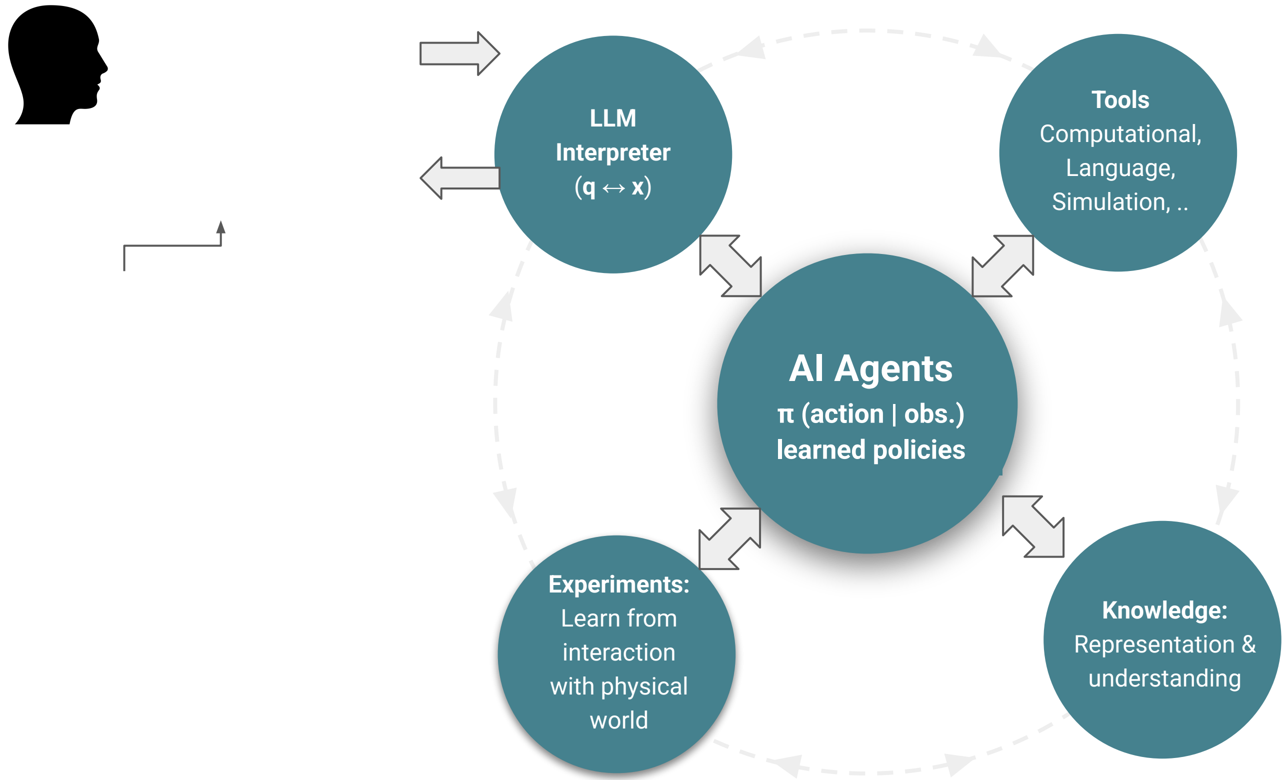
answer **a**

$\min_\Theta$ C

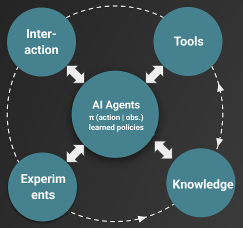
Benchmarking



What "generative design" may look like



Example 1: math-grounded reasoning



The Verification Gap

OpenAI's GPT5

ChatGPT 5 >

We are solving the equation:

$$5.9 = x + 5.11$$

Step 1: Isolate x

Subtract 5.11 from both sides:

$$x = 5.9 - 5.11$$

Step 2: Perform the subtraction carefully

$$5.90 - 5.11 = 0.79 - 0.00 - 0.21 = -0.21$$

(We can think of 5.90 - 5.11 as subtracting 5.11 from 5.90 directly, which equals -0.21.)

Final Answer:

$$x = -0.21$$

So the solution is:

$$x = -0.21$$

Ask anything

5

Leading LLM error rates of 10^{-2} are much too high for science, math, engineering

A New Algorithm for Computing Branch Number of Non-Singular Matrices over Finite Fields

P.R. Mishra[1], Yogesh Kum...

Next, we divide the second term on the right-hand side of (1) into the cases where $w_h(Mx) \le \lfloor \frac{n+1}{2} \rfloor$ and $w_h(Mx) > \lfloor \frac{n+1}{2} \rfloor$. Therefore, we have

$$\begin{aligned} &\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, \left\lfloor \frac{n+1}{2} \right\rfloor < w_h(x) \le n\right\} \\ &= \min\left\{\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, \left\lfloor \frac{n+1}{2} \right\rfloor < w_h(x) \le n, w_h(Mx) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\},\right. \\ &\left.\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, \left\lfloor \frac{n+1}{2} \right\rfloor < w_h(x) \le n, w_h(Mx) > \left\lfloor \frac{n+1}{2} \right\rfloor\right\}\right\}. \end{aligned} \tag{2}$$

Note that for the second term of the right-hand side of Equation (2), $h(M,x) = w_h(x)+w_h(Mx) > 2\lfloor \frac{n+1}{2} \rfloor + 1 \ge n+1$. However, we know that the upper bound for $\mathcal{B}(M)$ is $n+1$. Thus, we conclude that the second term of the right-hand side of (2) will not contribute to the computation of the branch number.

Therefore, from (1) and (2), we have

$$\begin{aligned} \mathcal{B}(M) = \min&\left\{\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\},\right. \\ &\left.\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, \left\lfloor \frac{n+1}{2} \right\rfloor < w_h(x) \le n, w_h(Mx) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\}\right\}. \end{aligned} \tag{3}$$

Again, we note that

$$\left\{h(M,x) \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor \frac{n+1}{2} \right\rfloor, w_h(Mx) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\} \subseteq \left\{h(M,x) \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\}.$$

Therefore,

$$\begin{aligned} &\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\} \le \\ &\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor \frac{n+1}{2} \right\rfloor, w_h(Mx) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\}. \end{aligned} \tag{4}$$

Note that the right-hand side of (4) is always greater than or equal to the left-hand side of (4). Therefore, if we include this extra term in (3), it will not affect

Therefore,

$$\begin{aligned} &\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\} \le \\ &\min\left\{h(M,x) \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor \frac{n+1}{2} \right\rfloor, w_h(Mx) \le \left\lfloor \frac{n+1}{2} \right\rfloor\right\}. \end{aligned} \tag{4}$$

Note that the right-hand side of (4) is always greater than or equal to the left-hand side of (4). Therefore, if we include this extra term in (3), it will not affect

GPT5: "Nothing wrong in the derivation" ❌

Thought for 2m 5s >

Short answer: the argument is essentially correct; I only see small nits/typos and a couple of places where you can tighten the logic. Here's a quick pass:

What's correct

- The reduction to the two regimes $w_h(x) \le \lfloor (n+1)/2 \rfloor$ and $w_h(x) > \lfloor (n+1)/2 \rfloor$, then discarding the subcase $w_h(x), w_h(Mx) > \lfloor (n+1)/2 \rfloor$, is valid: in that subcase

  $h(M,x) = w_h(x) + w_h(Mx) \ge 2(\lfloor \frac{n+1}{2} \rfloor + 1)$

  which is $> n+1$ for all n. Hence those x can't achieve the minimum (since there always exists some x with $h(M,x) \le n+1$; see the note below).
- The change of variables $y = Mx$ and the equality $h(M,x) = h(M^{-1}, y)$ are correct, and the final identity

  follows since $\min\{\min_S a, \min_S b\} = \min_S \min\{a,b\}$ over the same finite set S.

Ax Prover: Verification failed

```
1767    -- KEY SORRIES: THE ORIGINAL PROOF USES MINIMUMS OF EMPTY FINSETS, WHICH ARE NOT DEFINED
1768    have h_constrained_low_nonempty : ({ x ∈ (@low_weight_vectors p q n _ _ _) | wH (M.mulVec x) ≤ (n + 1) / 2 }).Nonempty := by sorry
1769    have h_first_set_nonempty : { y ∈ (@high_weight_vectors p q n _ _ _) | wH (M.mulVec y) ≤ (n + 1) / 2 }.Nonempty := by sorry
1770    have h_second_set_nonempty : { y ∈ (@high_weight_vectors p q n _ _ _) | wH (M.mulVec y) > (n + 1) / 2 }.Nonempty := by sorry
```

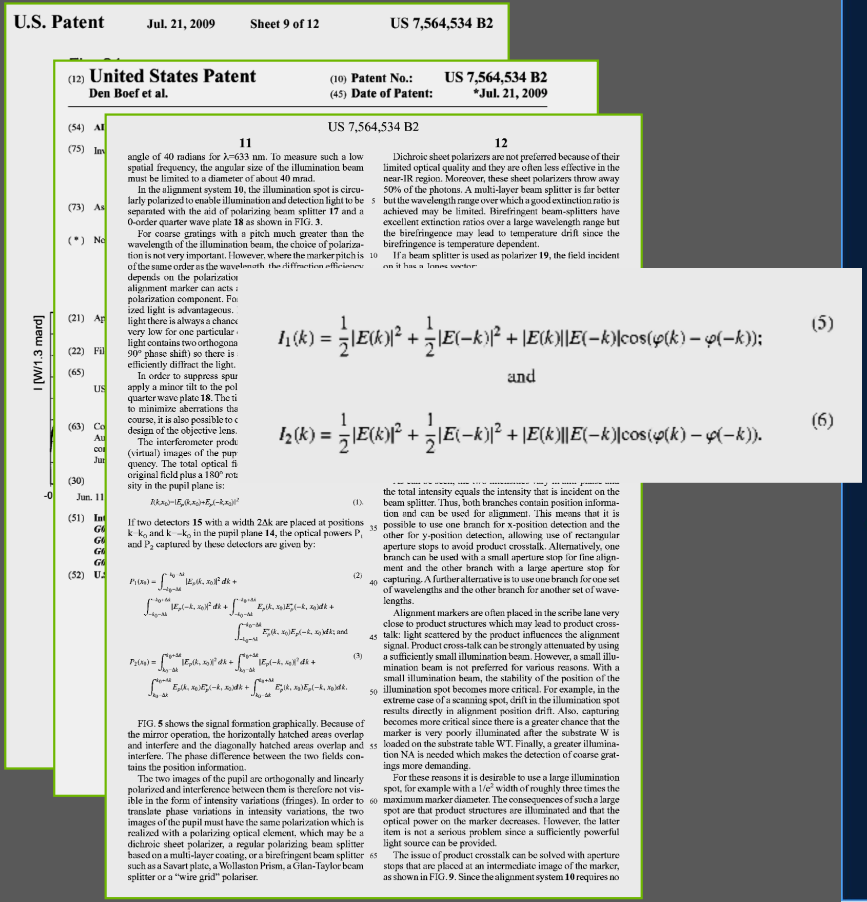
Ax-Prover: A Deep Reasoning Agentic Framework for Theorem Proving in Mathematics and Quantum Physics

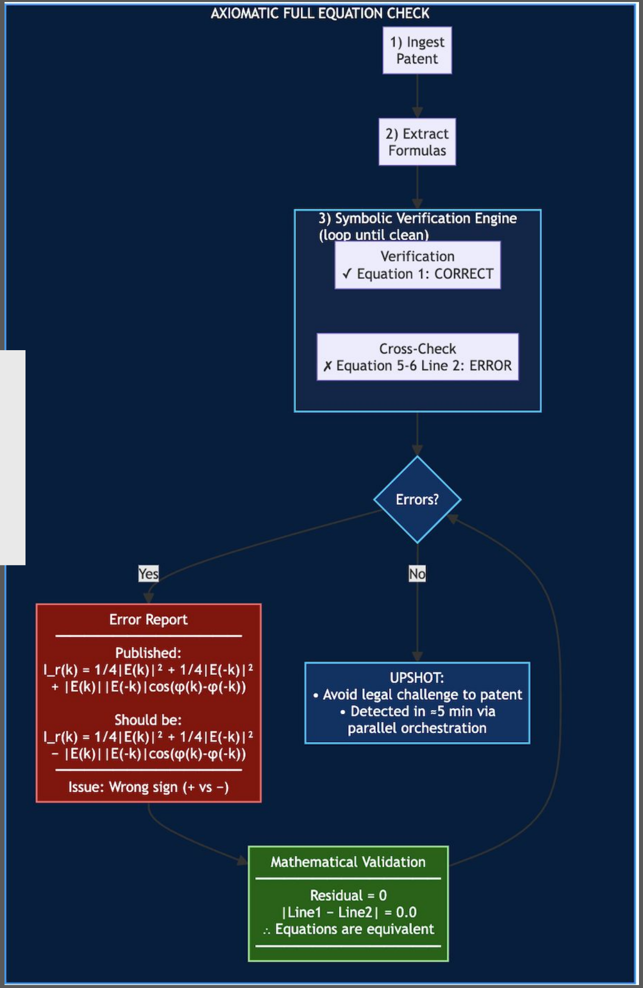
Marco Del Tredici[1], Jacob McCarran[1], Benjamin Breen[1], Javier Aspuru Mijares[1], Weichen Winston Yin[1], Jacob M. Taylor[1], Frank Koppens[2], and Dirk Englund[1,3]

[1]Axiomatic_AI
[2]Institut de Ciències Fotòniques (ICFO)
[3]Massachusetts Institute of Technology (MIT)

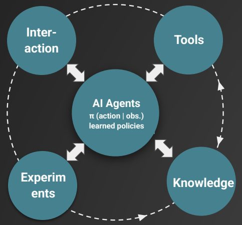
>cs>arXiv:2510.12787v2 (2025)

Concrete scenario:
VVUQ of patent (100+ equations)

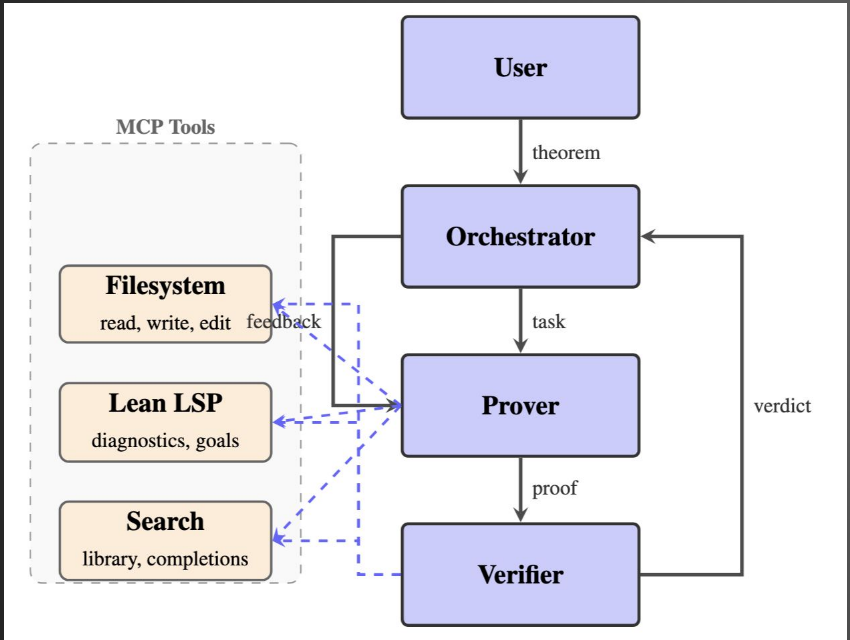




PATENT: 28,457 tokens



AX-Prover: A DEEP REASONING AGENTIC FRAMEWORK



Example of “basic” math (abstract algebra)

For example, a competition problem may ask to determine all positive integers a, b such that

$$\frac{a^2+b^2}{ab+1} \in \mathbb{Z},$$

Example cryptography proof

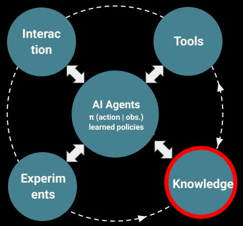
Theorem 1. *The branch number of an invertible matrix $M \in M_n(\mathbb{F}_q)$ is given as*

$$\mathcal{B}(M) = \min\left\{\min\left\{h(M,x), h(M^{-1},x)\right\} \mid x \in \mathbb{F}_q^n, 1 \le w_h(x) \le \left\lfloor\frac{n+1}{2}\right\rfloor\right\},$$

where $h(M,x) = w_h(x) + w_h(Mx)$.

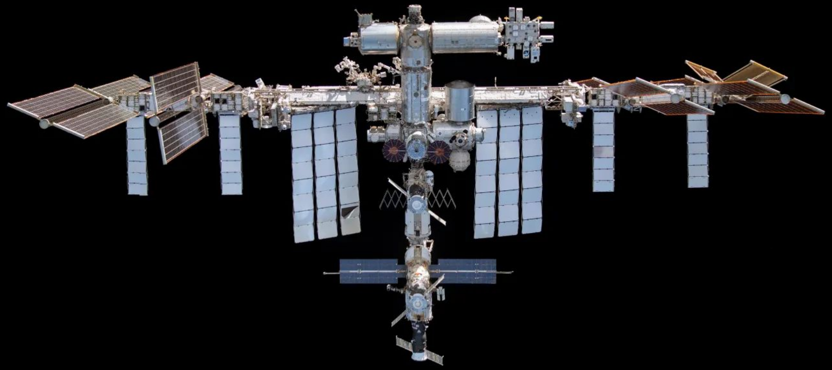
Dataset	Subset	Ax-Prover (pass@1)	Sonnet (pass@1)	DS-Prover (pass@1)	Kimina†
NuminaMath	solved-K	81%	7%	48%	100%
	solved-H	47%	8%	14%	0%
	unsolved	26%	1%	18%	0%
	total	51%	5%	28%	31%
AbstractAlgebra	easy	72%	10%	26%	12%
	intermediate	56%	6%	22%	14%
	total	64%	8%	24%	13%
QuantumTheorems	easy	100%	54%	88%	72%
	intermediate	92%	18%	48%	34%
	total	96%	40%	61%	57%

Axioamtic_AI + MIT team, arXiv:2510.12787 (2025)



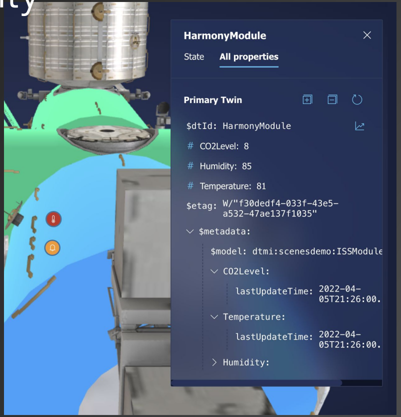
The Knowledge (Ingress) Gap

understand, model



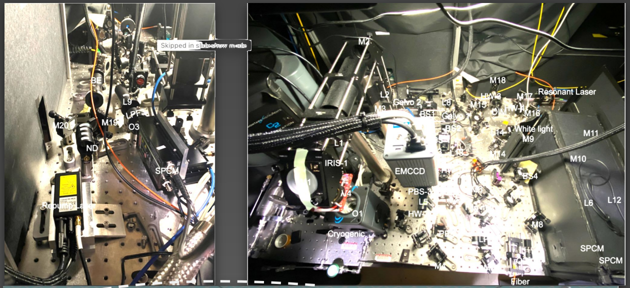
control, refine

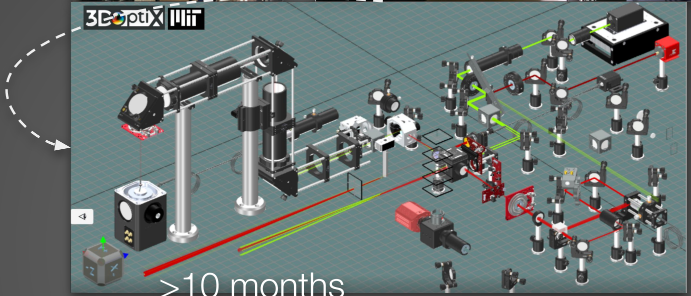
Reality



Quantum - extension to EDA

MIT postdoc: “Ansys is state of art but building models containing quantum physics was impossible.. [..] I ended up instead using python, then pytorch, .. and maybe eventually connect with to Ansys, Synopsis tools

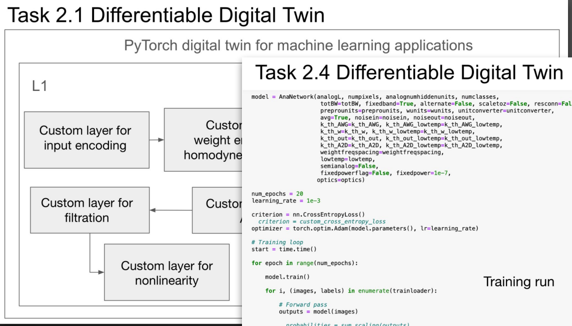


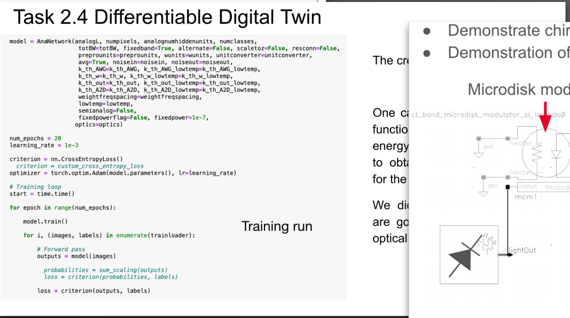


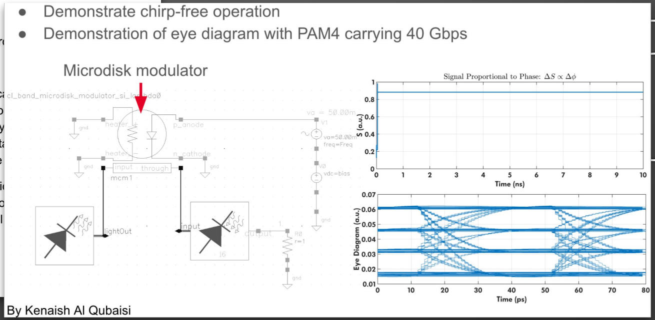
>10 months

Cadence Virtuoso ADE Suite

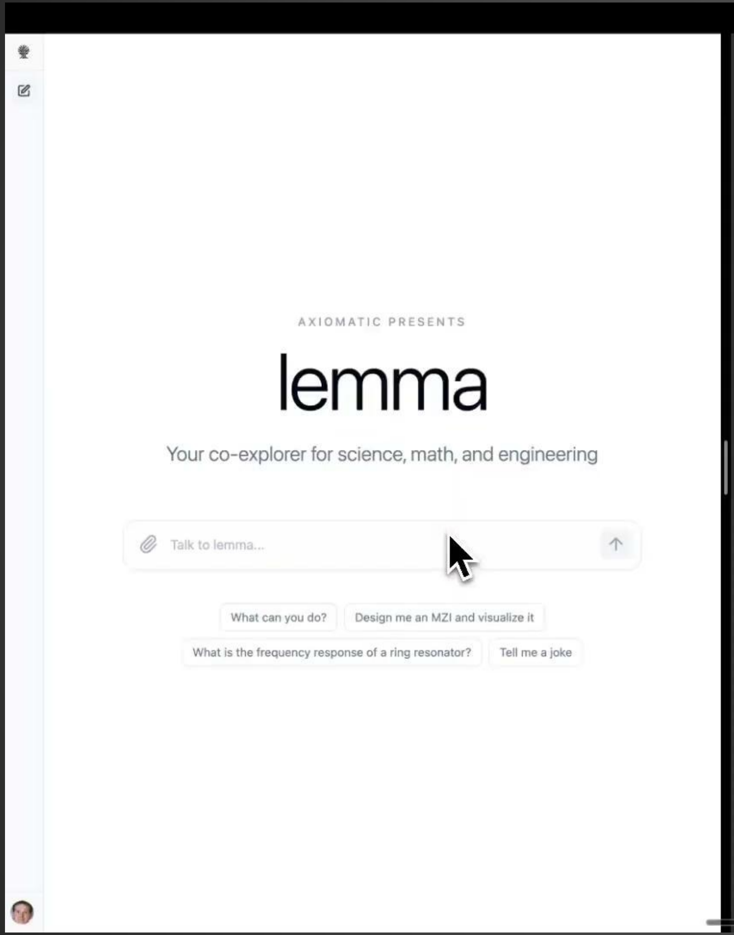
MIT postdoc : “ I didn’t know about digital twin modeling before joining this project, i wish they’d teach that in grad school”

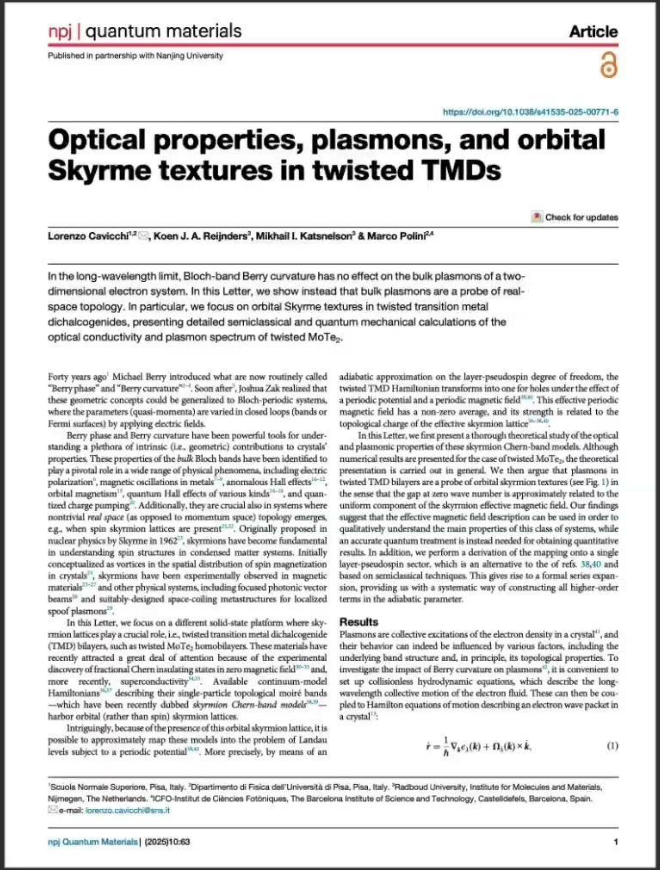






Example 1: Reading comprehension





npj | quantum materials

Published in partnership with Nanjing University

Article

https://doi.org/10.1038/s41535-025-00771-6

Optical properties, plasmons, and orbital Skyrme textures in twisted TMDs

Check for updates

Lorenzo Cavicchi[1,2] ✉, Koen J. A. Reijnders[3], Mikhail I. Katsnelson[3] & Marco Polini[2,4]

In the long-wavelength limit, Bloch-band Berry curvature has no effect on the bulk plasmons of a two-dimensional electron system. In this Letter, we show instead that bulk plasmons are a probe of real-space topology. In particular, we focus on orbital Skyrme textures in twisted transition metal dichalcogenides, presenting detailed semiclassical and quantum mechanical calculations of the optical conductivity and plasmon spectrum of twisted $MoTe_2$.

Forty years ago[1] Michael Berry introduced what are now routinely called "Berry phase" and "Berry curvature"[2–4]. Soon after[5], Joshua Zak realized that these geometric concepts could be generalized to Bloch-periodic systems, where the parameters (quasi-momenta) are varied in closed loops (bands or Fermi surfaces) by applying electric fields.

Berry phase and Berry curvature have been powerful tools for understanding a plethora of intrinsic (i.e., geometric) contributions to crystals' properties. These properties of the *bulk* Bloch bands have been identified to play a pivotal role in a wide range of physical phenomena, including electric polarization[6], magnetic oscillations in metals[7–9], anomalous Hall effects[10–12], orbital magnetism[13], quantum Hall effects of various kinds[14–19], and quantized charge pumping[20]. Additionally, they are crucial also in systems where nontrivial *real space* (as opposed to momentum space) topology emerges, e.g., when spin skyrmion lattices are present[21,22]. Originally proposed in nuclear physics by Skyrme in 1962[23], skyrmions have become fundamental in understanding spin structures in condensed matter systems. Initially conceptualized as vortices in the spatial distribution of spin magnetization in crystals[24], skyrmions have been experimentally observed in magnetic materials[25–27] and other physical systems, including focused photonic vector beams[28] and suitably-designed space-coiling metastructures for localized spoof plasmons[29].

In this Letter, we focus on a different solid-state platform where skyrmion lattices play a crucial role, i.e., twisted transition metal dichalcogenide (TMD) bilayers, such as twisted $MoTe_2$ homobilayers. These materials have recently attracted a great deal of attention because of the experimental discovery of fractional Chern insulating states in zero magnetic field[30–33] and, more recently, superconductivity[34,35]. Available continuum-model Hamiltonians[36,37] describing their single-particle topological moiré bands—which have been recently dubbed *skyrmion Chern-band models*[38,39]—harbor orbital (rather than spin) skyrmion lattices.

Intriguingly, because of the presence of this orbital skyrmion lattice, it is possible to approximately map these models into the problem of Landau levels subject to a periodic potential[38,40]. More precisely, by means of an adiabatic approximation on the layer-pseudospin degree of freedom, the twisted TMD Hamiltonian transforms into one for holes under the effect of a periodic potential and a periodic magnetic field[38,40]. This effective periodic magnetic field has a non-zero average, and its strength is related to the topological charge of the effective skyrmion lattice[36–38,40].

In this Letter, we first present a thorough theoretical study of the optical and plasmonic properties of these skyrmion Chern-band models. Although numerical results are presented for the case of twisted $MoTe_2$, the theoretical presentation is carried out in general. We then argue that plasmons in twisted TMD bilayers are a probe of orbital skyrmion textures (see Fig. 1) in the sense that the gap at zero wave number is approximately related to the uniform component of the skyrmion effective magnetic field. Our findings suggest that the effective magnetic field description can be used in order to qualitatively understand the main properties of this class of systems, while an accurate quantum treatment is instead needed for obtaining quantitative results. In addition, we perform a derivation of the mapping onto a single layer-pseudospin sector, which is an alternative to the of refs. 38,40 and based on semiclassical techniques. This gives rise to a formal series expansion, providing us with a systematic way of constructing all higher-order terms in the adiabatic parameter.

Results

Plasmons are collective excitations of the electron density in a crystal[41], and their behavior can indeed be influenced by various factors, including the underlying band structure and, in principle, its topological properties. To investigate the impact of Berry curvature on plasmons[42], it is convenient to set up collisionless hydrodynamic equations, which describe the long-wavelength collective motion of the electron fluid. These can then be coupled to Hamilton equations of motion describing an electron wave packet in a crystal[13]:

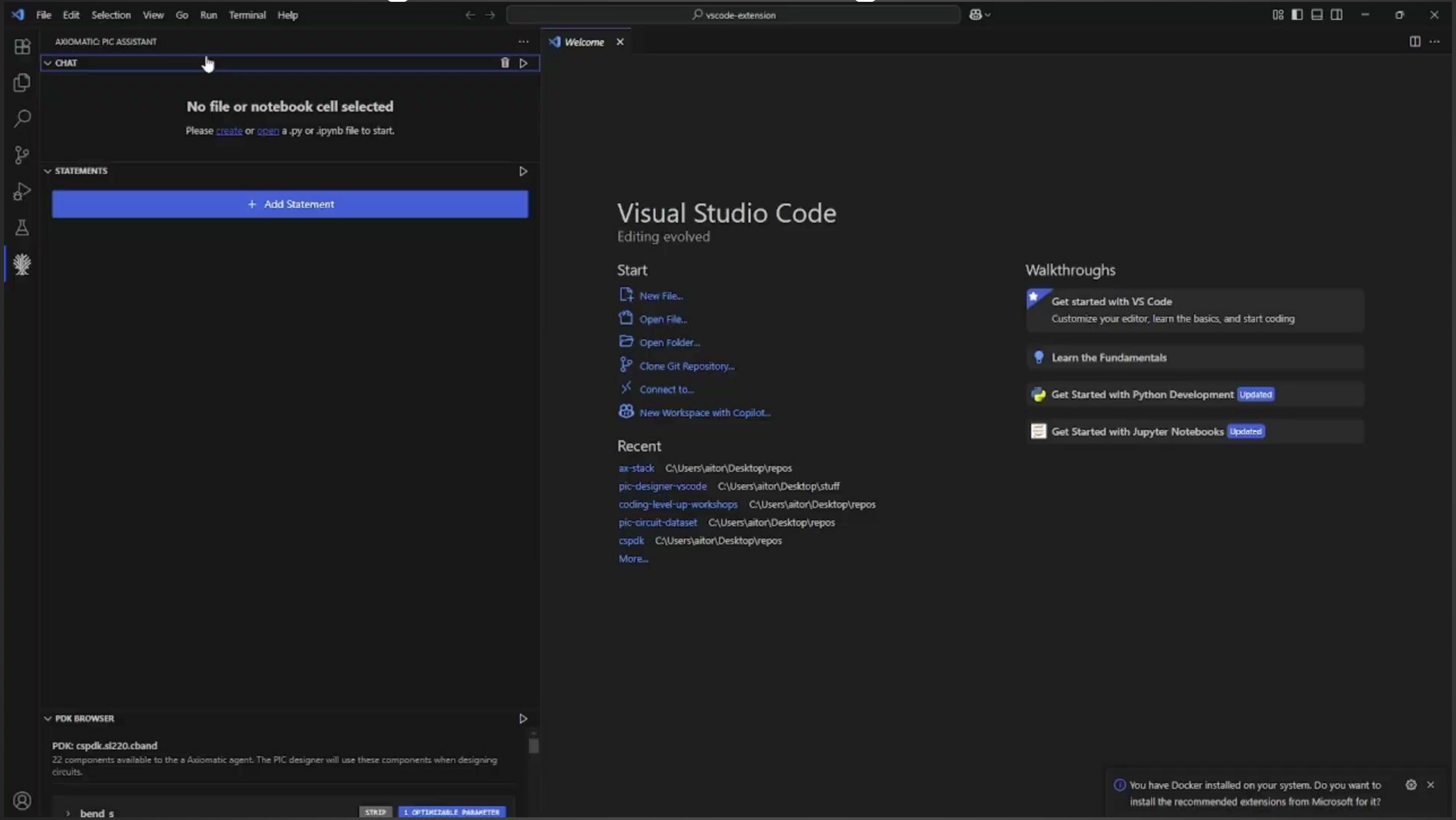
$$\dot{\boldsymbol{r}} = \frac{1}{\hbar}\nabla_{\boldsymbol{k}}\epsilon_{\lambda}(\boldsymbol{k}) + \boldsymbol{\Omega}_{\lambda}(\boldsymbol{k}) \times \dot{\boldsymbol{k}}, \quad (1)$$

[1]Scuola Normale Superiore, Pisa, Italy. [2]Dipartimento di Fisica dell'Università di Pisa, Pisa, Italy. [3]Radboud University, Institute for Molecules and Materials, Nijmegen, The Netherlands. [4]ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Castelldefels, Barcelona, Spain. ✉e-mail: lorenzo.cavicchi@sns.it

npj Quantum Materials | (2025)10:63 1

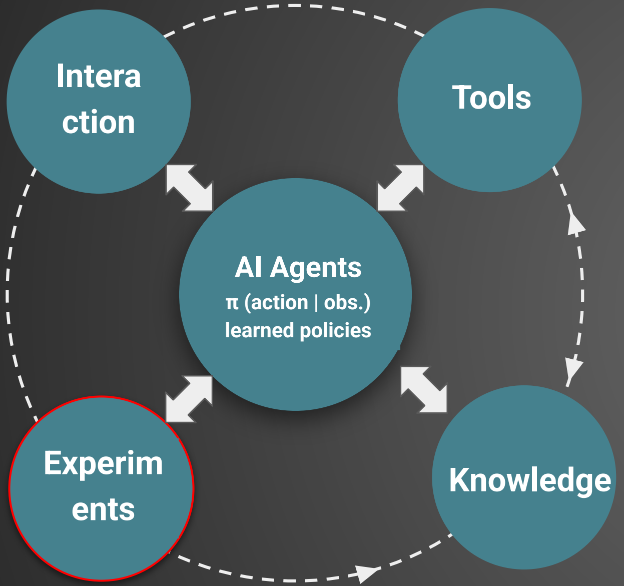
Photonic integrated circuit design, automated

U. Toronto - Axiomatic_AI - GDSFactory - MIT

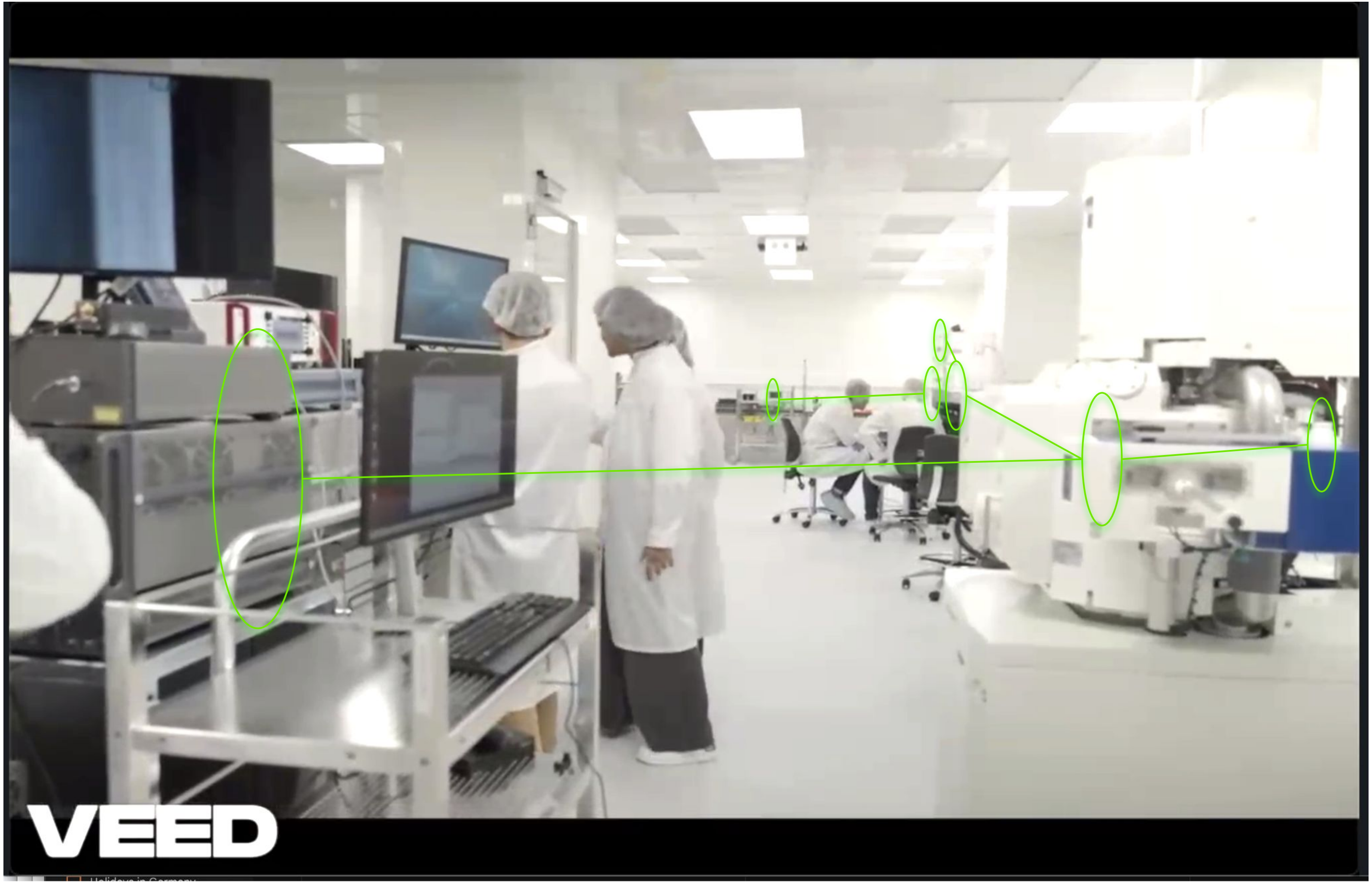


Opensource: https://github.com/JPPhotonics/PhIDO-Release

A Sharma, V Ansari, Y Fu, R Iyer, J Matres, T Tamas, O Akdeniz, D E , **J Poon, OFC 2024; to be published;**



4. Experiments


VEED

Digital twins to make the physical world intelligible

- Infinite space for compiling computing into physics and control.
- Math-grounded AI can help us explore this space of deep computational design

“To reach true AI, we need to build machines that learn models of the world in a self-supervised manner.”

Y. LeCun, “A Path Towards Autonomous Machine Intelligence” (2022)

The Verification Principle:

An AI system can create and maintain knowledge only to the extent that it can verify that knowledge itself.

Rich Sutton, “Verification: The Key to AI” (2002)