



Design of Super Functional Materials and Devices using Artificial Intelligence

In-Ho Lee

Korea Research Institute of Standards and Science Daejeon 34113, Korea

ihlee@kriss.re.kr

KIAS, June 26–30, 2023



Contents

- Background {artificial intelligence}
- Evolutionary learning
- Artificial neural networks
- Applications
- [RDFSEARCH, AMADEUS]



What's next?

As DeepMind co-founder, Demis Hassabis explains, Zero was not programmed to understand Go specifically, it could be reprogrammed to discover information in other fields: drug discovery, **protein folding**, quantum chemistry, particle physics, and **material design**.

- 2011: Relu
- 2012: CNN, Alex Net
- 2013: VAE
- 2014: GAN
- 2015: TF
- 2016: alpha Go
- 2017: Zero, cycleGAN
- 2018: alpha fold
- 2020: alpha fold 2
- 2021: Ramanujan,
Floor planning,
Nowcasting,
DM21
- 2022: alpha tensor
ChatGPT

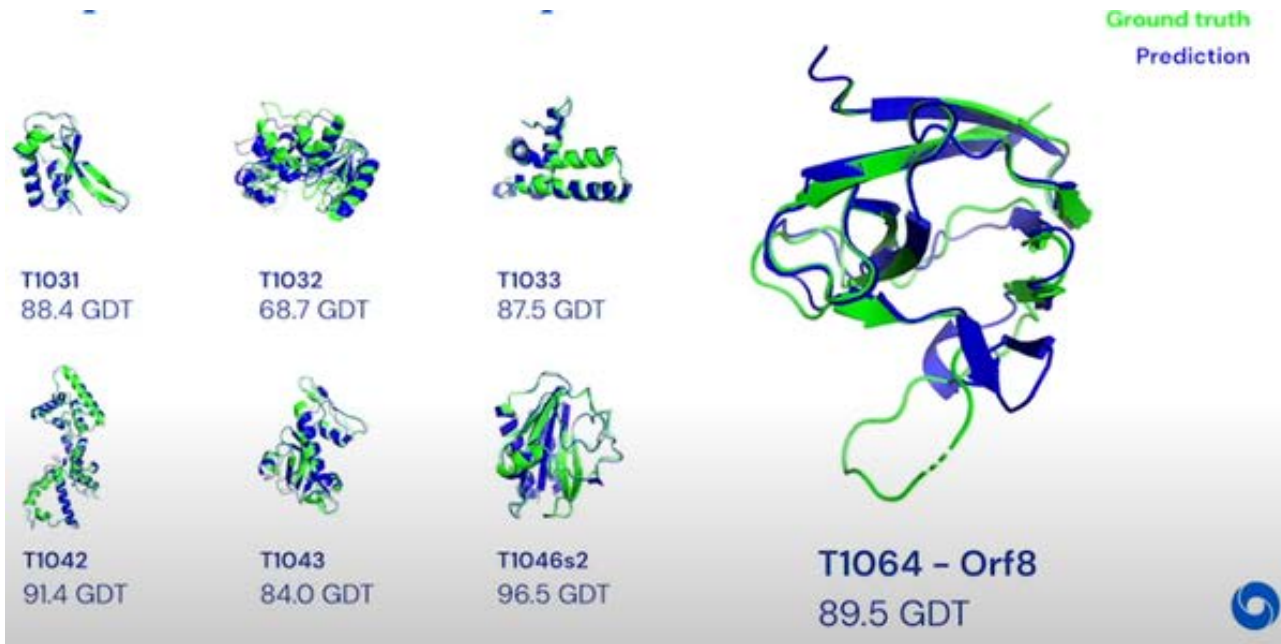


Breakthrough Prize

Demis Hassabis

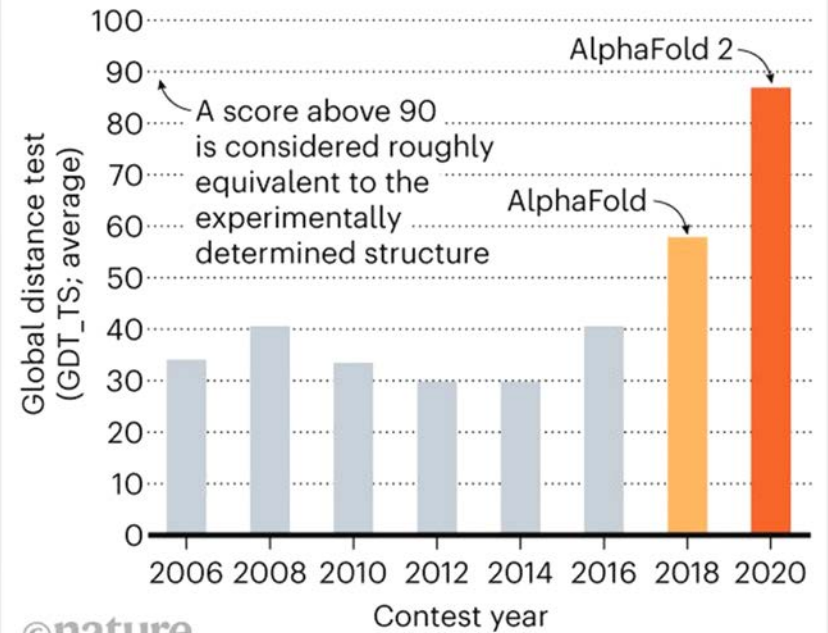


Protein folding problem



STRUCTURE SOLVER

DeepMind's AlphaFold 2 algorithm significantly outperformed other teams at the CASP14 protein-folding contest — and its previous version's performance at the last CASP.



RoseTTAFold
ESMFold, Meta's Rival To AlphaFold

Isomorphic Labs, 갤럭스, 굳인텔리전스, 히츠, 아론티어, 스탠다임



얼마나 창의적인가? 새로운 수학기공식들

$$\pi = \sum_{n=0}^{\infty} \frac{1}{16^n} \left(\frac{4}{8n+1} - \frac{2}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6} \right).$$

$$\frac{8}{\pi^2} = 1 - \frac{2 \times 1^4 - 1^3}{7 - \frac{2 \times 2^4 - 2^3}{19 - \frac{2 \times 3^4 - 3^3}{37 - \frac{2 \times 4^4 - 4^3}{\dots}}}}$$

$$\frac{12}{7\zeta(3)} = 1 \times 2 - \frac{16 \times 1^6}{3 \times 12 - \frac{16 \times 2^6}{5 \times 32 - \frac{16 \times 3^6}{7 \times 62 - \frac{16 \times 4^6}{\dots}}}}$$

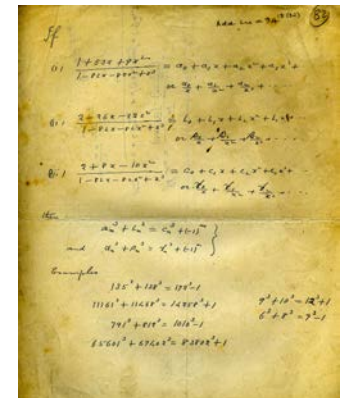
$$\frac{8}{7\zeta(3)} = 1 \times 1 - \frac{1^6}{3 \times 7 - \frac{2^6}{5 \times 19 - \frac{3^6}{7 \times 37 - \frac{4^6}{\dots}}}}$$

$$\frac{2}{-1+2G} = 3 + 0 \times 7 - \frac{6 \times 1^3}{3 + 1 \times 10 - \frac{8 \times 2^3}{3 + 2 \times 13 - \frac{10 \times 3^3}{\dots}}}$$

(4)



Srinivasa Ramanujan: The mathematical genius who credited his 3900 formulae to visions from Goddess Mahalakshmi.



Floor planning

Can We Design Personalized Chips?

Systemic Complexity: 18-24 months and 100s of millions to bring a new chip to market

Chess



Number of states: 10^{123}

Win / Lose

Go



$\sim 10^{360}$

Win / Lose

Placement



$> 10^{90,000}$

Better / Worse

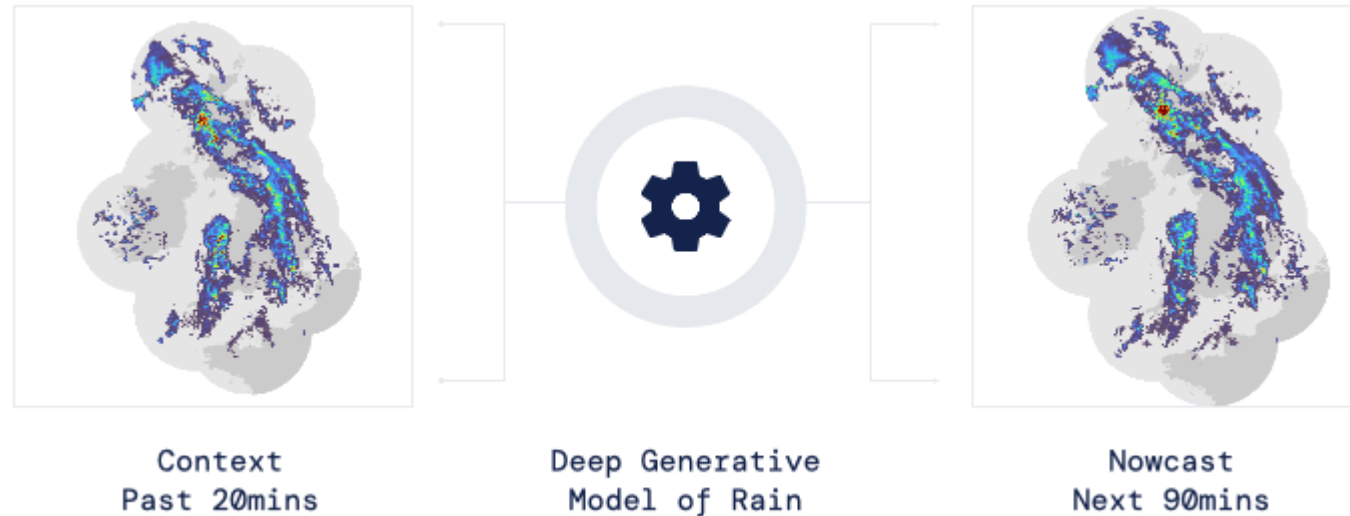
Real Chips



$> 10^{???,???$

Impossible?

Generative models for nowcasting



2012년이 딥러닝의 원년이다.

2011년 Relu 함수의 출현 : vanishing gradient problem 해결, 표현력이 높은 딥러닝이 시작됨.

2012년 Dropout 방법의 출현 : overfitting problem 해결, 정규화된 훈련이 가능함.

2012년 CNN 출현 : 방대한 이미지 데이터 처리 문제 해결, 비대칭 데이터에 대한 처리가 가능함.

2013년 VAE 출현 : 생성모델로서 다양한 가능성을 제안함.

2014년 GAN 출현 : 생성모델로서 많은 응용 가능성을 제안함.

2015년 TensorFlow 구현의 출현과 보급 : 실질적인 딥러닝의 민주화가 단행됨.

2016년 알파고 출현 : 바둑 정복

2017년 알파고 제로 출현 : tabula rasa, first-principles 방법을 이용하여 바둑 정복, 진정한 바둑의 정복이 가능함. 기보없이 바둑 정복함. 프로그래머에게 절대로 지지 않는 사실상 바둑의 신이 탄생함. 바둑의 역사 5000년 [Mastering the game of Go without human knowledge | Nature](http://incredible.egloos.com/7372719) <http://incredible.egloos.com/7372719>

2017년 cycleGAN 출현 <http://incredible.egloos.com/7530913>

2018년 알파폴드 출현

2020년 알파폴드 2 출현 <http://incredible.egloos.com/7479504> Highly accurate protein structure prediction with AlphaFold | Nature

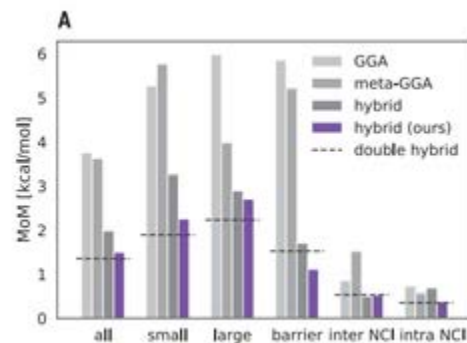
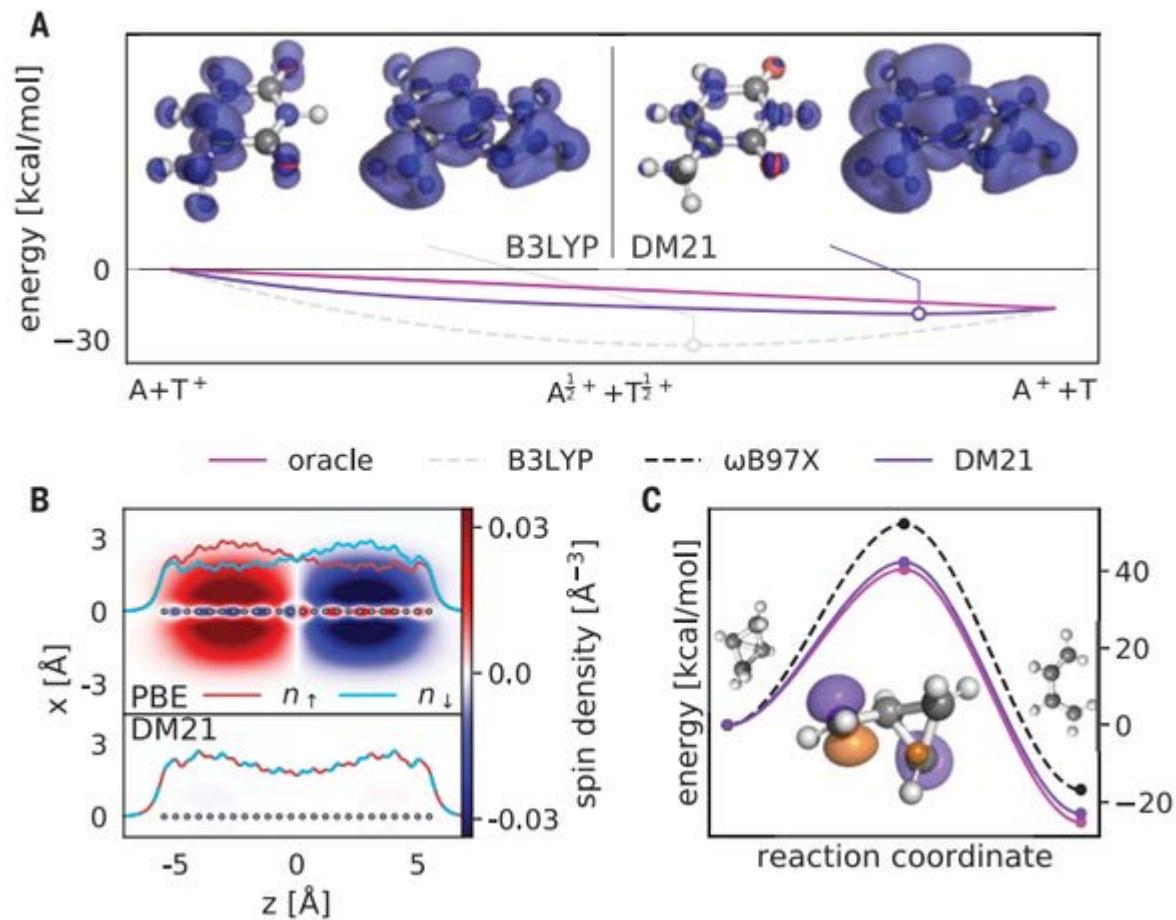
2021년 라마누잔 기계 (수학공식 생성기) <http://incredible.egloos.com/7511552> Generating conjectures on fundamental constants with the Ramanujan Machine | Nature

2021년 플로어플랜닝 (반도체 칩 설계) [A graph placement methodology for fast chip design | Nature](https://www.nature.com/articles/s41586-021-03854-z)

2021년 나우캐스팅 (단기 일기예보) <https://www.nature.com/articles/s41586-021-03854-z>

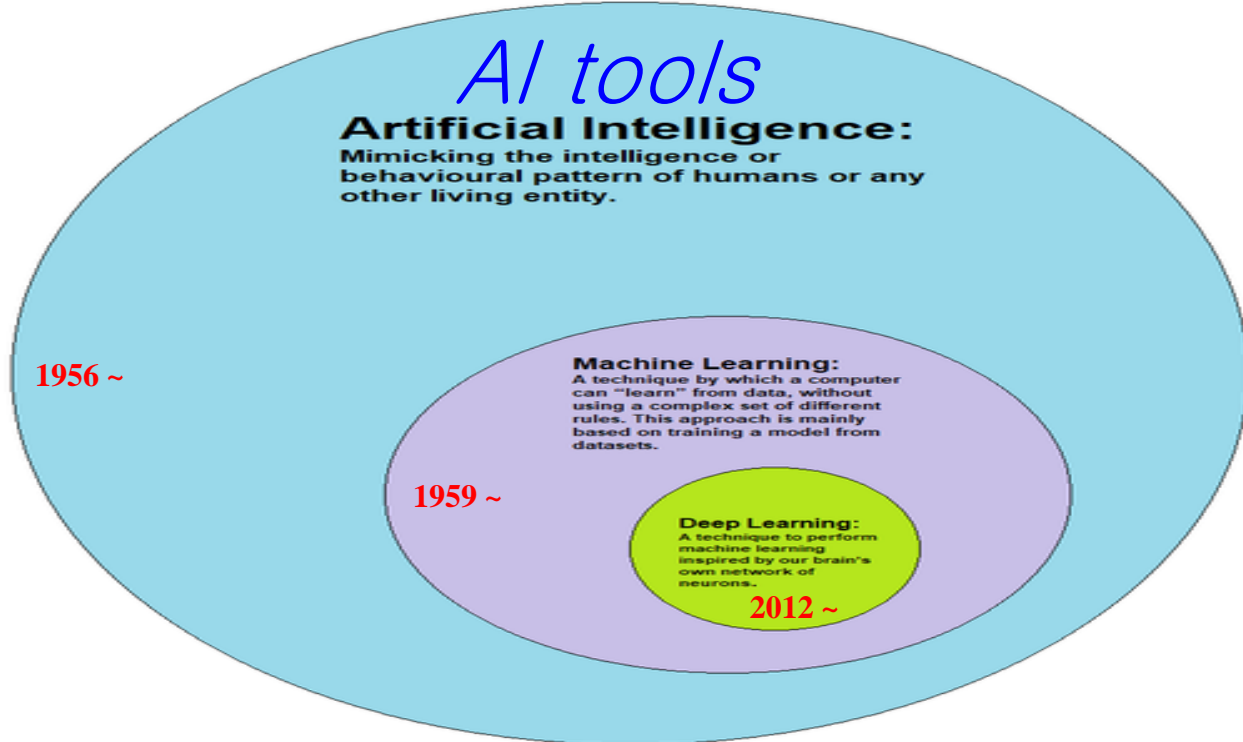
2022년 알파텐서, ChatGPT

Density functionals

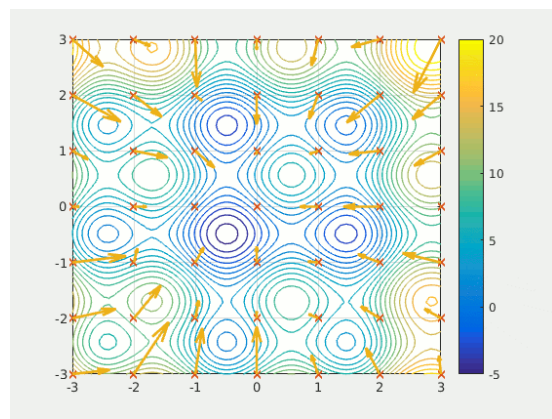


B

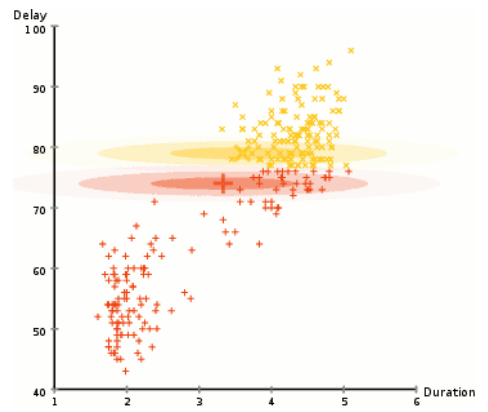
	GMTKN55	BBB	QM9
MoM			
SCAN(-D3 _{BJ})	3.63	52.90	3.57
ω B97X(-V)	2.45	70.28	2.13
M06-2X	2.08	64.50	2.20
PW6B95:D3 ₀	1.98	57.81	2.24
DM21	1.50	13.20	1.66



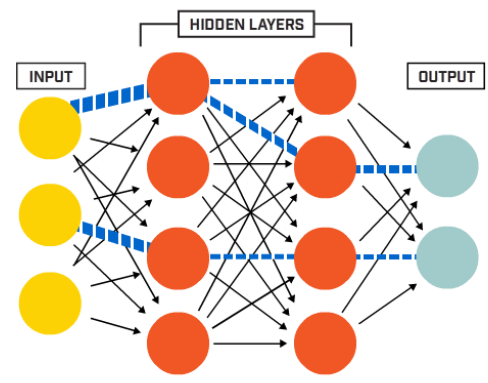
Search and optimization



Classifiers and statistical learning methods



Artificial neural networks



The overall research goal of artificial intelligence is to create technology that allows computers and machines to function in an intelligent manner.

Forms of Machine Learning

1. Supervised learning

Four learning types

2. Unsupervised learning

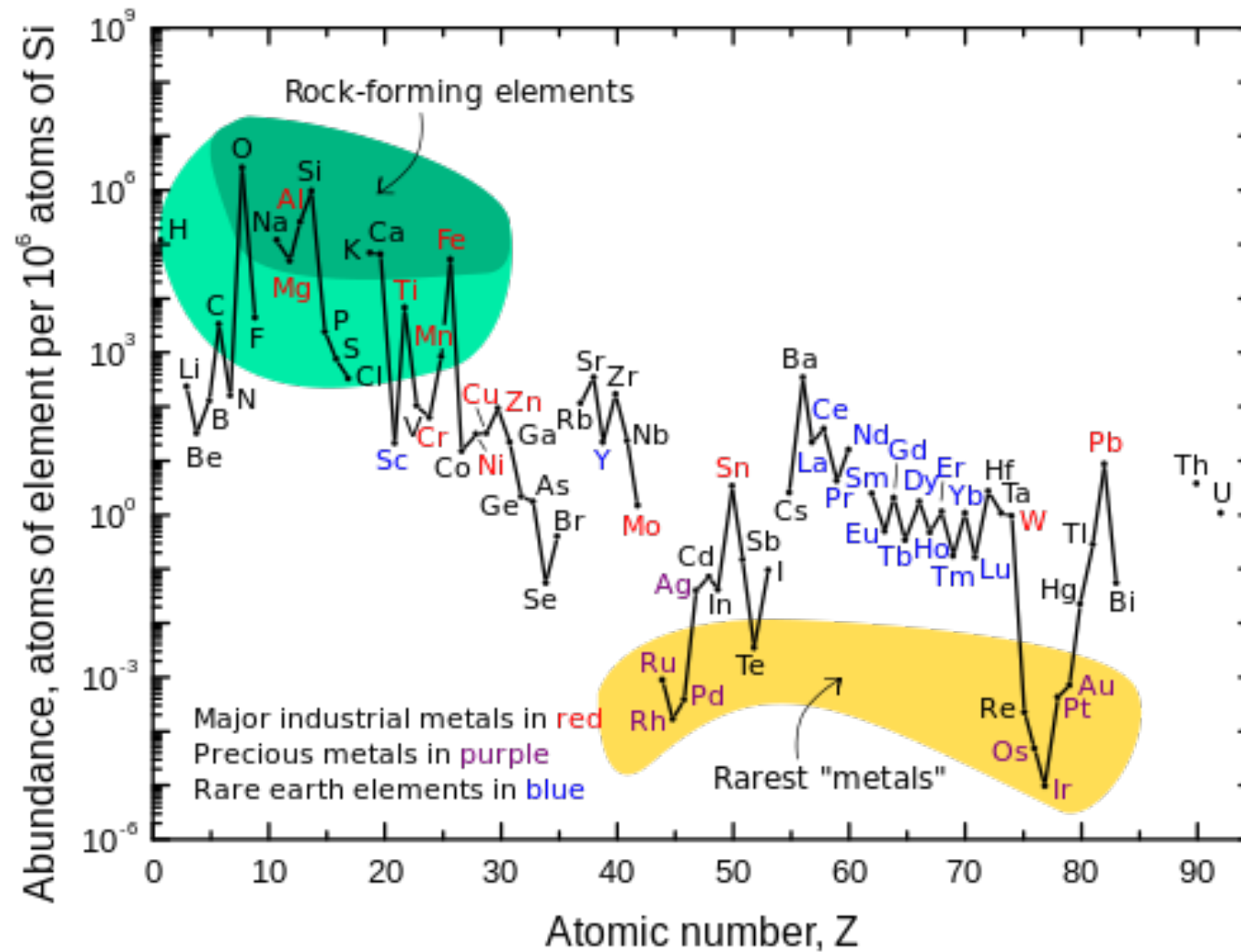
3. Reinforcement learning

4. Evolutionary learning

Three elements

Learning task = {1. Representation} \oplus {2. Evaluation} \oplus {3. Optimization}

Abundance, crust, ...



Contents

- Background {artificial intelligence}
- Evolutionary learning
- Artificial neural networks
- Applications
- [RDFSEARCH, AMADEUS]

Approaches

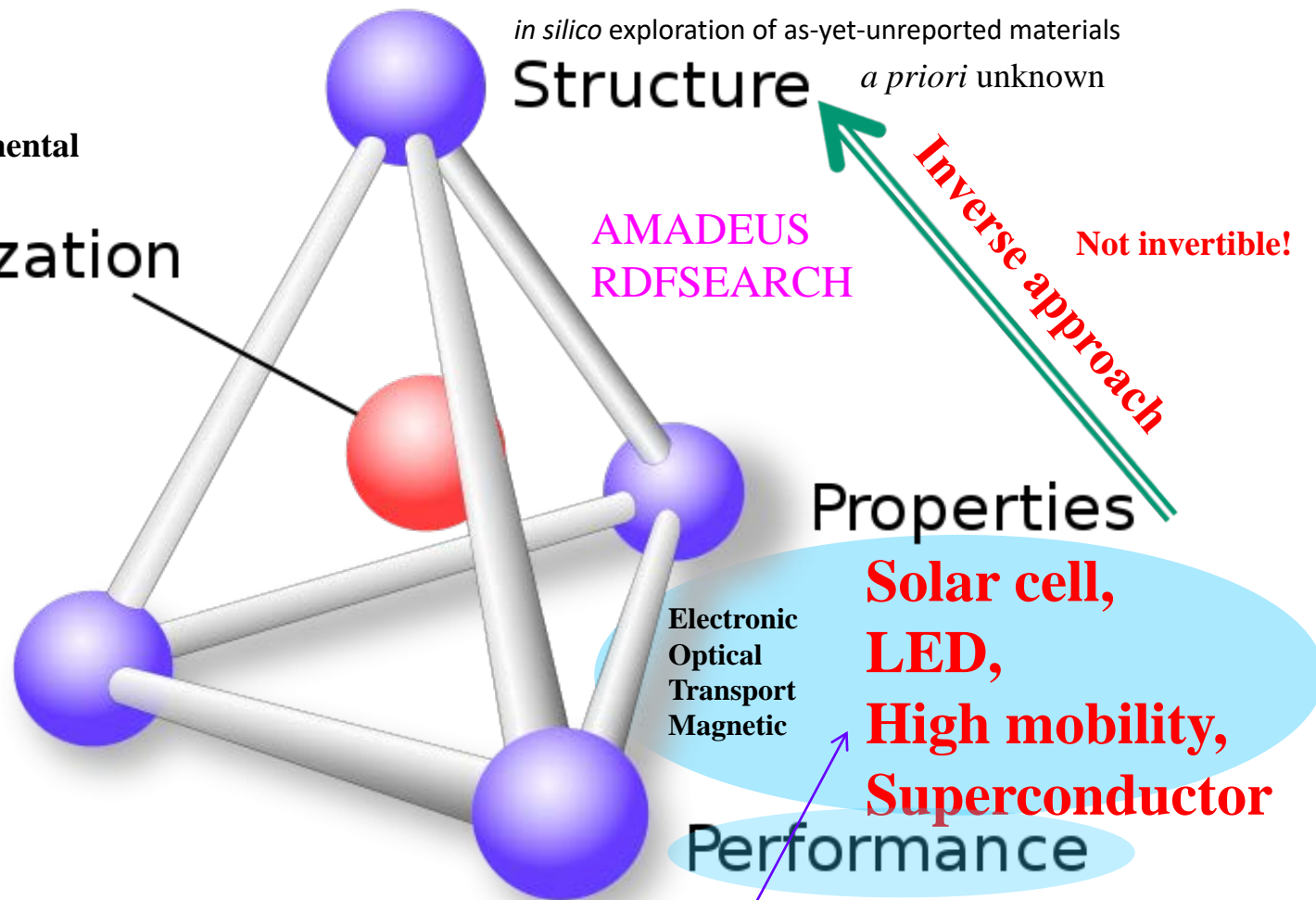
- Known materials → High-throughput calculations [first-principles]
- Known materials → encoding → regression/classification models → prediction for existing and/or known materials → first-principles calculations : Materials repurposing/repositioning [artificial neural networks]
- Unknown materials → Generative models[exploration] [[database](#)]
- Unknown materials → Design via global optimization[design] [[tabula rasa](#), first-principles]

The understanding of processing-structure-properties relationships is called the materials paradigm.

Materials Project → computational
Cambridge Structural Database → experimental

Characterization

Processing



in silico exploration of as-yet-unreported materials

Structure

a priori unknown

AMADEUS
RDFSEARCH

Properties

Solar cell,
LED,
High mobility,
Superconductor

Electronic
Optical
Transport
Magnetic

Performance

Topological materials, Chern insulator?

VESTA
VASP
Quantum Espresso
Mechanical stability
Phonopy
Wannier90
EPW
WannierTools
BerkeleyGW
BoltzTrap
PyProcar

Conformational Space Annealing (CSA)

- Based on **genetic algorithm (GA)** → maintain a pool of conformations (bank) and new conformations are obtained by mating (crossover) and mutation (small perturbation)
- Consider only conformational space of local minima (as in **MCM**)
- **Diversity** of the bank is directly controlled by introducing the concept of *distance* between the conformations
- Narrows the search to smaller region while **maintaining diversity of sampling**
- Key idea : **GA + annealing in conformational space**

Protein structure prediction (modeling) {CASP experiments}

X-ray data → crystal structure determination

NMR → molecule structure determination

Bioinformatics

LJ clusters

Travelling salesman problem

Ab initio materials design (super functional materials)

Pathway sampling (Action-CSA)

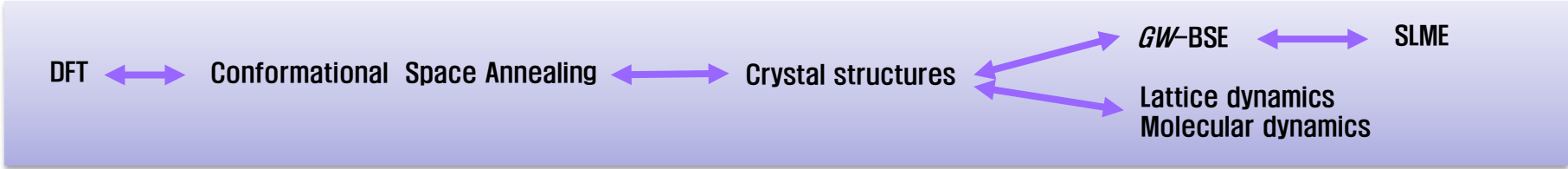
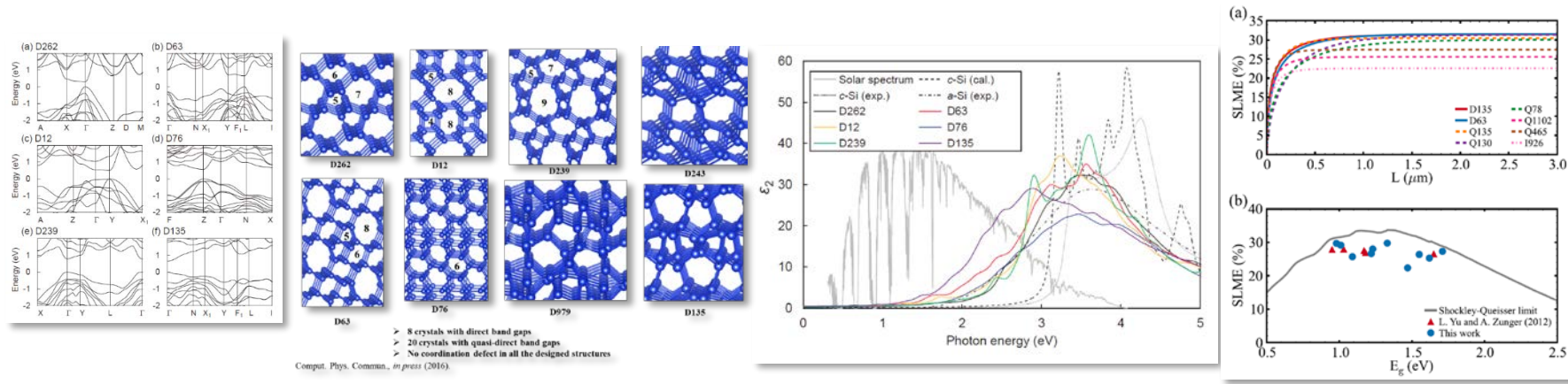
.....



Ab initio materials design using conformational space annealing (AMADEUS)

Atomic configuration → Electronic structure (1)

Electronic structure → Atomic configuration (2)



Applications

- **Solar cell (Si)**: Phys. Rev. B **90**, 115209 (2014), Sci. Rep. **5**, 18086 (2015), Comp. Phys. Commun. **203**, 110 (2016).
- **LED (C)**: Phys. Rev. B **93**, 085201 (2016).
- **High-mobility [green phosphorus] (P)**: J. Phys. Chem. Lett. **8**, 4627 (2017).
- **Topological materials (Si, C, $\text{Ag}_2\text{Se}_{0.5}\text{Te}_{0.5}$)**: J. Phys. Chem. C **123**, 1839 (2019), NPG Asia Materials **9**, e361 (2017), Nanoscale **11**, 5171 (2019).
- **Structure searches (B)**: Sci. Rep. **7**, 7279 (2017), NPG Asia Materials **9**, e400 (2017), arXiv:1902.08390
- **Superconductivity (Si)**: Phys. Rev. Lett. **120**, 157001 (2018).
- **Topological superconductivity (Si)**: Phys. Rev. Mat. **5**, 104802 (2021); Phys. Rev. B **107**, 115127 (2023).

Contents

- Background {artificial intelligence}
- Evolutionary learning
- Artificial neural networks
- Applications
- [RDFSEARCH, AMADEUS]

ThisPersonDoesNotExist.com uses AI to generate endless fake faces

This **X** does not exist



Materials Project → computational
Cambridge Structural Database → experimental

- [This X Does Not Exist](#)
- [Artificial Intelligence Songwriter – These Lyrics Do Not Exist](#)
- [This Chemical Does Not Exist](#)
- [thisartworkdoesnotexist.com \(512×512\)](#)
- [This Person Does Not Exist](#)
- [This Music Video Does Not Exist](#)

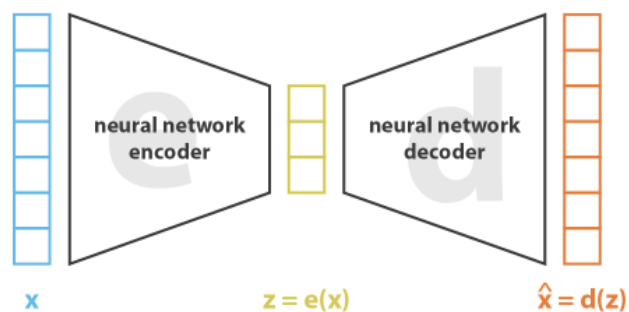
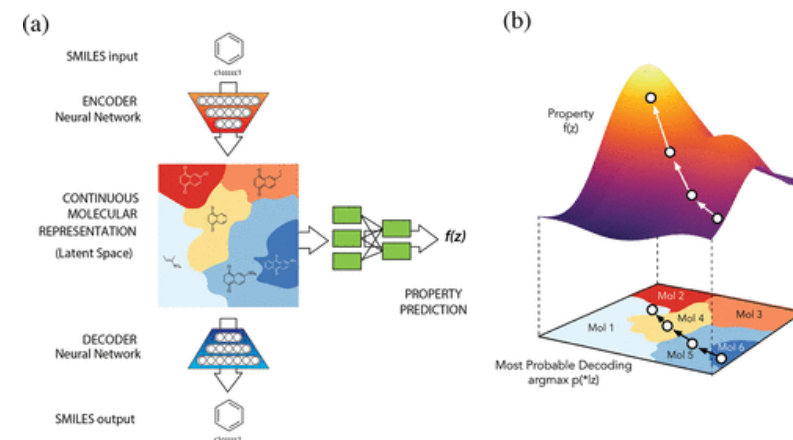
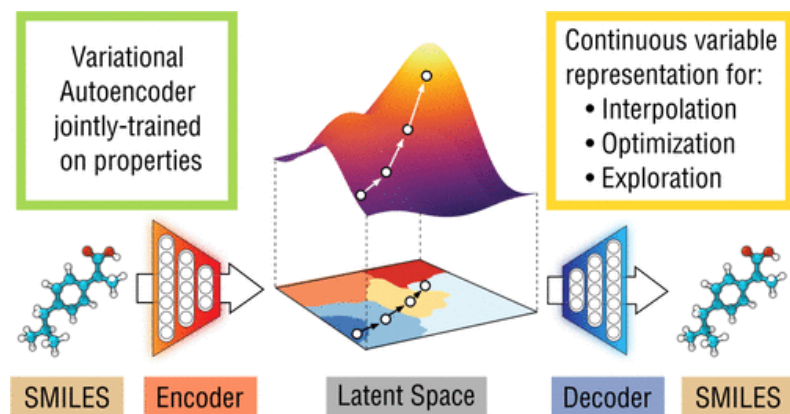


Neural Style Transfer
Gram matrix

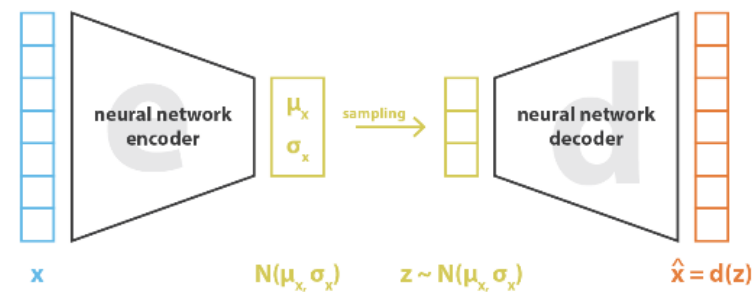
Boltzmann machine, GAN, VAE, RNN, PixelCNN

Generative models

Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules

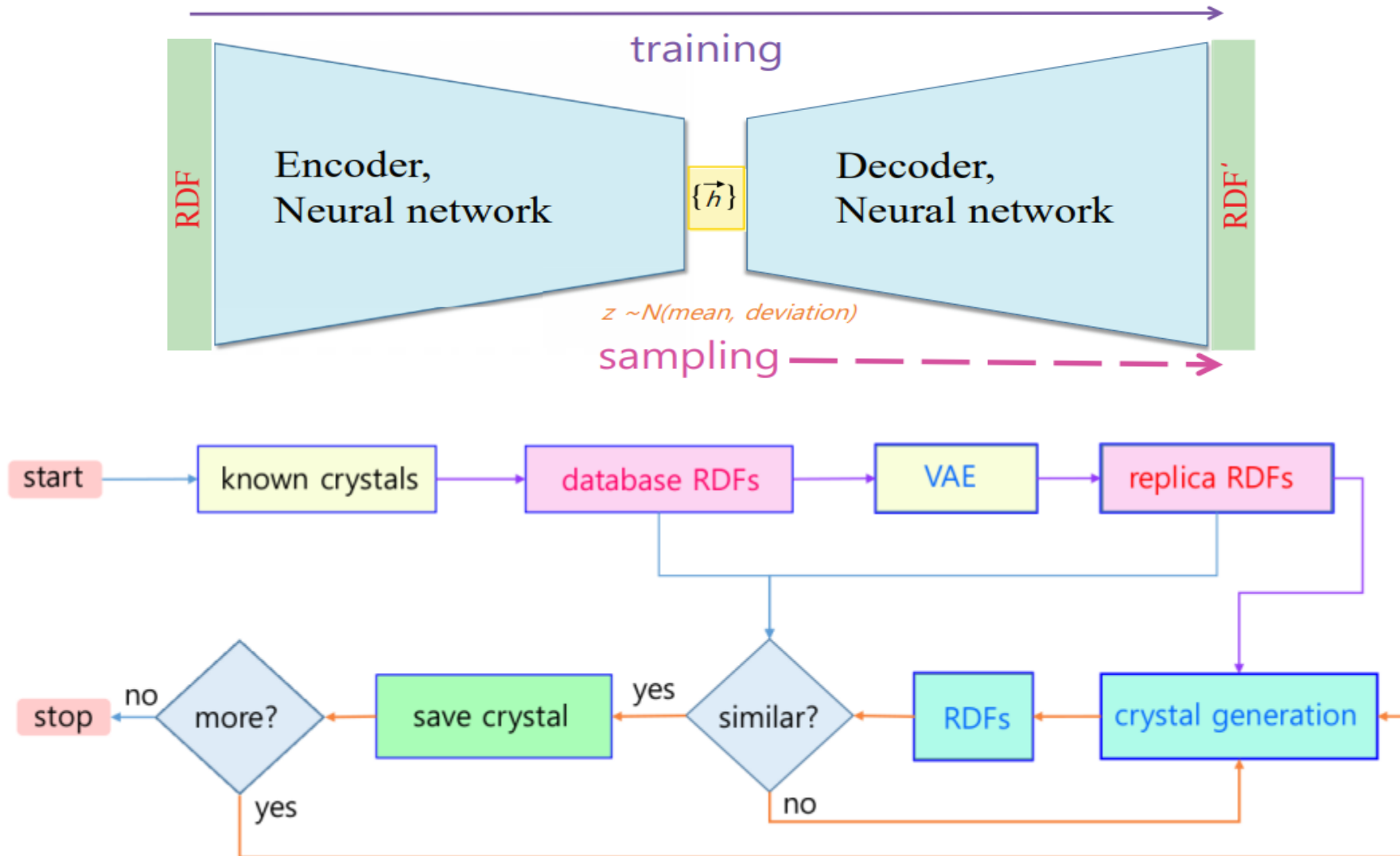


$$\text{loss} = \|x - \hat{x}\|^2 = \|x - d(z)\|^2 = \|x - d(e(x))\|^2$$



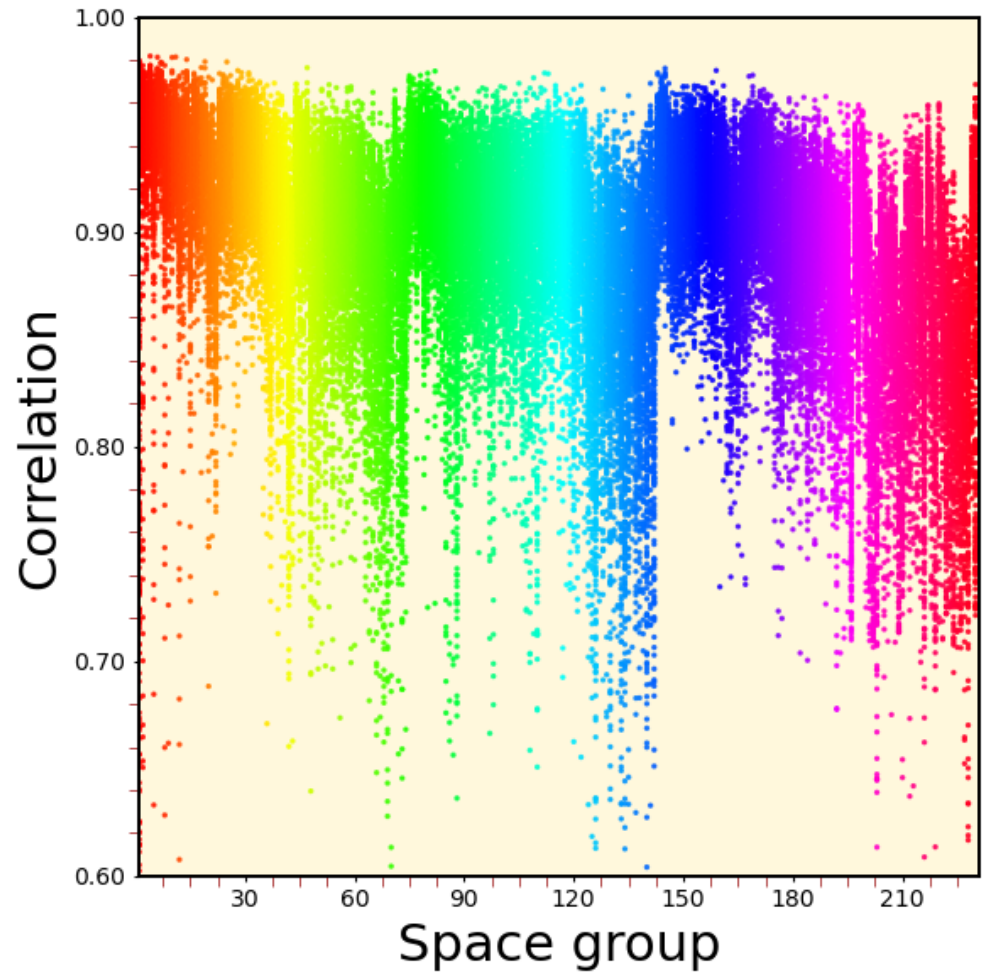
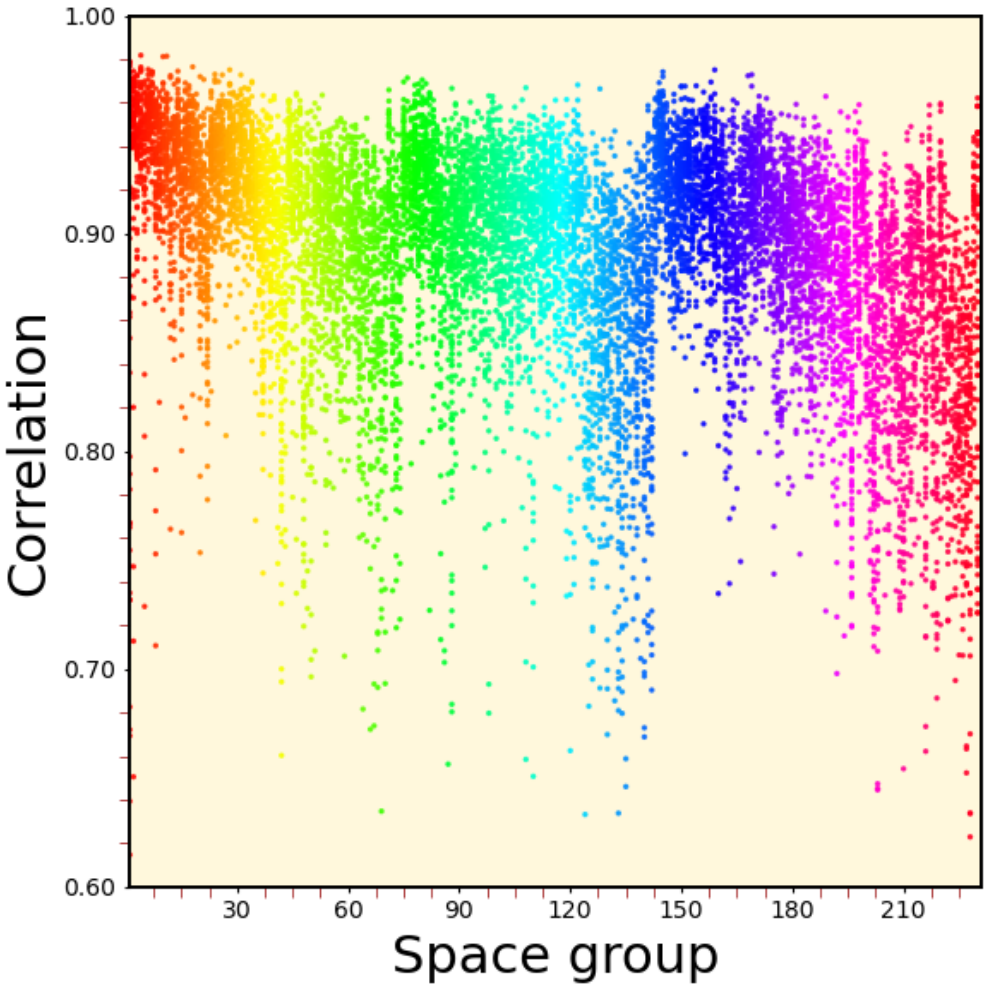
$$\text{loss} = \|x - \hat{x}\|^2 + \text{KL}[N(\mu_x, \sigma_x), N(0, 1)] = \|x - d(z)\|^2 + \text{KL}[N(\mu_x, \sigma_x), N(0, 1)]$$

Crystal structure prediction in a continuous representation space

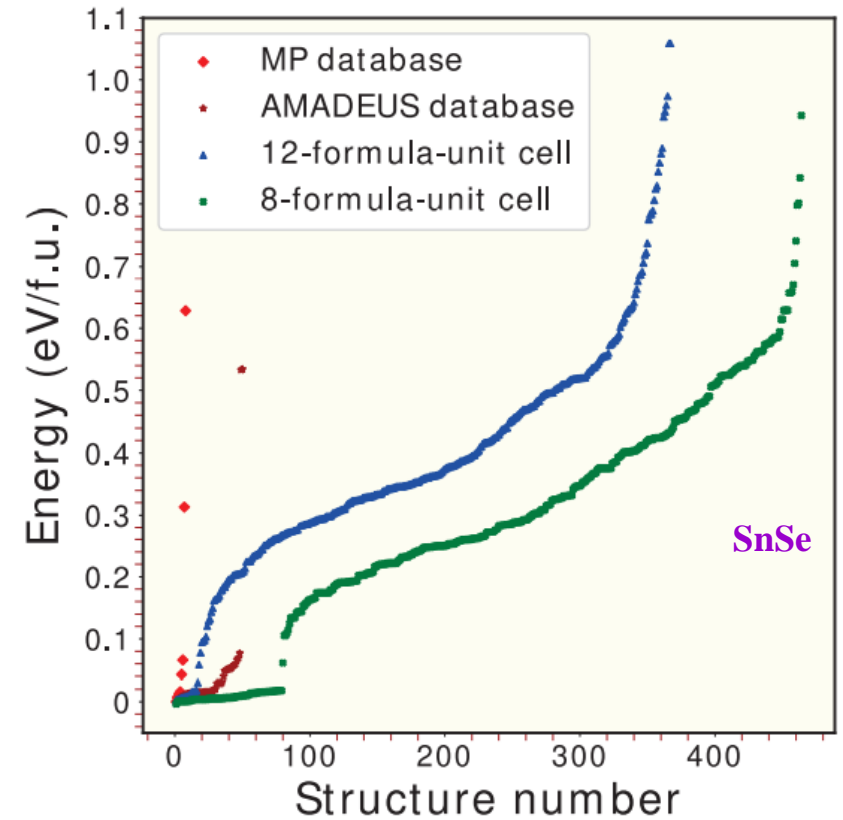
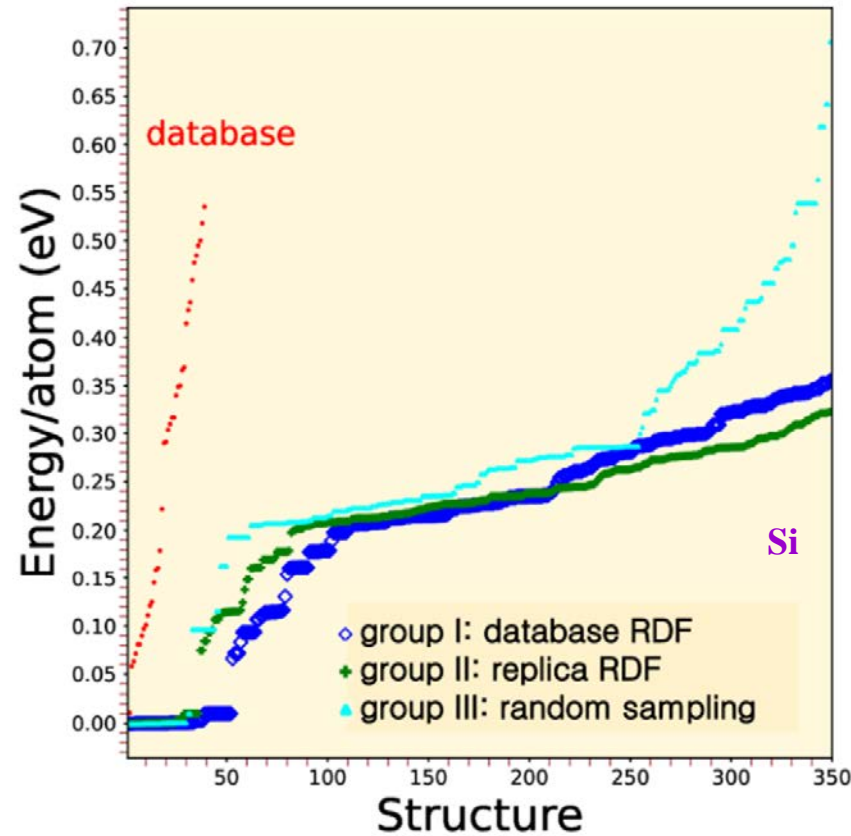
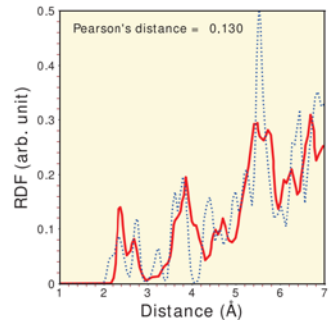
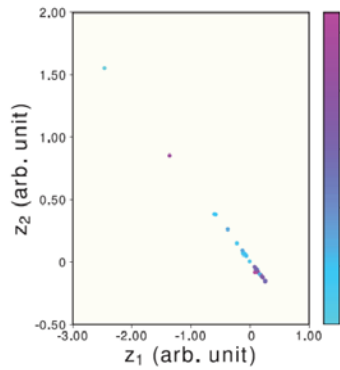
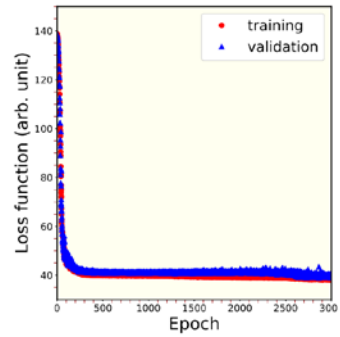




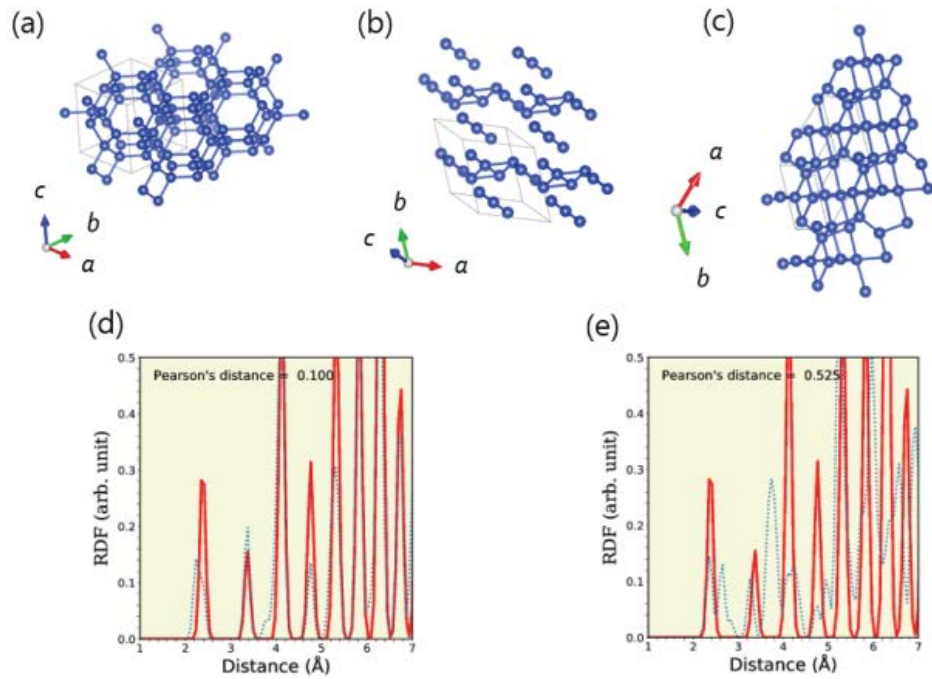
Crystal structure prediction in a continuous representation space



Data-driven continuous representation



$$r_P = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^m (y_i - \bar{y})^2}}$$



Pearson's distance between $C2/m$ and $C2/m^*$ = 0.102

$C2/m^* > C2/m$, 40 meV/atom

$C2/m$: 2.28 g/cm³, 0.13 eV

$C2/m^*$: 2.39 g/cm³, 0.28 eV

Microexplosion experiment: $t32$, $t32^*$, $m32$, $m32^*$ [\sim BC8]

$C2/m^*$ is stable by 3-10 meV/atom than these four phases.

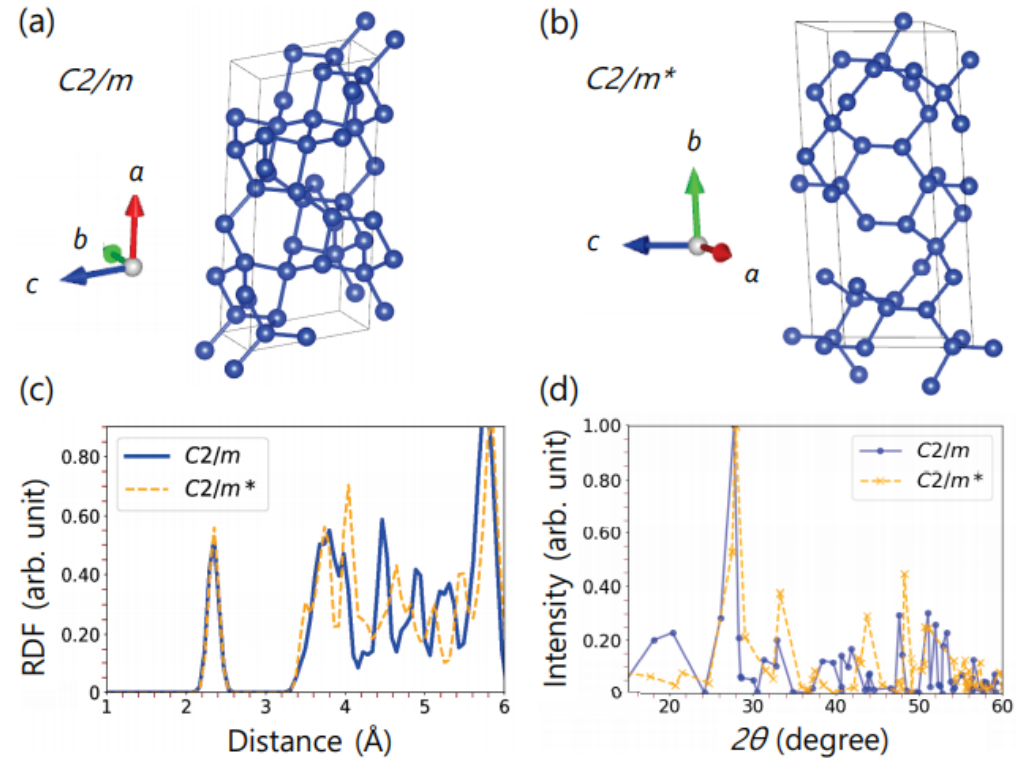
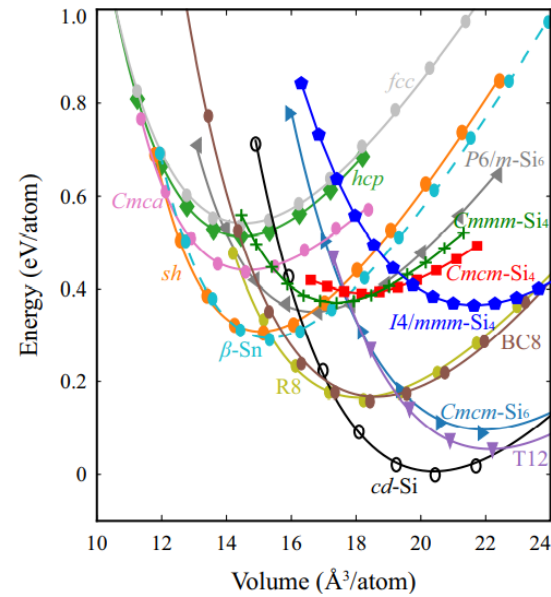
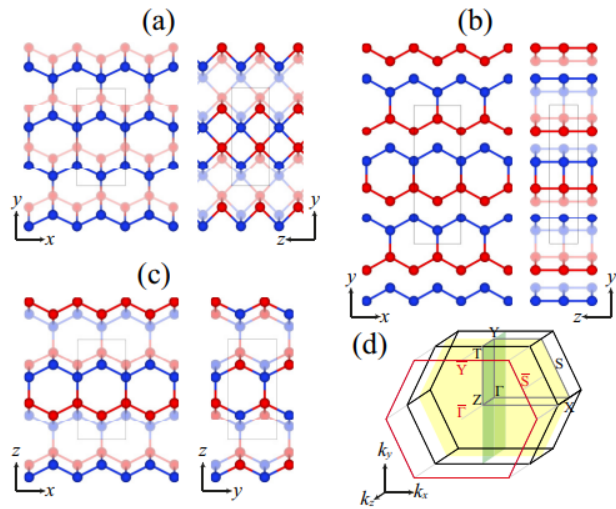


Table 1: Lattice parameters and Wyckoff positions for three Si allotropes, $Cmmm$ -Si, $C2/m$ -Si, and $C2/m^*$ -Si.

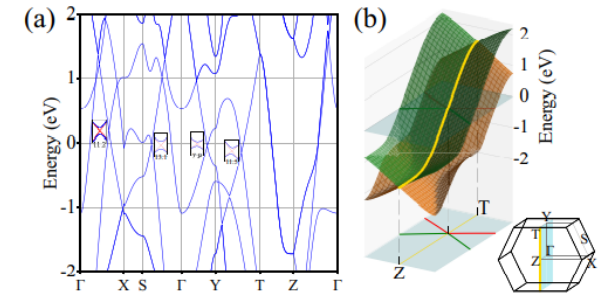
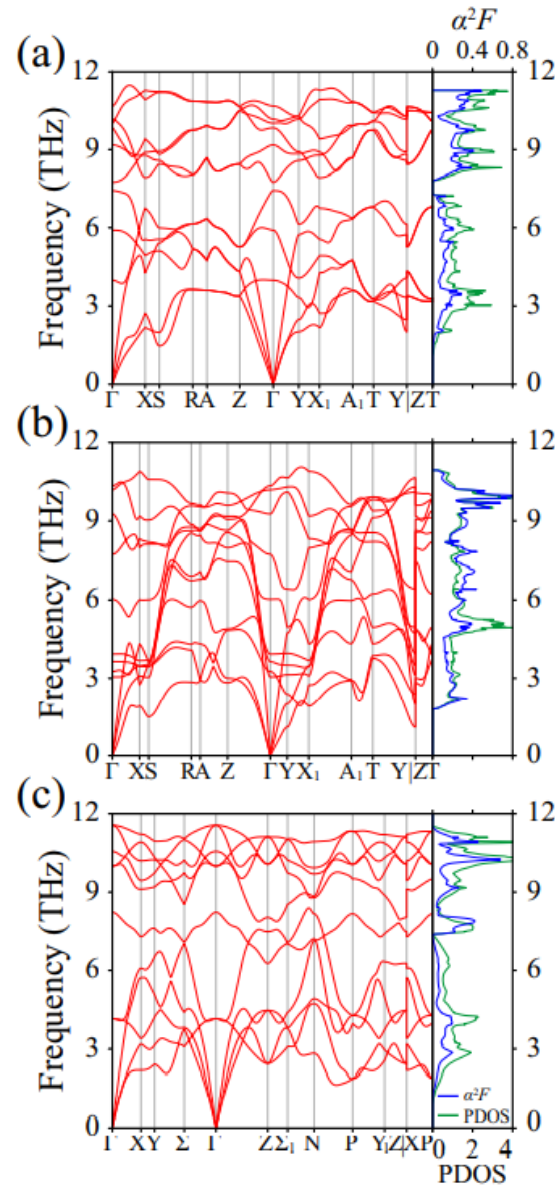
Allotrope	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	Wyckoff positions
$Cmmm$ -Si	8.21	6.59	5.20	90.0	90.0	90.0	4 1 $mm2$ (0, 0.50000, 0.75536)
							8 m ..2 (0.25000, 0.25000, 0.72747)
							4 g 2 mm (0.86311, 0.00024, 0.99999)
$C2/m$ -Si	13.73	3.82	6.29	90.0	83.3	90.0	4 i m (0.94460, 0, 0.87347)
							4 i m (0.44097, 1.00000, 0.65322)
							4 i m (0.78744, 0, 0.05727)
							4 i m (0.27303, 1.00000, 0.58547)
$C2/m^*$ -Si	3.77	16.56	6.93	90.0	133.7	90.0	4 i m (0.08137, 0, 0.86120)
							8 j 1 (0.07598, 0.87869, 0.36283)
							4 g 2 (0.00029, 0.29131, 0.00058)

Ab initio prediction of topological superconductivity in metallic Si allotropes



Allotrope	C_{11}	C_{22}	C_{33}	C_{44}	C_{55}	C_{66}	C_{12}	C_{13}	C_{23}
<i>Cmc</i> -Si ₄	147.10	119.16	109.73	33.59	8.02	33.58	60.97	63.10	8.92
<i>Cmmm</i> -Si ₄	205.60	145.33	300.06	26.29	69.00	26.46	77.07	-3.59	25.81
<i>I4/mmm</i> -Si ₄	108.73	108.73	140.78	36.93	36.93	36.23	68.13	40.17	40.17

Allotrope	Pressure (GPa)	V (Å ³ /atom)	$N(0)$ (states/Ry/atom/spin)	λ	ω_{log} (K)	T_c (K)
<i>sh</i> -Si	15	13.50	2.44	0.660	250	7.9
<i>P6/m</i> -Si ₆	0	16.80	2.40	0.799	263	12.2
<i>Cmc</i> -Si ₄	0	18.18	1.87	0.462	253	2.2
<i>Cmmm</i> -Si ₄	0	17.38	2.52	0.808	238	11.4
<i>I4/mmm</i> -Si ₄	0	21.64	1.87	0.417	234	1.2

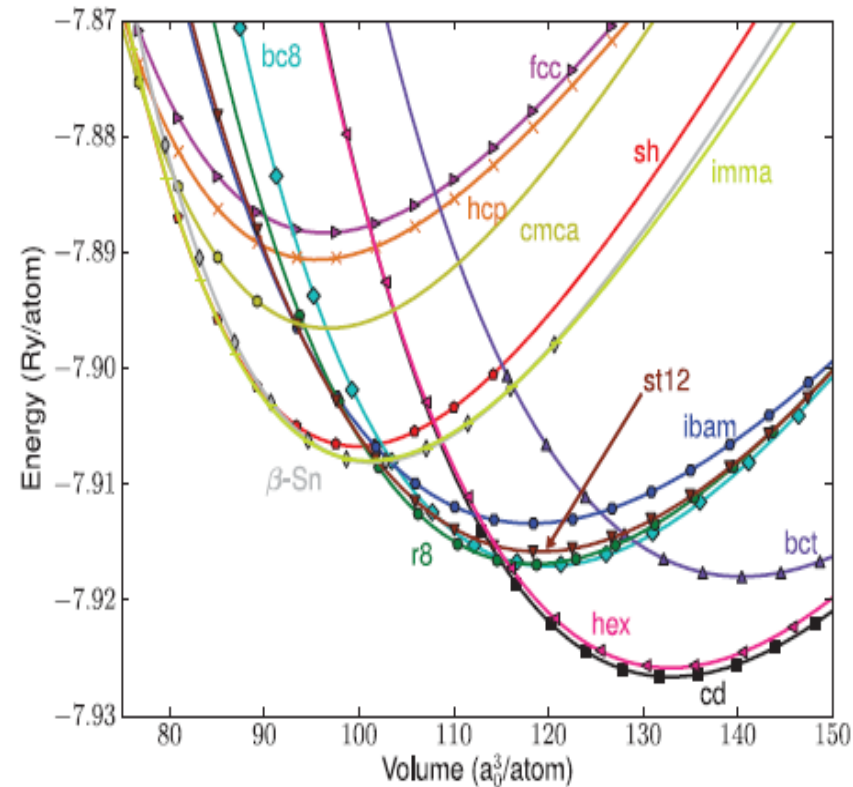


Contents

- Background {artificial intelligence}
- Evolutionary learning
- Artificial neural networks
- Applications
- [RDFSEARCH, AMADEUS]

Reported crystalline phases of Si under pressure

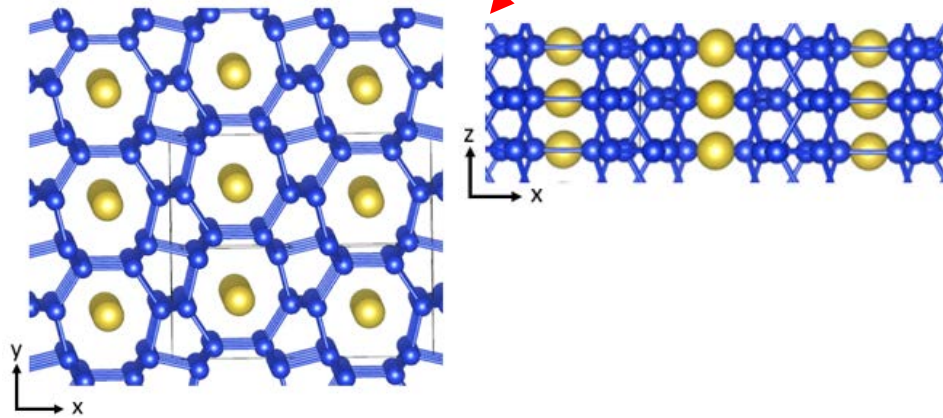
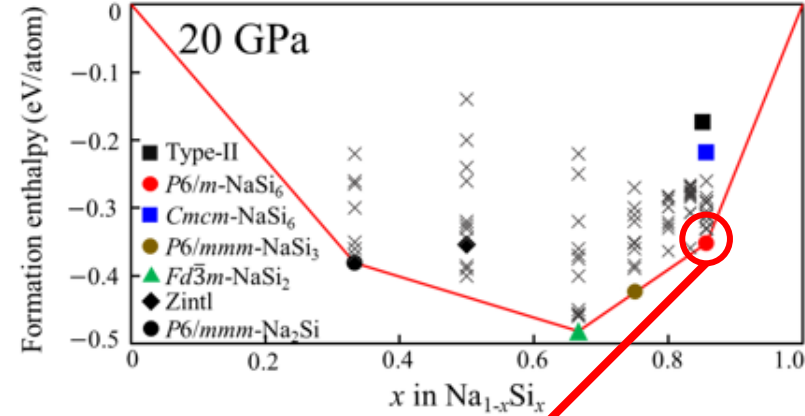
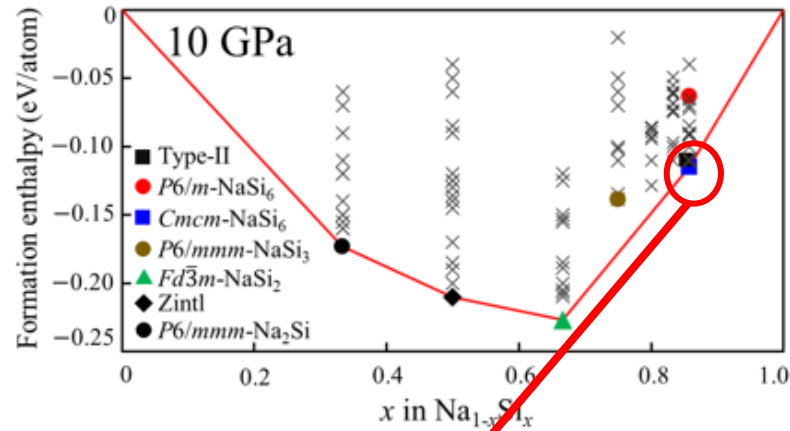
Si-I (cubic diamond)
Si-II (β -Sn) : metal
Si-III (BC8) : narrow gap
Si-IV (hexagonal diamond)
Si-V (simple hexagonal) : metal
Si-VI (*Cmca*-Si) : metal
Si-VII (hcp) : metal
Si-VIII : not identified yet
Si-IX : not identified yet
Si-X (fcc) : metal
Si-XI (*Imma*-Si) : metal
Si-XII (R8) : indirect gap 0.24 eV
Si-XIII (T12) : indirect gap
BT8, ST12 : indirect gap
clathrates (type-I, type-II), *Cmcm*-Si₆



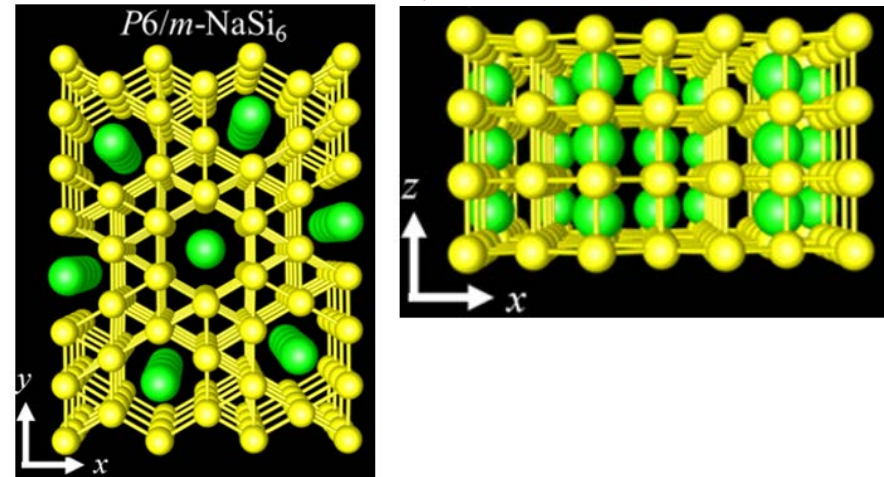
B.D. Malone, M.L. Cohen, Phys. Rev. B (2012).

- Si allotropes with direct band gaps have not been synthesized.
- High pressure metallic phases (β -Sn, sh, hcp, fcc) do not maintain their crystal structures after pressure release.
- Metallic Si phases at ambient pressure had not been reported before April 11, 2018.

NaSi₆ clathrate at high pressure



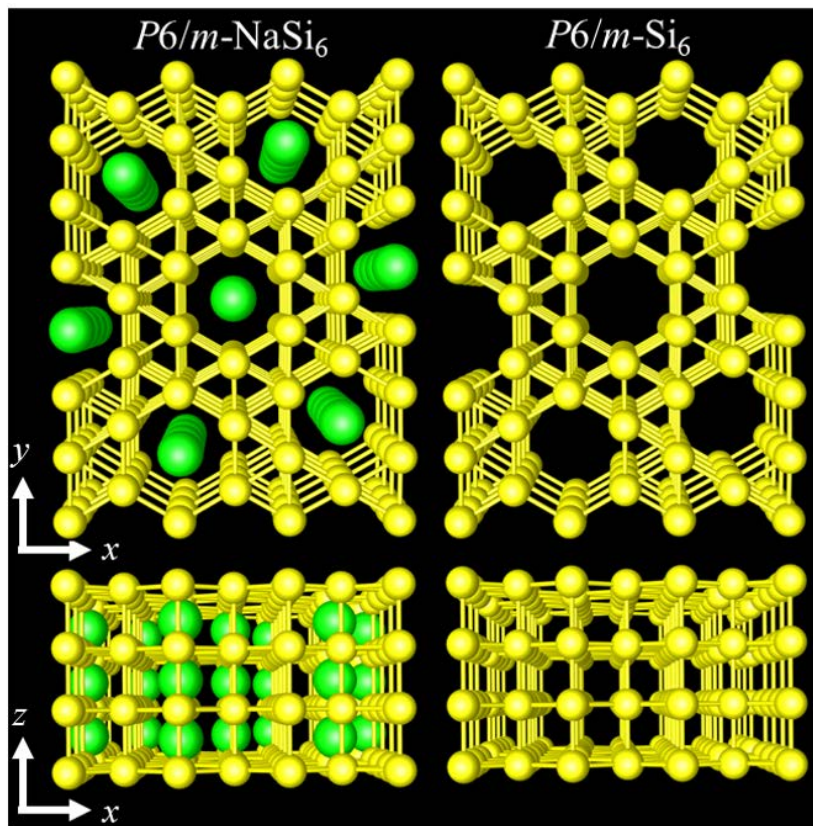
➤ ***Cmcm*-NaSi₆ at 10 GPa**
[Cryst. Growth Des. **13**, 303 (2013)]



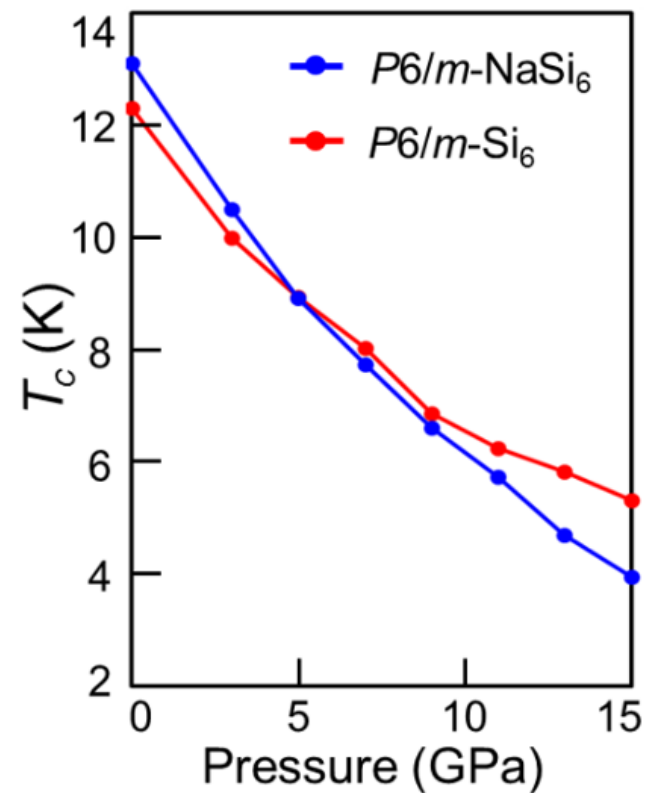
➤ ***P6/m*-NaSi₆ at 20 GPa**
➤ **Transition pressure from *Cmcm*-NaSi₆ to *P6/m*-NaSi₆ : 12.4 GPa**

Superconducting $P6/m$ - NaSi_6 and $P6/m$ - Si_6 clathrates

(first proposed)



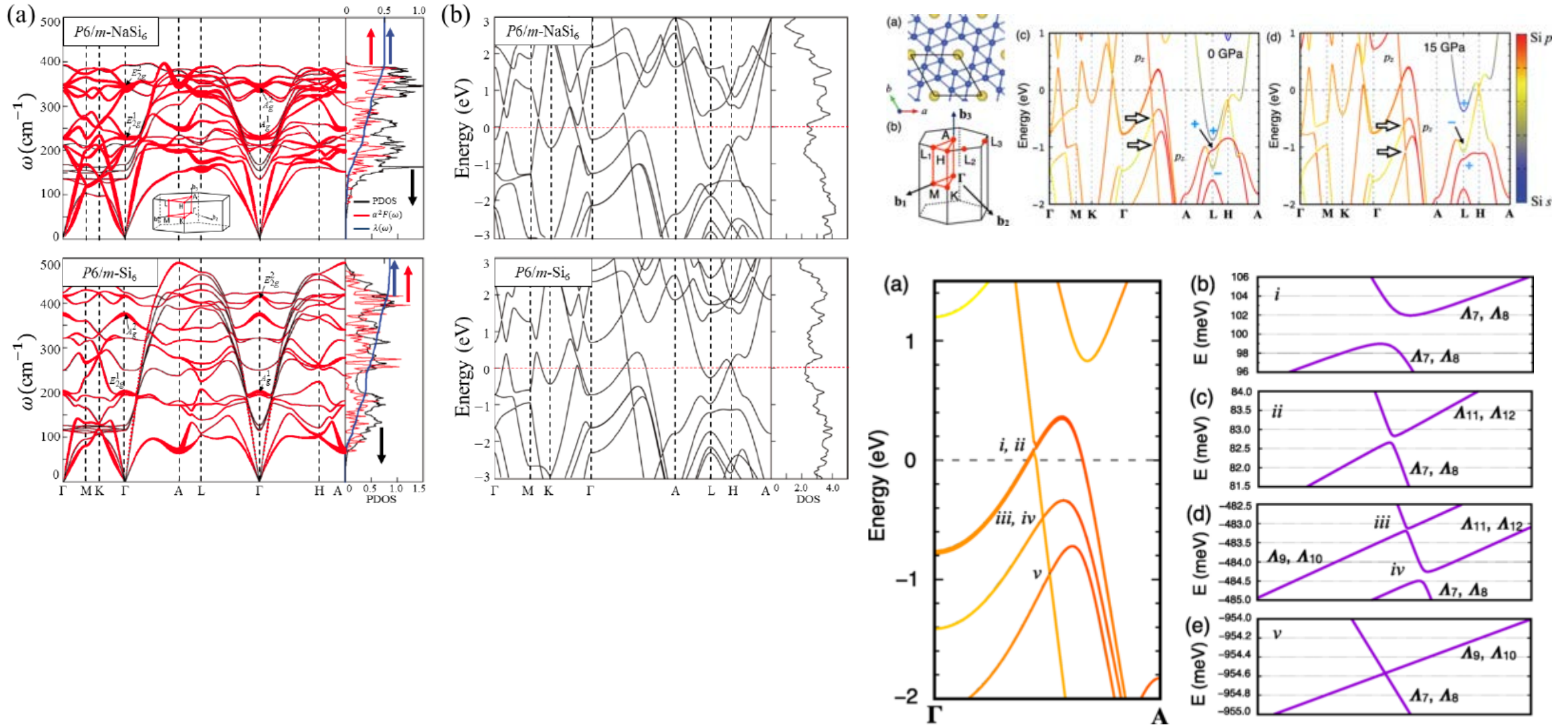
Pressure dependence of T_c



- $P6/m$ space group
- Simple hexagonal with open channels
- Na atoms along open channels
- Stable at ambient pressure
- Metallic

- Superconducting at ambient pressure
 - $T_c = 13.1$ K for $P6/m$ - NaSi_6
 - $T_c = 12.2$ K for $P6/m$ - Si_6

Na_{1-x}Si_x



A new superconducting open-framework allotrope of silicon at ambient pressure

Allotrope	Pressure (GPa)	$(\text{\AA}^3/\text{Si-atom})$ (states/Ry/Si-atom/spin)			(K)	(K)
		V	$N(0)$	λ	ω_{\log}	T_c
$P6/m\text{-NaSi}_6$	0	18.4	2.82	0.897	225	13.1
	15	16.1	2.60	0.498	331	4.0
$P6/m\text{-Si}_6$	0	16.8	2.40	0.799	263	12.2
	15	14.8	2.22	0.556	299	5.3
$sh\text{-Si}$	15	13.5	2.44	0.660	250	7.9

$T_c=4\text{--}8$ K
 high-pressure β -Sn phase of Si,
 high-pressure sh phase of Si,
 type-I clathrate of $(\text{Na, Ba})_x\text{-Si}_{46}$,
 $\text{Ba}_8\text{Si}_{46}$

Design of single-layer metasurface filter by CSA for 5G mm-wave communications

It is anticipated that 28 GHz and 38 GHz mmWave bands will play a key role in 5G deployments and beyond-5G systems.

[pre-assigned frequency response characteristic functions required in the 5G frequency band]

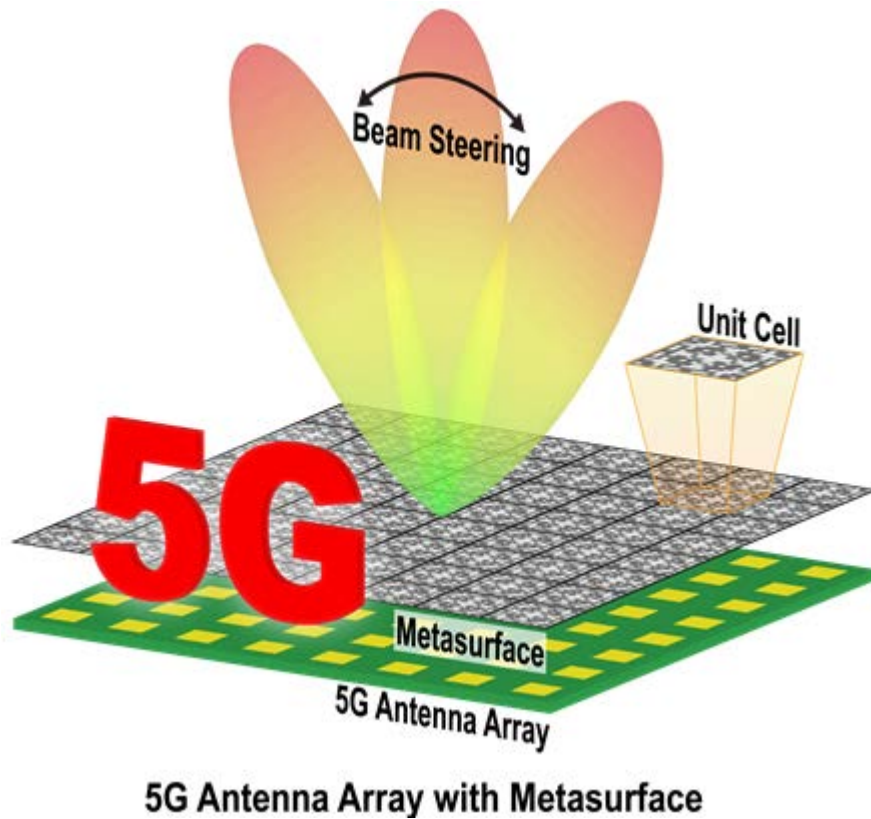
[low in-band loss and high out-of-band suppression] [the widest bandwidth]

[angular stability] [polarization insensitivity]

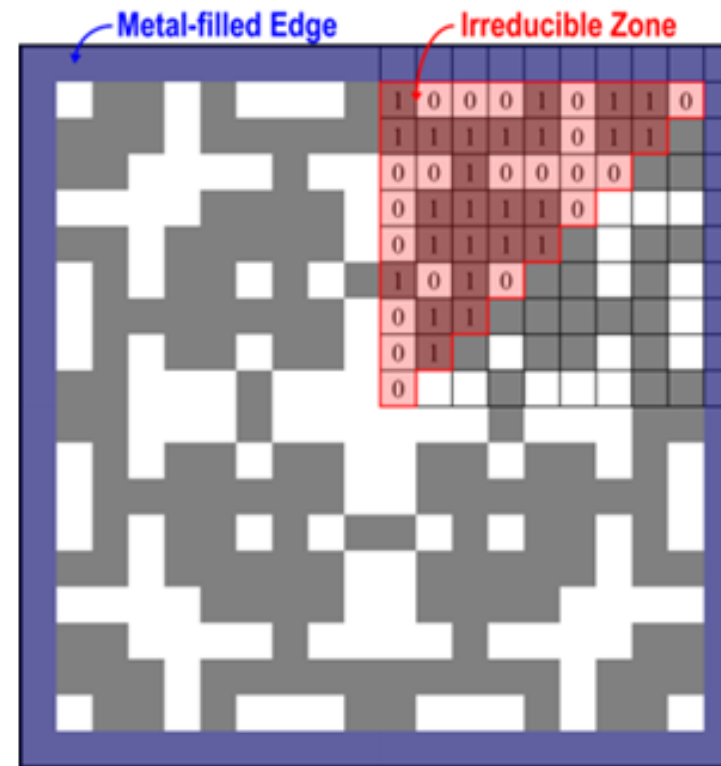
Objective function

$$f(\vec{x}) = \sum_{k=1}^N |T_{target}(v_k) - T_{EM}(v_k; \vec{x})|.$$

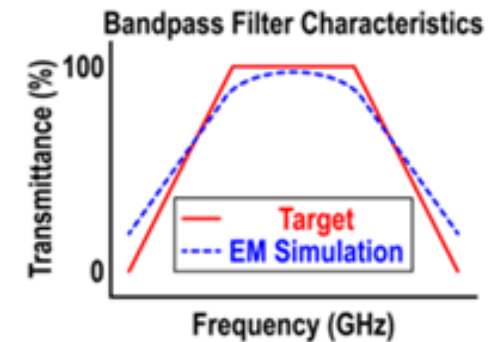
The more pixels the better.



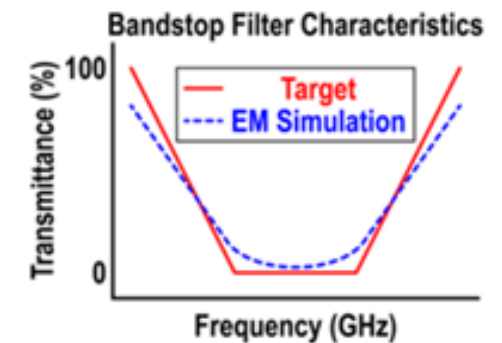
Coexistence of 5G With Fixed Satellite Service at 28 GHz



50 μm thick flexible polyethylene terephthalate film
6 μm thick copper foil

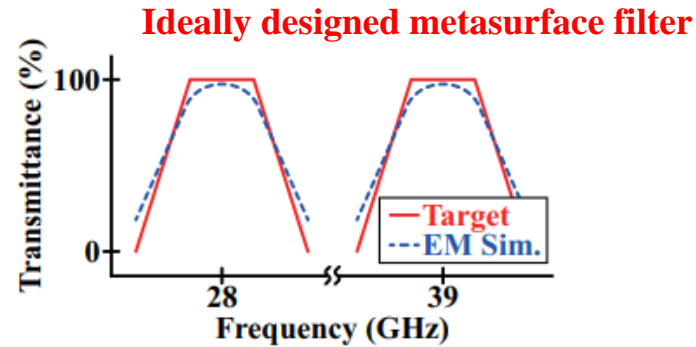
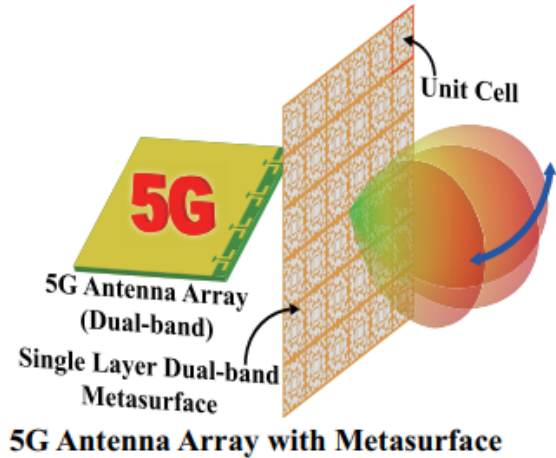


(b)



(c)

Design of dual-band single-layer metasurfaces for millimeter-wave 5G communication systems



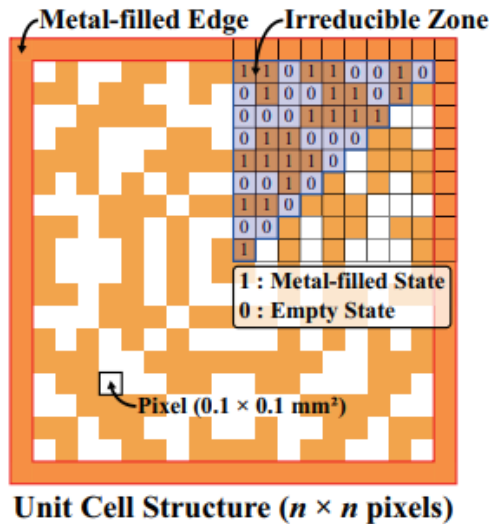
Dual Band-pass Filter Characteristic

Objective function

$$f(\vec{x}) = \sum_{k=1}^N |T_{target}(v_k) - T_{EM}(v_k; \vec{x})|.$$

generic flexible single-layer pixel-based unit-cell

50 μm
6 μm



Transmittance of the surface of a material is its effectiveness in transmitting **radiant energy**. It is the fraction of incident electromagnetic power that is transmitted through a sample, in contrast to the **transmission coefficient**, which is the ratio of the transmitted to incident **electric field**.

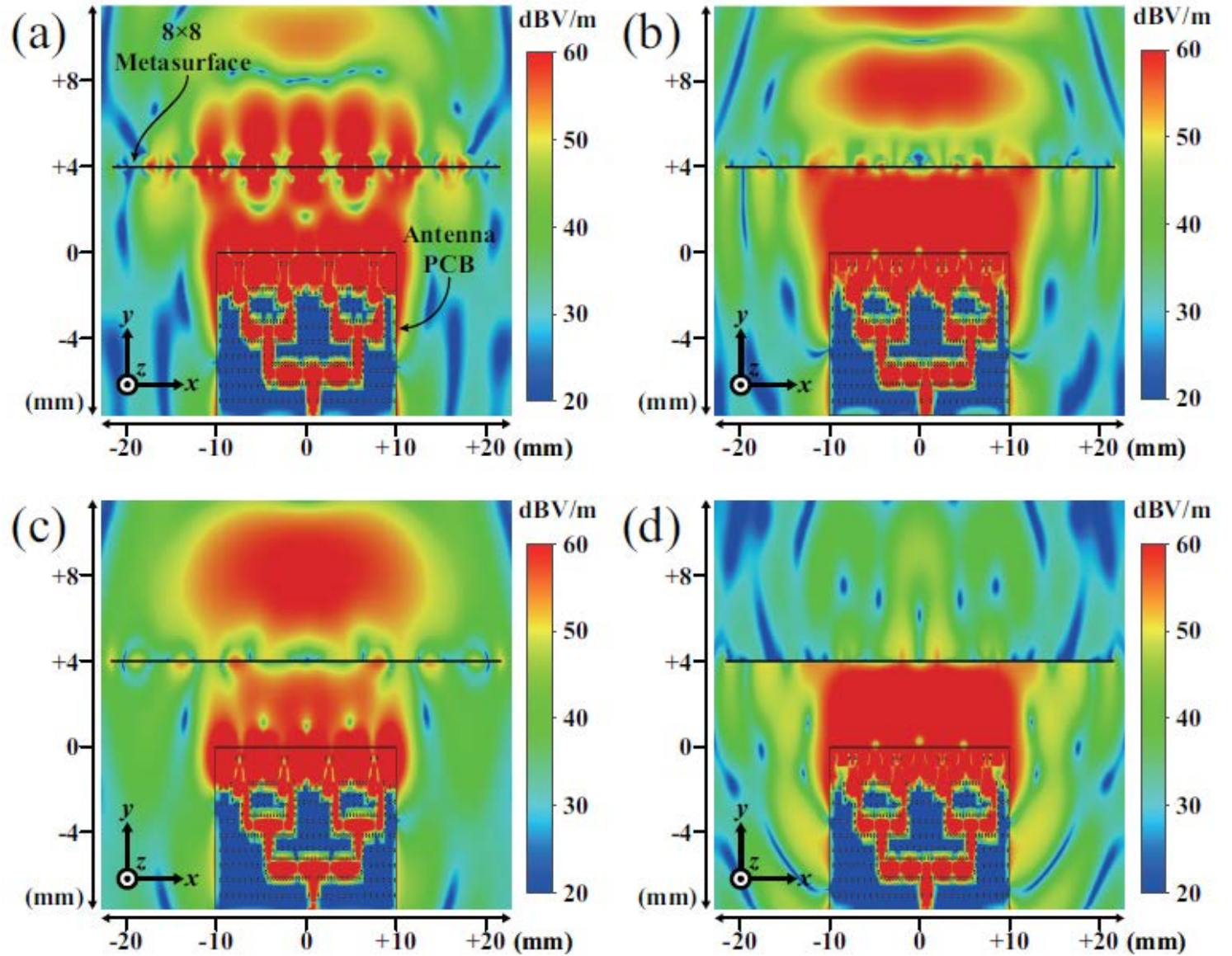
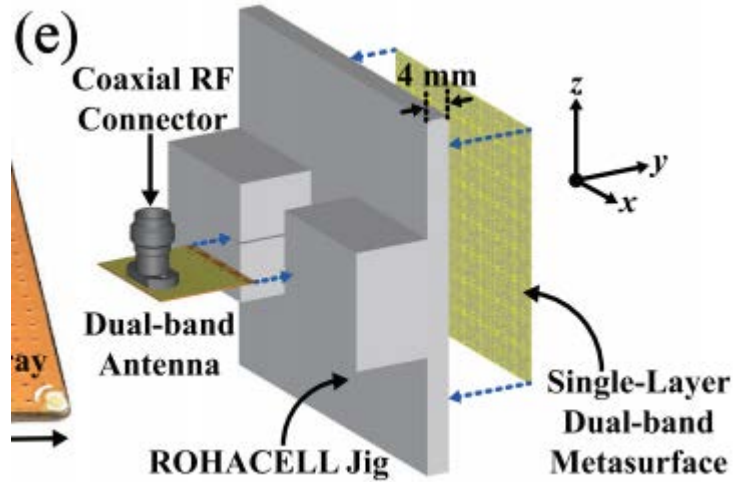
Approaches reported in the literature, which can be classified into three types of dual-band solutions—multilayer metasurfaces, fractal elements, and multi-resonant metasurfaces—have been utilized for enabling interference alleviation techniques. Typical metasurfaces rely on the introduction of multiple resonators that resonate separately by frequency.

The **decibel** (symbol: **dB**) is a relative **unit of measurement** equal to one tenth of a **bel (B)**.

Table: Gain / loss in decibels

Loss/gain as a ratio	Loss/gain in decibels	Loss/gain as a ratio	Loss/gain in decibels
$\frac{P_{\text{output}}}{P_{\text{input}}}$	$10 \log \frac{P_{\text{output}}}{P_{\text{input}}}$	$\frac{P_{\text{output}}}{P_{\text{input}}}$	$10 \log \frac{P_{\text{output}}}{P_{\text{input}}}$
1000	30 dB	0.1	-10 dB
100	20 dB	0.01	-20 dB
10	10 dB	0.001	-30 dB
1 (no loss or gain)	0 dB	0.0001	-40 dB

Simulated electric field intensity distributions
(x-component of electric field) of prototype configurations.
5G antenna array with 28/39 GHz dual BPF
 at (a) 29.5 GHz (passband) and (b) 39.0 GHz (passband).
5G antenna array with 28 GHz BFP/39 GHz BSF
 at (c) 27.0 GHz (passband) and (d) 39.0 GHz (stop band).





Contents

- Background {artificial intelligence}
- Evolutionary learning
- Artificial neural networks
- Applications
- [RDFSEARCH, AMADEUS]

- <http://events.kias.re.kr/h/ka2022/?pageNo=4483>
- <https://github.com/inholeegithub/winter2022>

The screenshot shows the KIAS website with a navigation menu (Home, Program, Past Schools, Registration) and a large red banner. The banner text reads: "The 2nd KIAS Electronic Structure Calculations Winter School" and "January 18(Tue) ~ 21(Fri), 2022 | Online".

[강사: 이인호 박사\(KRISS\) 1/2 제2회 KIAS 전자구조계산 겨울학교 - YouTube](#)

[강사: 이인호 박사\(KRISS\) 2/2 제2회 KIAS 전자구조계산 겨울학교 - YouTube](#)

The screenshot shows the GitHub repository page for 'inholeegithub/winter2022'. It includes a file browser with folders 'amadeus', 'rdfsearch', and 'README'. The README content is as follows:

README

In the winter of 2022, I am pleased to share materials on the two crystal structure exploration/design methods, for the purpose of the practice along with the presentation at the electronic structure calculation conference held at KIAS.

The first method is AMADEUS and the second method is RDFSEARCH. AMADEUS is a combination of the first-principles electronic structure calculation and the global optimization method. RDFSEARCH can generate a set of new crystal structures as a database-based generative model.

I am looking forward to seeing your interest and feedback.

KRISS, January 14, 2022
In-Ho Lee
ihlee@kriss.re.kr

On the right side of the repository page, there are statistics: 0 stars, 1 watching, 0 forks, 0 releases, 0 packages, and a language usage chart showing Fortran (72.0%), Python (22.8%), Shell (2.6%), Jupyter Notebook (2.0%), and Other (0.6%).

Conclusions

- Conformational Space Annealing (CSA)
- AMADEUS (many applications)
- Action-CSA (reaction path search)
- Replica RDFs (VAE)
- Solar cell (Si)
- LED (C)
- High-mobility (P)
- Topological materials (Si, C, $\text{Ag}_2\text{Se}_{0.5}\text{Te}_{0.5}$)
- Topological superconductivity (Si)
- Structural searches (2D, 3D types of B)
- **Metasurfaces** for 5G communication systems (License-out)



