

ХИМИЧЕСКАЯ ИНФОРМАТИКА: ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ

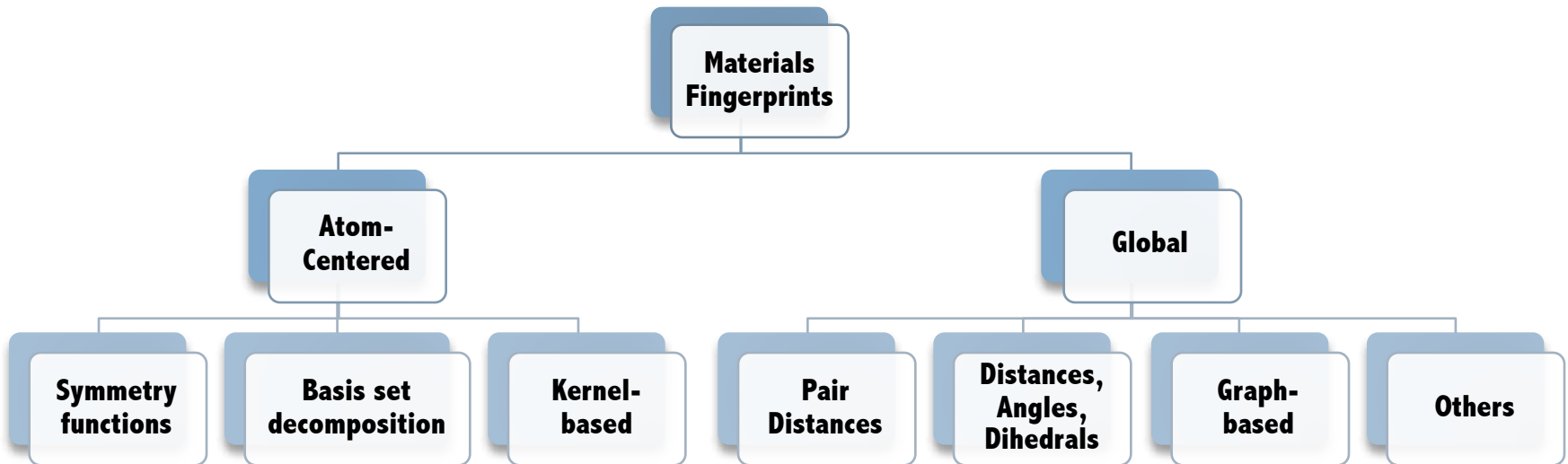
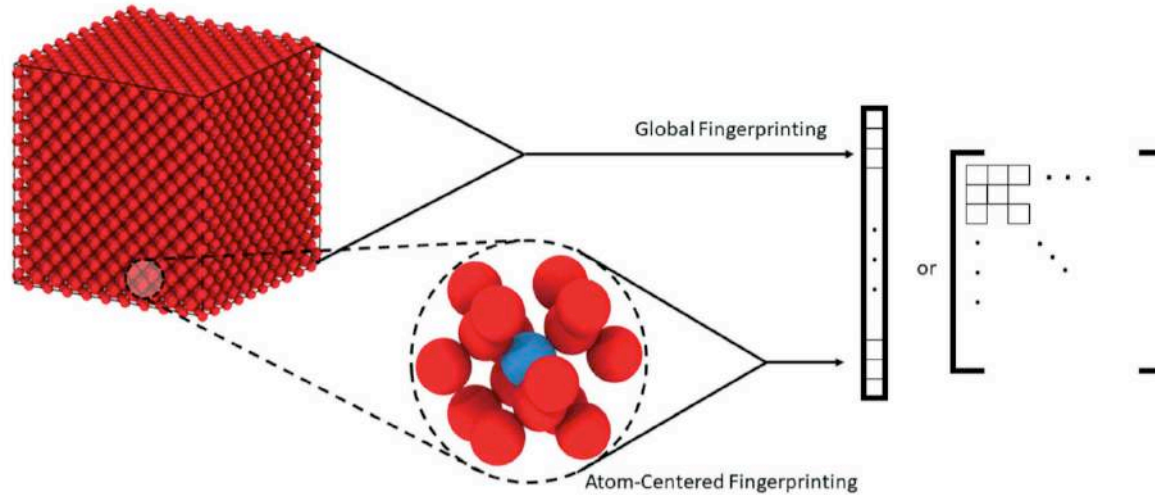
Лекция 8

Информатика материалов
(первая лекция)

МНОГООБРАЗИЕ ЗАДАЧ И ОБЪЕКТОВ



MATERIALS FINGERPRINTS



Reveil et al Classification of spatially resolved molecular fingerprints for machine learning applications and development of a codebase for their implementation Mol Syst Des & Engin (2018)

ATOM-CENTERED SYMMETRY FUNCTIONS

Отпечатки, основанные на определении функции симметрии базируются на суммировании по связям и углам для всех атомов в пределах заданного радиуса

Атомно-центрированные функции симметрии (ACSFs): Behler and Parinello

$$G_i^{rad} = \sum_{j \neq i}^N e^{-\eta(r_{ij}-\mu)^2} f_c(r_{ij})$$

$$f_c(r_{ij}) = \begin{cases} \frac{1}{2} \left[\cos\left(\frac{\pi r_{ij}}{r_c}\right) + 1 \right], & r_{ij} \leq r_c \\ 0, & r_{ij} \geq r_c \end{cases}$$

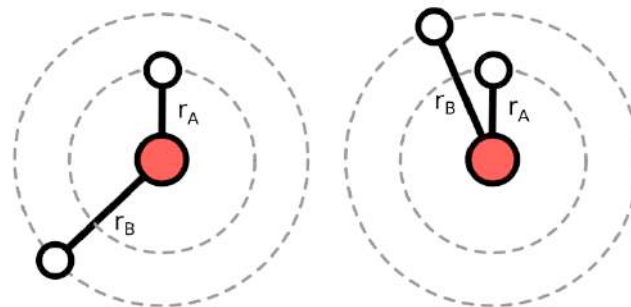
$$G_i^{ang} = 2^{1-\zeta} \sum_{j \neq i}^N \sum_{k \neq i, j}^N (1 + \lambda \cos \Theta_{ijk})^\zeta e^{-\eta(r_{ij}-\mu)^2} e^{-\eta(r_{ik}-\mu)^2} e^{-\eta(r_{jk}-\mu)^2} f_{ij} f_{ik} f_{jk}$$

G_i^{rad} and G_i^{ang} - atom-centered function

r_{ij} - distances to all atoms j within a cutoff radius r_c

η and μ - width and center of Gaussians (empirically defined)

f_c - cutoff function (smoothing impact as a function of distance)

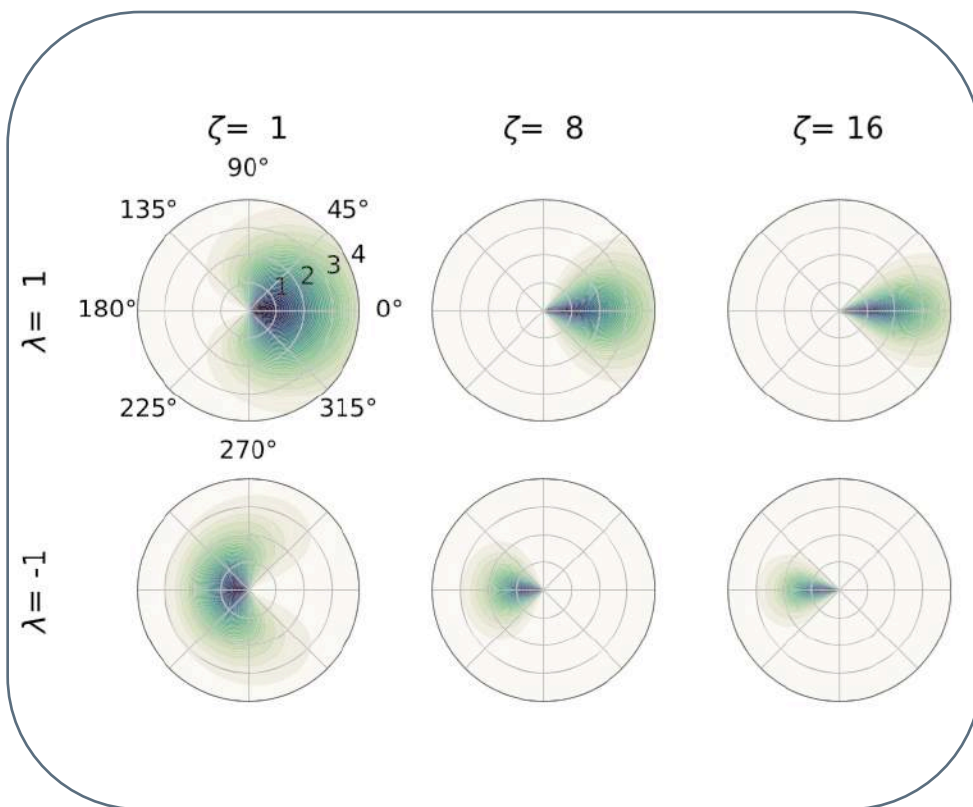
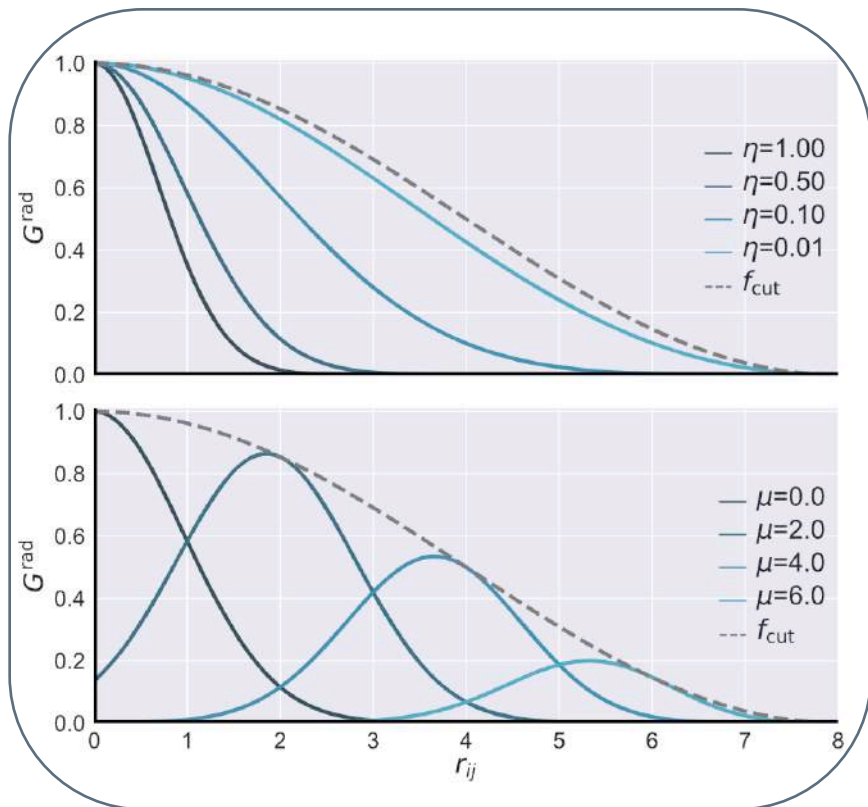


Limitation: Empirical parametrization of symmetry functions

Reveil et al **Classification of spatially resolved molecular fingerprints for machine learning applications and development of a codebase for their implementation** Mol Syst Des & Engin (2018)

Behler et al **Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces** Phys. Rev. Lett. 98, 146401

ATOM-CENTERED SYMMETRY FUNCTIONS



Reveil et al **Classification of spatially resolved molecular fingerprints for machine learning applications and development of a codebase for their implementation** Mol Syst Des & Engin (2018)

Behler et al **Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces** Phys. Rev. Lett. 98, 146401

BISPECTRUM FINGERPRINTS

$$\rho_i(r) = \delta(r) + \sum_j \delta(1 - r_{ij}) f_c(|r_{ij}|) \quad \text{where } f_c(r_{ij}) = \begin{cases} \frac{1}{2} \left[\cos\left(\frac{\pi r_{ij}}{r_c}\right) + 1 \right], & r_{ij} \leq r_c \\ 0, & r_{ij} \geq r_c \end{cases}$$

In order to achieve rotational invariance the spherical coordinates can be used.

Bispectrum – three-point correlation function that can be used to describe the atomic neighborhood

Atomic density is projected onto the surface of the four-dimensional unit sphere:

$$r \equiv \begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{aligned} \phi &= \arctan\left(\frac{y}{x}\right) \\ \theta &= \arccos(z/|r|) \\ \theta_0 &= \pi|r|/r_0 \end{aligned}$$

where ϕ, θ, θ_0 – polar angles;

$$r_0 > r_{cut}/\pi$$

4D spherical harmonics (Wigner matrices $U_{m'm}^j$) – complete basis for the interior of the 3D sphere

The projection of the atomic density on the surface of the 4D sphere can be represented through the set of the ordinary Clebsch-Gordan coefficients that are expansion coefficients of total angular momentum eigenstates:

$$c_{m'm}^j = \langle U_{m'm}^j | \rho \rangle$$

$$B_{j_1, j_2, j} = \sum_{m'_1, m_1 = -j_1}^{j_1} \sum_{m'_2, m_2 = -j_2}^{j_2} \sum_{m', m = -j}^j (c_{m'm}^j) c_{j_1, m'_1, j_2, m_2}^{jm'} c_{j_1, m'_1, j_2, m'_2}^{jm'} c_{m'_1 m_1}^{j_1} c_{m'_2 m_2}^{j_2}$$

SMOOTH OVERLAP OF ATOMIC POSITIONS (SOAP) SIMILARITY KERNELS

Локальная атомная конфигурация \tilde{x}_l может быть выражена определена посредством сглаженных атомных плотностей, центрированных на атоме l :

$$\rho_i(r) = \sum_j \omega(z_i) \varphi_\sigma(r - r_{ij}) f_c(|r_{ij}|) \quad \text{where } f_c(r_{ij}) = \begin{cases} \frac{1}{2} \left[\cos\left(\frac{\pi r_{ij}}{r_c}\right) + 1 \right], & r_{ij} \leq r_c \\ 0, & r_{ij} \geq r_c \end{cases}$$

z_i – atomic number

ω_i – associated with given atomic type weight

Атомные плотности заданы посредством гауссовых функций с дисперсией σ^2

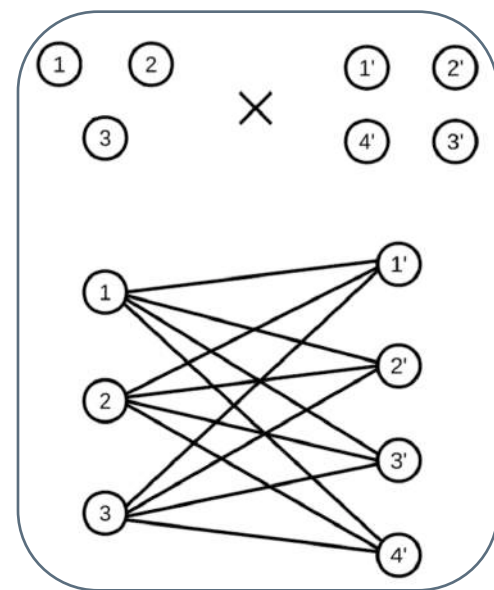
$$\varphi_\sigma(r) = \exp(-r^2/2\sigma^2)$$

Определение ядра как перекрывания двух локальных атомных плотностей, интегрированное с учетом всех возможных трехмерных вращений и инвариантного по отношению к вращению:

$$S(\tilde{x}, \tilde{x}') = \int_{\mathbb{R}^3} \rho_{\tilde{x}}(r) \rho_{\tilde{x}'}(r) dr \quad k(\tilde{x}, \tilde{x}') = \int_{SO(3)} S(R\tilde{x}, \tilde{x}')^p dR$$

Нормализованное ядро (кERNEL) с увеличением в областях наибольшего значения, принимаемого функцией при $\zeta > 1$:

$$\tilde{K}(\tilde{x}, \tilde{x}') = \left[\frac{k(\tilde{x}, \tilde{x}')}{\sqrt{k(\tilde{x}, \tilde{x})k(\tilde{x}', \tilde{x}')}} \right]^\zeta$$



Graph interpretation of the SOAP kernel. Two local environments comprised of 3 and 4 atoms are compared by forming a bipartite graph that links pairs of atoms

STEINHARDT ORDER PARAMETERS

The complex vector $q_{lm}(i)$ of particle i is defined as follows:

$$q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{lm}(r_{ij})$$

$Y_{lm}(r_{ij})$ – spherical harmonics

Steinhardt order parameters:

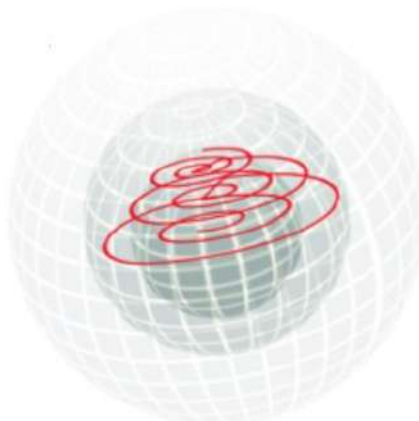
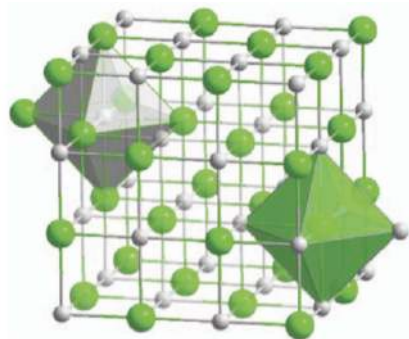
$$q_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^l |q_{lm}(i)|^2}$$

$$S_{ij} = \sum_{m=-6}^6 q_{6m}(i) q_{6m}^*(i)$$

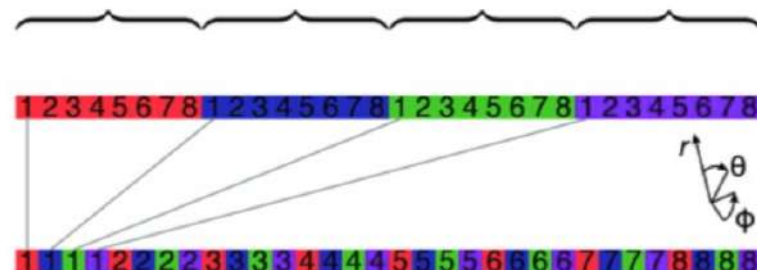
Two particles are connected if $S_{ij} > 0.5$

A particle is solidlike if $N_{connect} > 6$, otherwise - liquidlike

MORTON SPACE-FILLING CURVES



Atomic Number (q) Radial (r) Polar (θ) Azimuthal (ϕ)



Atomic layer:

Atomic number

Atomic and coordination number

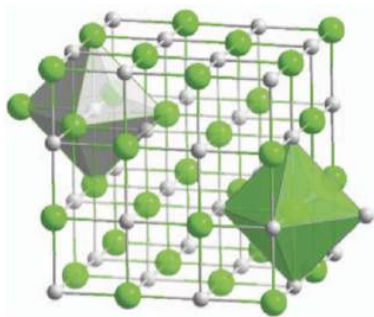
Coulomb matrix

$$c = \sum_{i=1}^{N_{\beta}} \frac{Z_{\alpha} Z_{\beta}}{r_{\alpha\beta_i}}$$

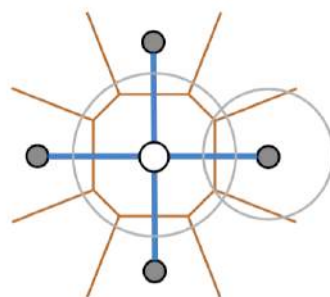
Basic steps:

- Orientation along the principal moments of inertia
- Cartesian to spherical coordinates mapping
- Four-dimensional grid formation
- Compressing to one-dimensional vector

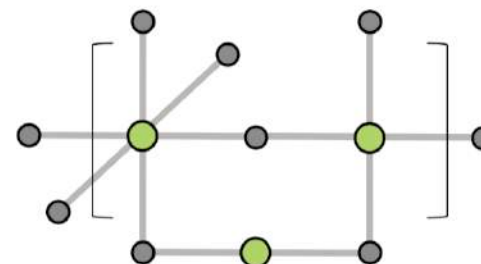
ФРАГМЕНТНЫЕ ДЕСКРИПТОРЫ ДЛЯ ПРОГНОЗИРОВАНИЯ СВОЙСТВ НЕОРГАНИЧЕСКИХ КРИСТАЛЛИЧЕСКИХ СОЕДИНЕНИЙ



Структурный тип



Полиэдр Вороного-Дирихле



Генерация графа, ассоциирование вершин графа со значением выбранных ключевых параметров, описывающих свойства образующих его химических элементов

Генерация фрагментов

Nodes (atoms)



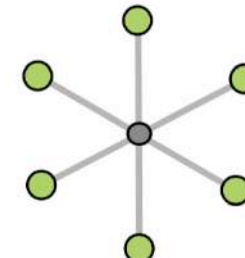
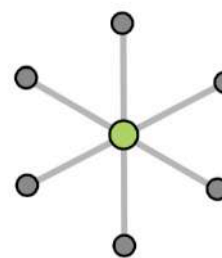
Edges (bonds)



Path fragments of length l ,
 $l = 2, 3, \dots$



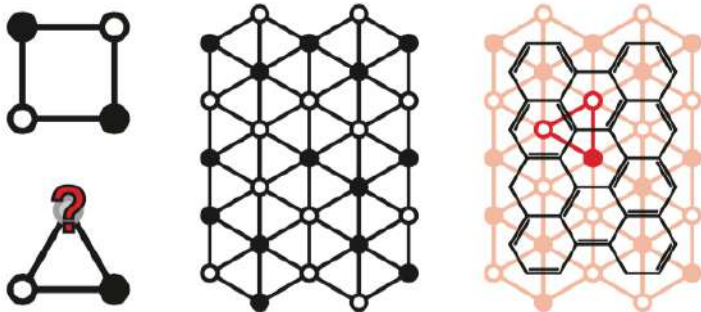
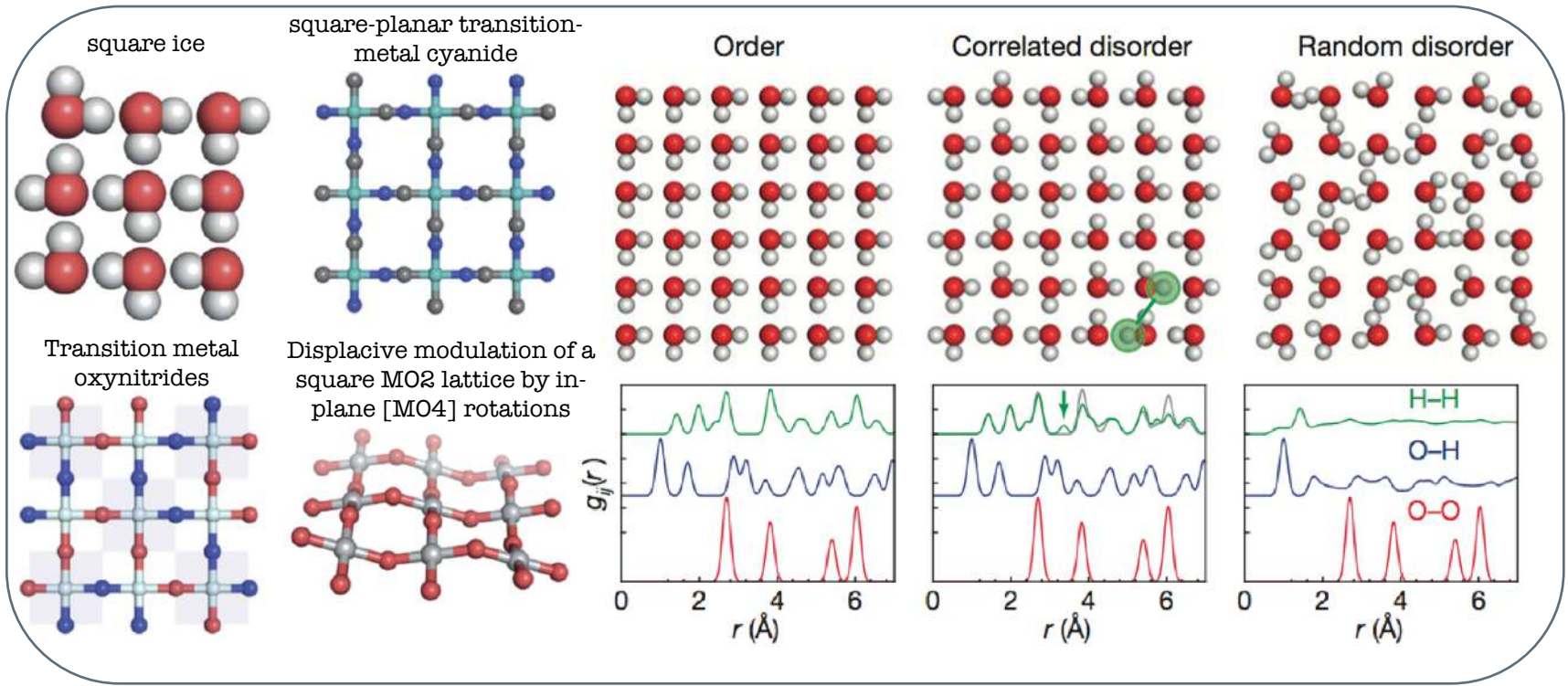
Circular fragments (polyhedrons)



$$T^E = \sum_{i=1}^{n-1} \sum_{j=i+1}^n |q_i - q_j| M_{ij}$$

$$T_{\text{bond}}^E = \sum_{\{i,j\} \in \text{bonds}} |q_i - q_j| M_{ij}$$

КОРРЕЛИРОВАННЫЙ БЕСПОРЯДОК



$$S_{\text{config}} \simeq R \ln \left(\frac{n}{2^{d/2}} \right)$$

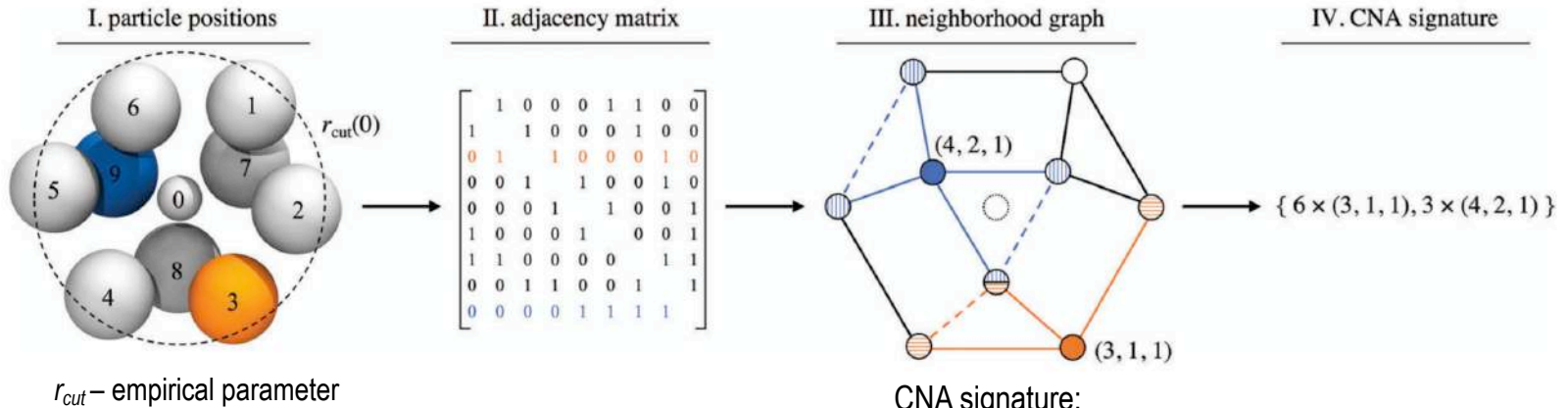
Число Полинга $P = \left(\frac{n}{2^{d/2}} \right)$

(склонность системы к беспорядку)

COMMON NEIGHBOR ANALYSIS: SIGNATURE

Controlling microscopic nature:

- Nanoscale colloidal crystals
- Micro-sized colloids

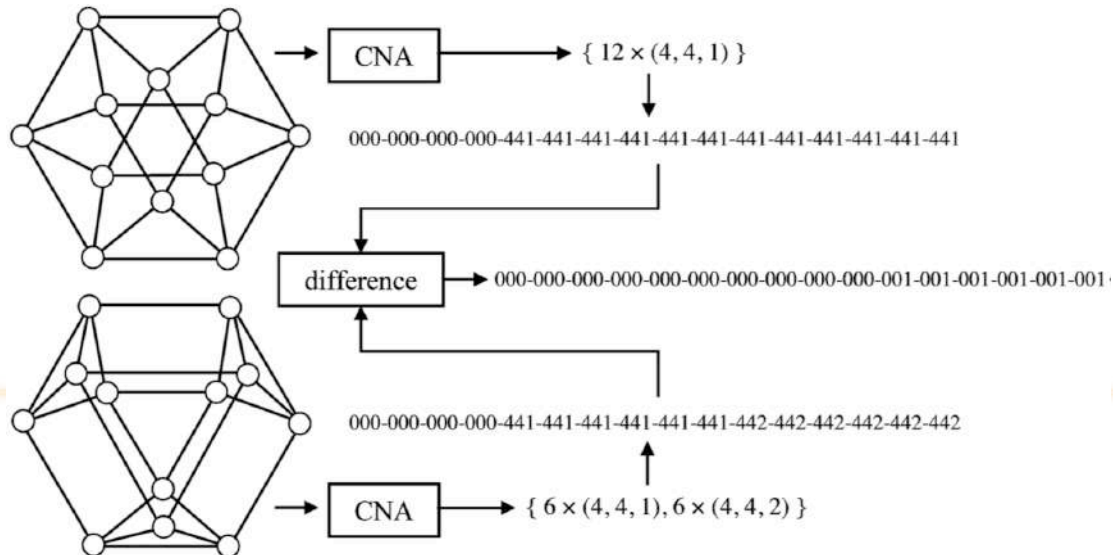


CNA signature:

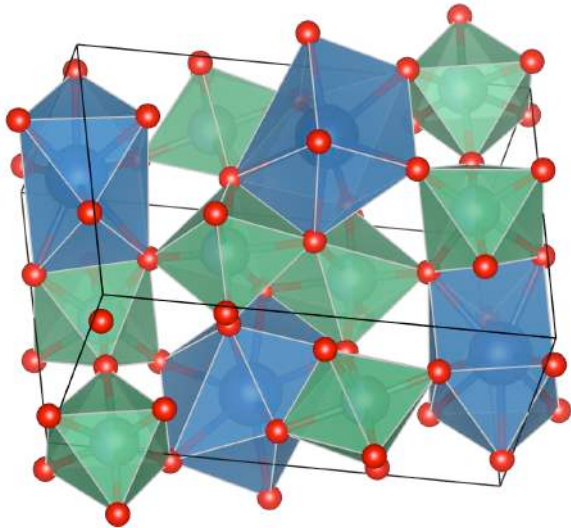
Number of adjacent nodes

Number of edges between those nodes

Length of the longest path among only those edges



ФРАГМЕНТНЫЕ ДЕСКРИПТОРЫ ДЛЯ ПРОГНОЗИРОВАНИЯ СВОЙСТВ НЕОРГАНИЧЕСКИХ КРИСТАЛЛИЧЕСКИХ СОЕДИНЕНИЙ



Aeschynite $A^3+B^4+C^5+O_6$ (SG: $Pnma$)
A=RE(57-63), B=Ti⁴⁺, C=TM

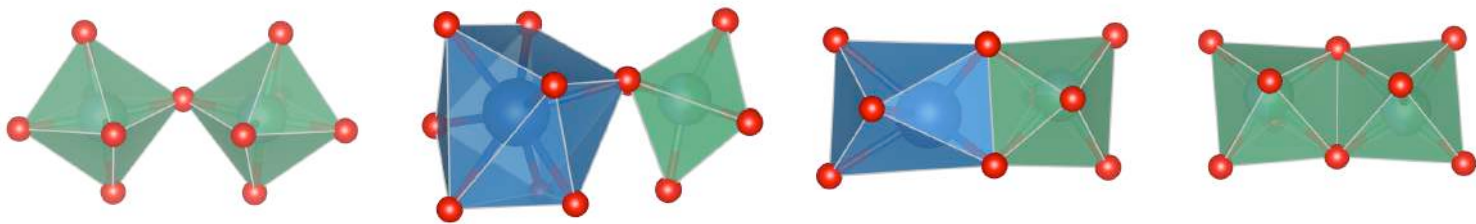
Structure parameter

$$P = m \cdot n / 2^d$$

m – relative number of polyhedra connected similarly

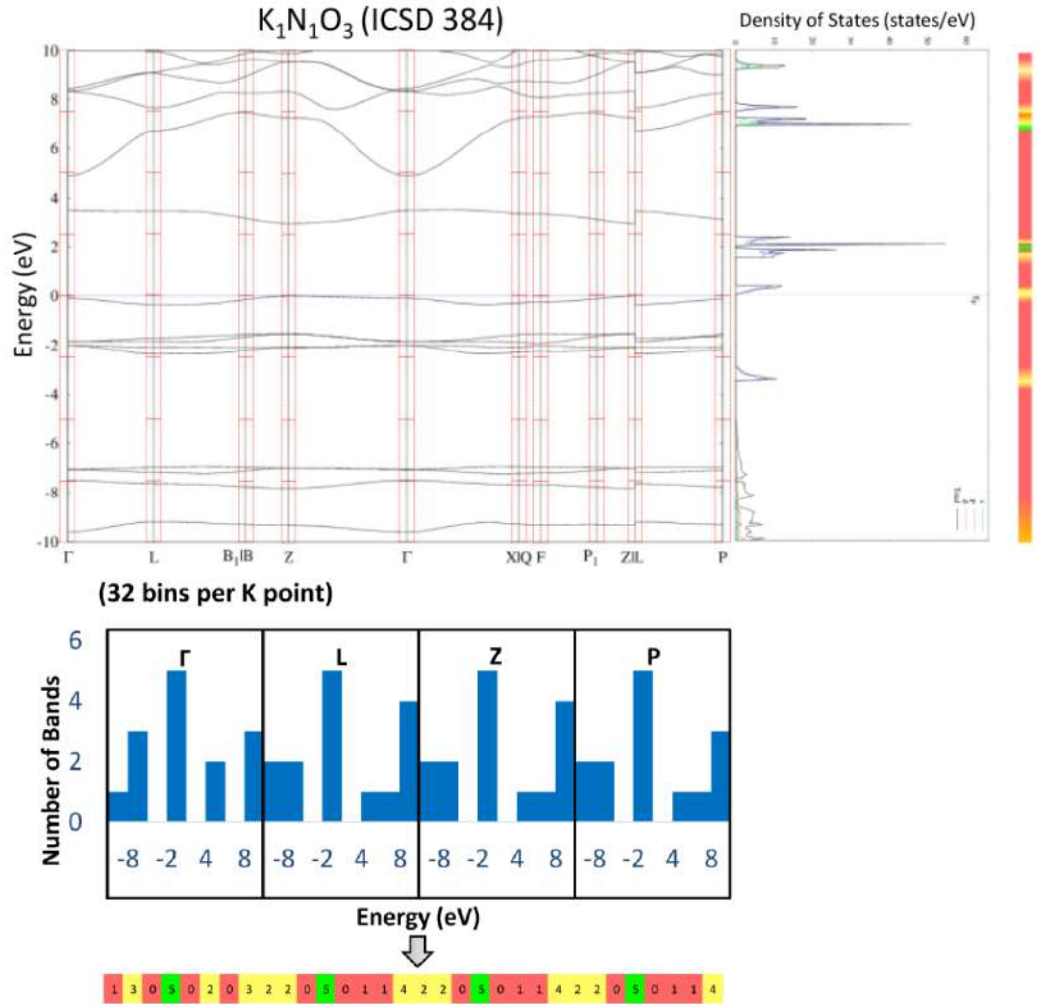
n – order of the point group symmetry

d – connectivity type (nodes, edges, grains)



$$P = \sum_i m_i \cdot n_i / 2^d$$

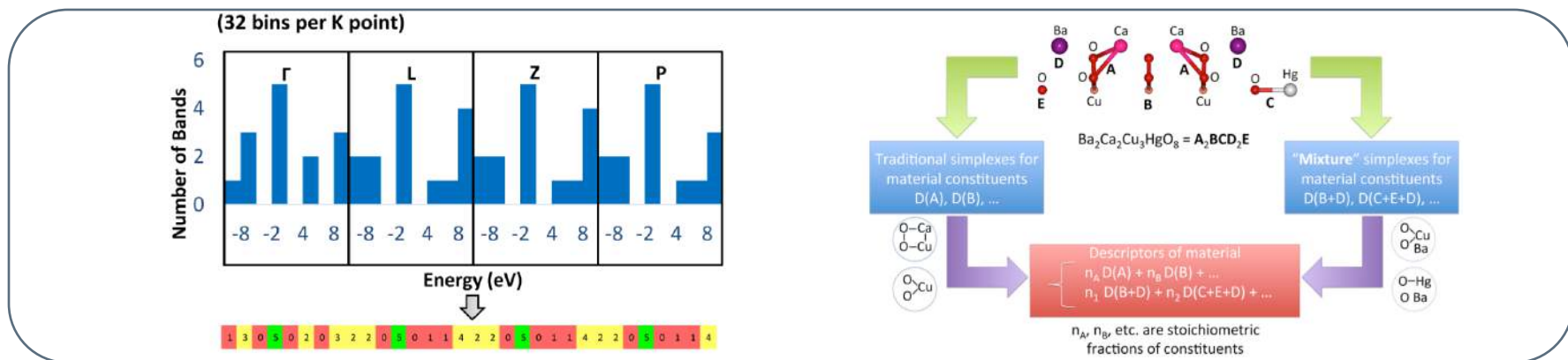
ELECTRONIC FINGERPRINTS



СКРИНИНГ СОЕДИНЕНИЙ С ЗАДАНЫМИ СВОЙСТВАМИ

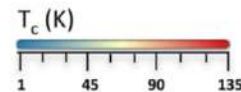
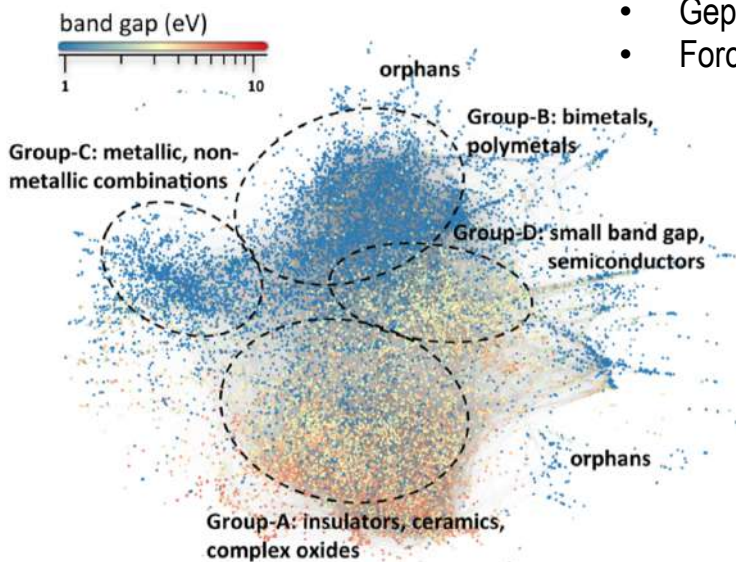


MATERIALS CARTOGRAPHY: REPRESENTING AND MINING MATERIALS SPACE FOR SCREENING OF MATERIALS WITH TARGET PROPERTIES



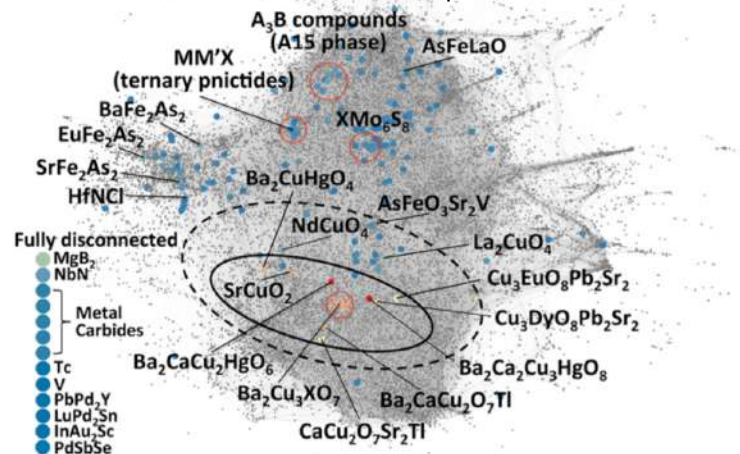
Network representation:

- Gephi package
- ForceAtlas algorithm



Data:

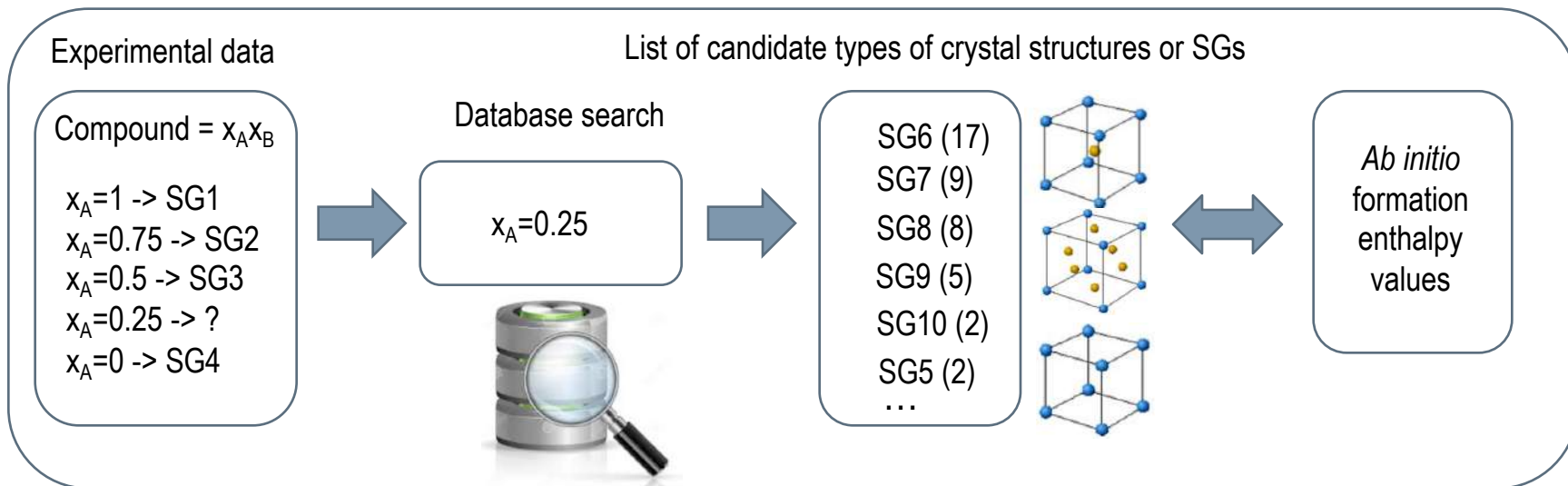
- Handbook of Superconductivity
- Handbook of Chemistry and Physics
- SuperCon Database



ПРОГНОЗИРОВАНИЕ СТРУКТУРЫ



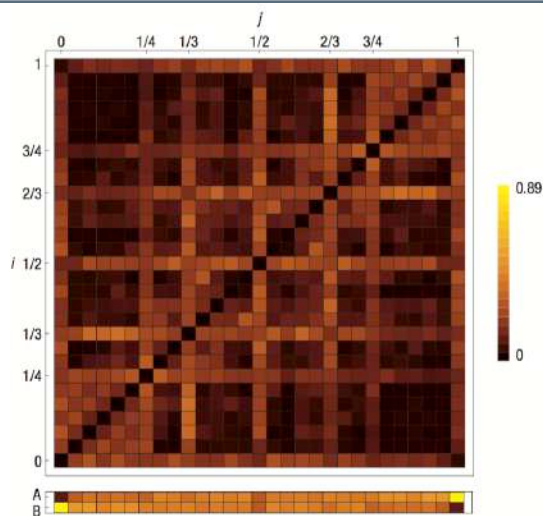
ПРОГНОЗИРОВАНИЕ СТРУКТУРЫ



$$p(\mathbf{X}) = \frac{1}{Z} \prod_i p(x_i) \prod_{j < k} g^{(2)}(x_j, x_k)$$

$$g^{(2)}(x_i, x_j) = \frac{p(x_i, x_j)}{p(x_i)p(x_j)}$$

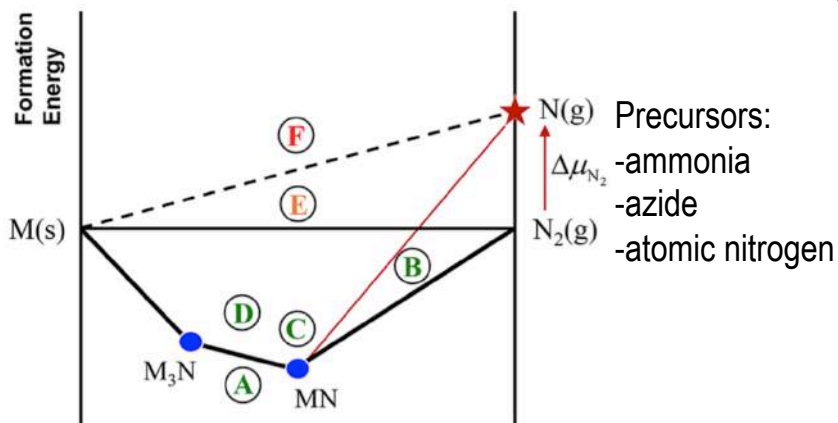
Cumulative function g – correction term for using the information from the compounds with correlated behavior



ПОИСК НОВЫХ МЕТАСТАБИЛЬНЫХ АЗОТ-НАСЫЩЕННЫХ НИТРИДОВ

	nitrides	oxides	all inorganic solids
no. of unique ICSD phases	1253	13 497	29 902
percent metastable ($T = 0$ K)	59.8%	56.0%	51.6%
$\Delta H(E - E_{GS})$ – median	67 meV/atom	15.4 meV/atom	14.9 meV/atom
$\Delta H(E - E_{GS})$ – 90th percentile	190 meV/atom	62 meV/atom	70 meV/atom
median cohesive energy	-6.38 eV/atom	-6.26 eV/atom	-4.88 eV/atom
electronegativity	3.02	3.44	

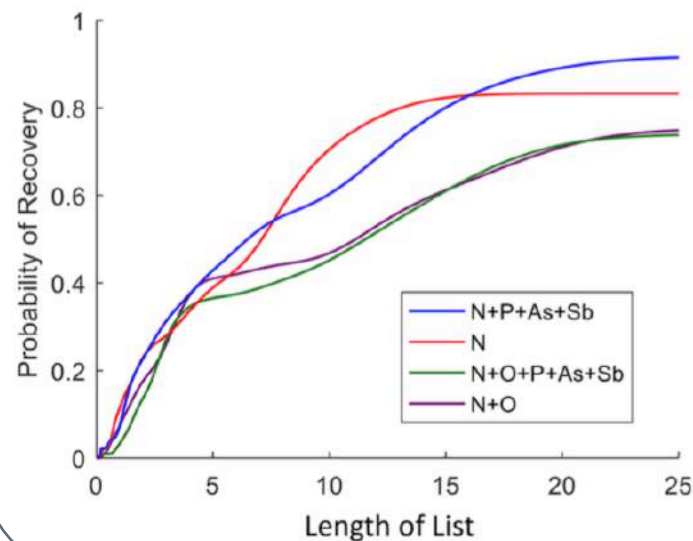
Stabilization of nitrogen-rich metastable nitrides



Negative Formation Energy w.r.t Ambient Conditions

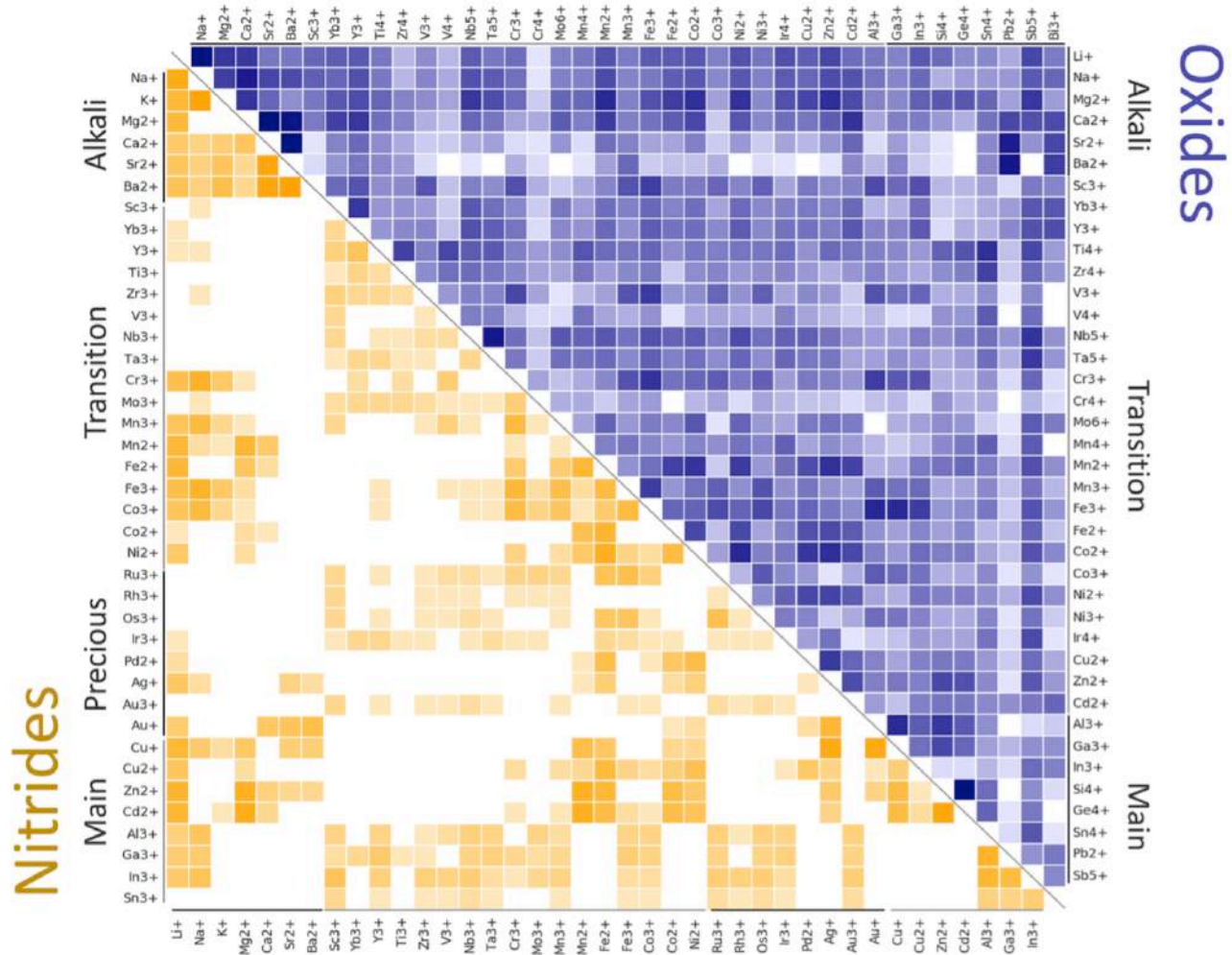
- A.** Stable Compound
- B.** Stabilizable by Reactive Nitrogen Precursors
- C.** Metastable Polymorph
- D.** Metastable Phase-Separating
- E.** Negative Formation Energy w.r.t Sputtering
- F.** No Thermodynamic Driving Force

Probability of finding nitride with known structure from suggested candidates (TP rate)



ПОИСК НОВЫХ МЕТАСТАБИЛЬНЫХ АЗОТ-НАСЫЩЕННЫХ НИТРИДОВ

Substitution matrix for nitrides and oxides predicted by DMSP approach



ПОИСК НОВЫХ МЕТАСТАБИЛЬНЫХ АЗОТ-НАСЫЩЕННЫХ НИТРИДОВ

nitride compd	prototype ICSD ID	prototype structure	space group	E abv hull (eV/atom)	decomp products	peritride	critical μ_{N_2} (eV/atom)
RuN ₂ * ^b	167872	RhN ₂	$F\bar{4}3m$	0.10	Ru + N ₂	N	0.20
Nb ₃ N ₅	76460	Ta ₃ N ₅	$Cmcm$	0.03	Nb ₅ N ₆ + N ₂	N	0.20
Sn ₃ N ₄ * ^c	89525	Sn ₃ N ₄	$Fd\bar{3}m$	0.16	Sn + N ₂	N	0.29
V ₃ N ₄	100135	Si ₃ N ₄	$P2_1/m$	0.04	VN + N ₂	N	0.30
ReN ₂ * ^d	187441	ReN ₂	$C2/m$	0.23	Re ₃ N + N ₂	Y	0.41
RuN ₂ * ^b	240754	RuN ₂	$Pmnn$	0.30	Ru + N ₂	N	0.45
PtN ₂ * ^e	166462	PtN ₂	$Pa\bar{3}$	0.30	Pt + N ₂	Y	0.45
ReN	162871	BN	Cm	0.16	Re ₃ N + N ₂	N	0.48
OsN ₂ * ^f	260545	OsN ₂	$P6/mmm$	0.35	Os + N ₂	N	0.52
FeN ₂	240759	OsN ₂	$Pnmm$	0.18	FeN + N ₂	Y	0.54
IrN ₂ * ^f	240755	IrN ₂	$P2_1/c$	0.38	Ir + N ₂	Y	0.57
Ti ₃ N ₄	78944	Zr ₃ N ₄	$Pnma$	0.09	TiN + N ₂	N	0.60
Re ₃ N ₄	156339	Ge ₃ N ₄	$Pnma$	0.27	Re ₃ N + N ₂	N	0.64
PdN ₂ * ^g	191244	PdN ₂	$Pnmm$	0.44	Pd + N ₂	Y	0.65
Re ₃ N ₅	95782	P ₃ N ₅	$Imm2$	0.34	Re ₃ N + N ₂	N	0.68
Re ₂ N	181874	Re ₂ N	$P6_3/mmc$	0.08	Re ₃ N + N ₂	N	0.72
SbN	162883	BN	$P2_1/c$	0.40	Sb + N ₂	N	0.80
Cu ₂ N ₂	60168	MoN	$P\bar{1}$	0.40	Cu + N ₂	Y	0.81
CrN ₂	240754	RuN ₂	$Pnmm$	0.27	CrN + N ₂	Y	0.82
Na ₃ N* ^h	421115	Na ₃ N	$Pm\bar{3}m$	0.21	Na + N ₂	N	0.84
Cr ₃ N ₄	156339	Ge ₃ N ₄	$Pnma$	0.13	CrN + N ₂	N	0.92
Mn ₃ N ₄	100135	Si ₃ N ₄	$P2_1/m$	0.13	MnN + N ₂	N	0.92
Mo ₃ N ₅	95782	P ₃ N ₅	$Imm2$	0.22	MoN + N ₂	N	0.97
Ti ₃ N ₄	92156	Si ₃ N ₄	$P31c$	0.14	TiN + N ₂	N	1.00
Cu ₃ N* ⁱ	53313	Cu ₃ N	$Pm\bar{3}m$	0.26	Cu + N ₂	N	1.04
Pb ₃ N ₂	91273	NiSr ₂ N ₂	$Pnma$	0.42	Pb + N ₂	N	1.05
TiN ₂ * ^j	N/A	TiN ₂	$I4/mcm$	0.39	TiN + N ₂	N	1.18
MoN ₂ * ^k	260549	OsN ₂	$P4/mbm$	0.38	MoN + N ₂	Y	1.23
Pb ₃ N ₄	41952	C ₃ N ₄	$R3m$	0.75	Pb + N ₂	Y	1.32
Pb ₃ N ₂	182699	Nb ₂ N ₃	$Pnma$	0.54	Pb + N ₂	N	1.35
Mo ₂ N ₃	16528	MoNCl ₃	$P\bar{1}$	0.24	MoN + N ₂	N	1.37
BiN	162876	BN	$Pnma$	0.72	Bi + N ₂	N	1.44
Pd ₃ N ₂	162795	Ca ₃ N ₂	$C2/m$	0.58	Pd + N ₂	N	1.46

Perspectives:

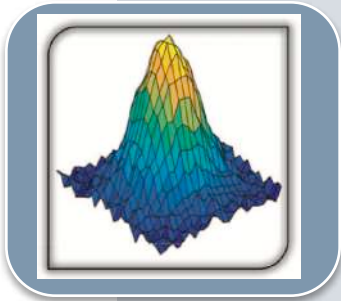
- Down-conversion nitride materials
- Localized surface plasmon resonance materials

АНАЛИЗ ЛОКАЛЬНОЙ СТРУКТУРЫ МАТЕРИАЛА

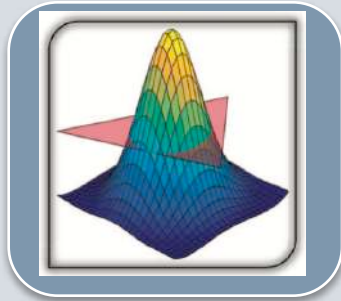


Анализ локальной структуры материала

Определение локальной структуры материала при недостаточной информативности структурного фактора (несколько кристаллографических фаз, присутствие сложных дефектов): многомерный статистический анализ координационных сфер отдельных атомов с использованием метрики между атомом и его окружением для определения структуры и симметрии



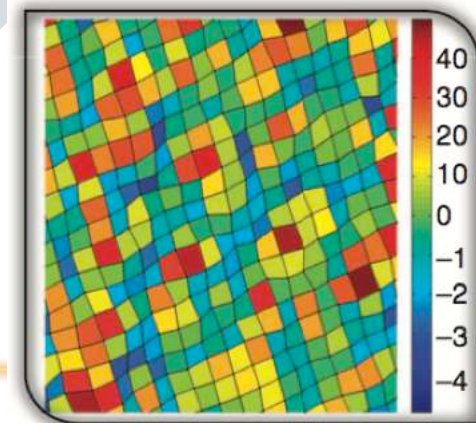
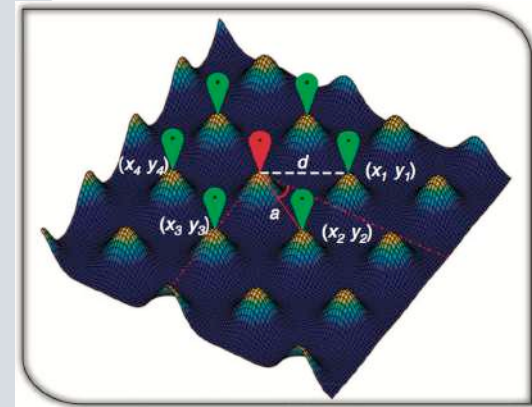
Аппроксимация формы атома при помощи идентификации репрезентативной части изображения, построение корреляционной карты



Удаление шума

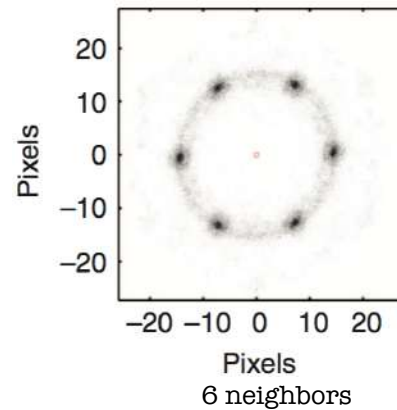
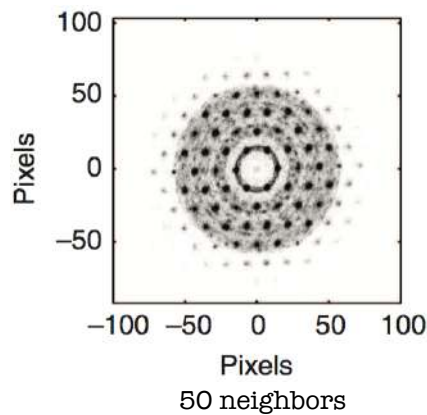
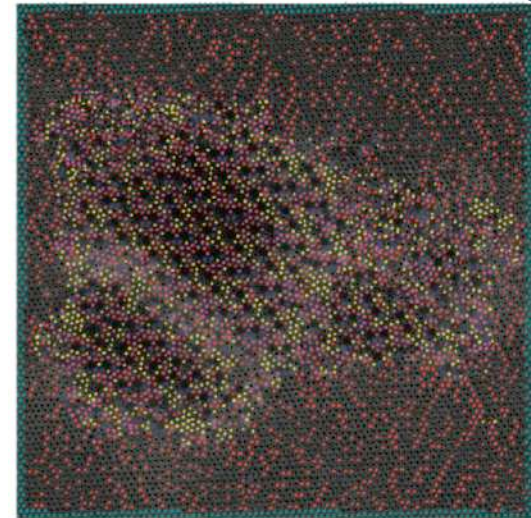
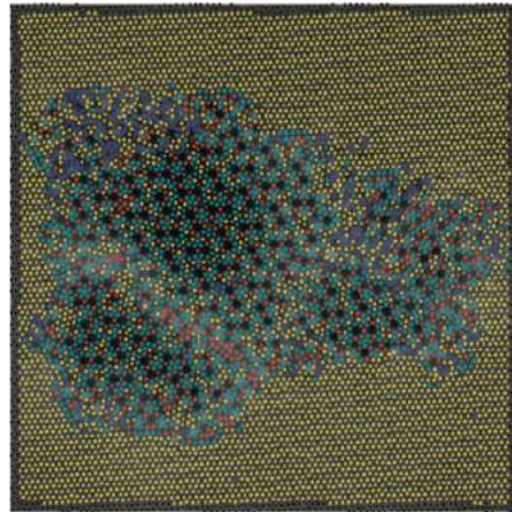
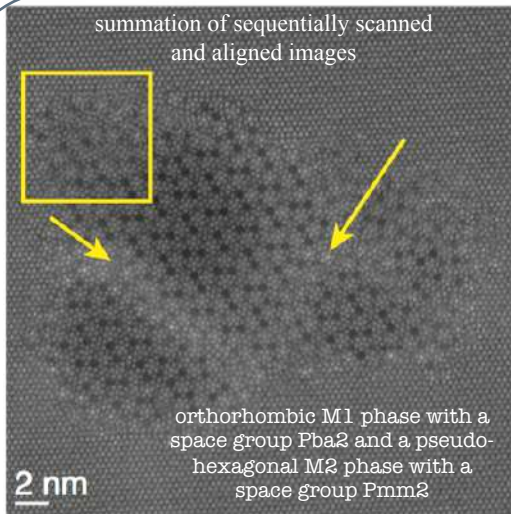


Идентификация центроидов



Анализ локальной структуры материала

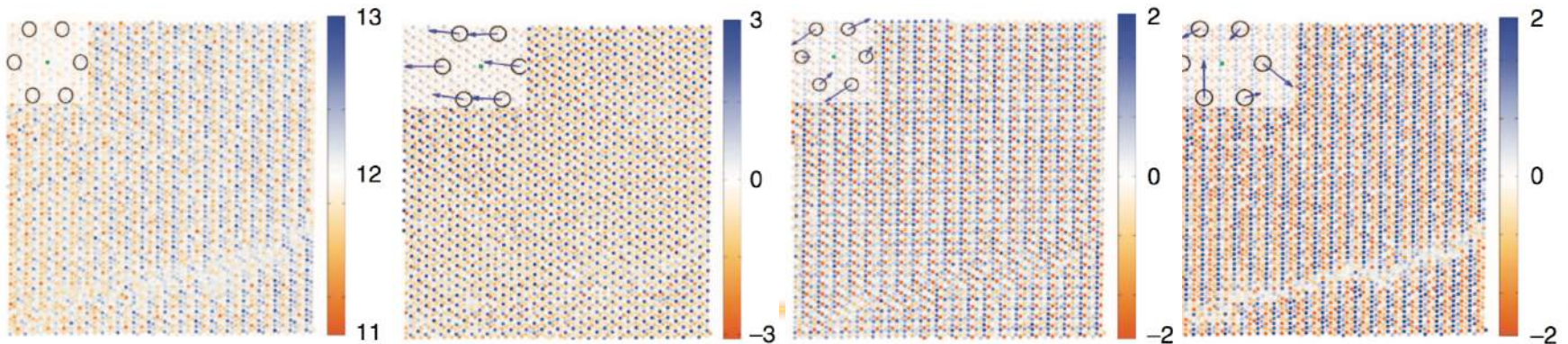
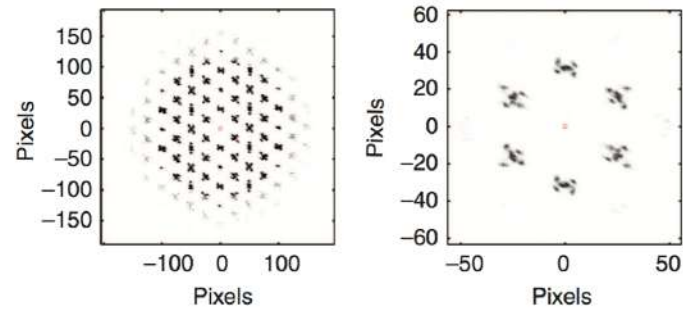
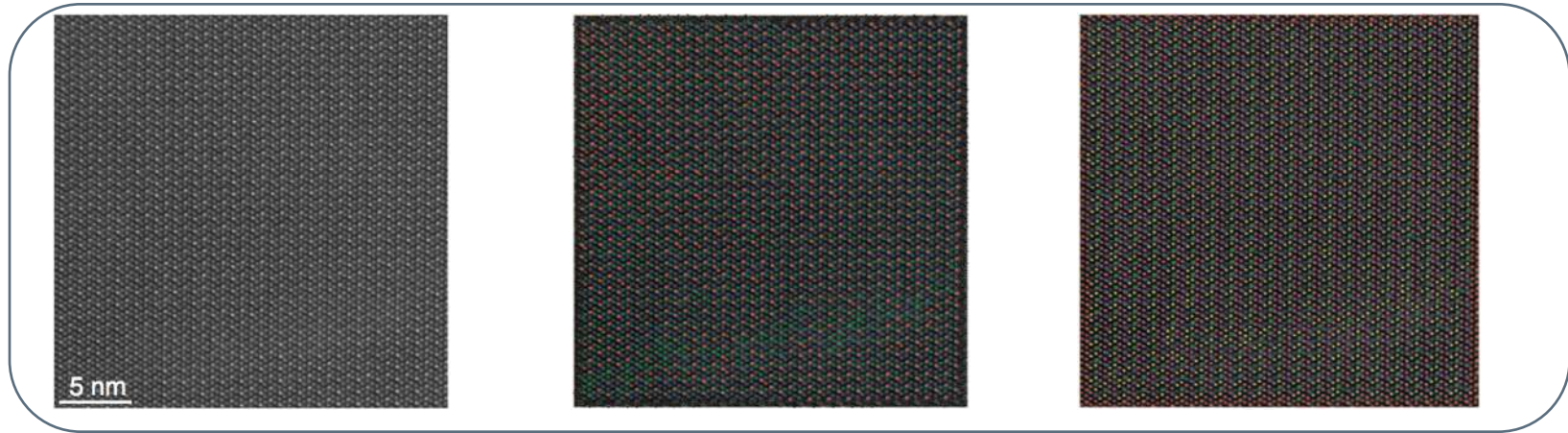
Two-phase Mo-V-Te-Ta oxide



Near-coordination sphere for each atom: $\mathbf{N}_i = ([x_1, y_1], \dots, [x_j, y_j])$

K-means clustering:
$$\arg \min_S \sum_{i=1}^K \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \mu_i\|^2$$

Анализ локальной структуры материала



Principal Component Analysis

Belianinov et al Identification of phases, symmetries and defects through local crystallography Nature Communications (2015), 6, 7801

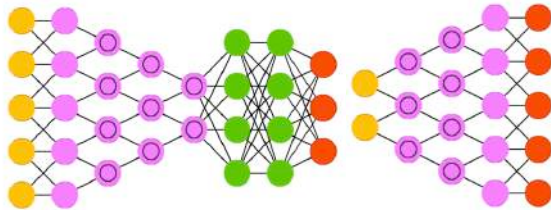
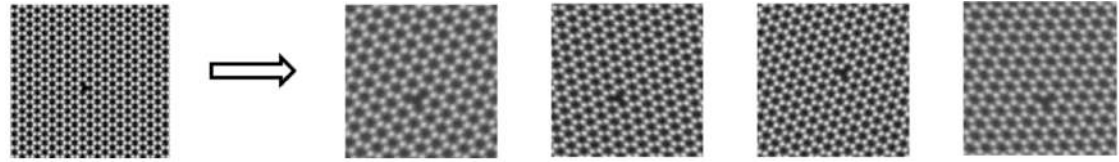
АНАЛИЗ ЛОКАЛЬНОЙ СТРУКТУРЫ МАТЕРИАЛА: ИСПОЛЬЗОВАНИЕ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

Возможности:

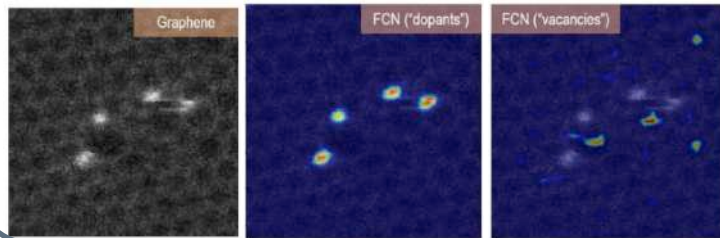
- Анализ кристаллических решеток и структуры и типа дефектов
- Уточнение типа структуры
- Извлечение значимых структурных/химических параметров

Обучающий набор данных (искусственные данные, мимикрирующие данные STEM):

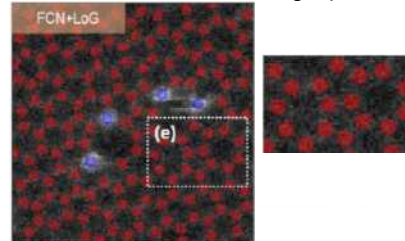
- Идеальная кристаллическая решетка
- Решетка с присутствием вакансии
- Атом-допант с большим атомным номером



- Finding atomic positions from raw experimental data
- Pixel-wise recognition of atomic defects and lattice reconstructions
- Possible chemical structure-related interpretation of output

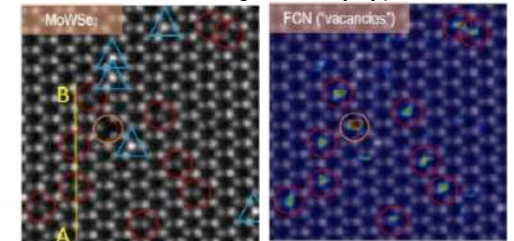


Stone-Wales defects in graphene



Identifying "not-seen" defects:

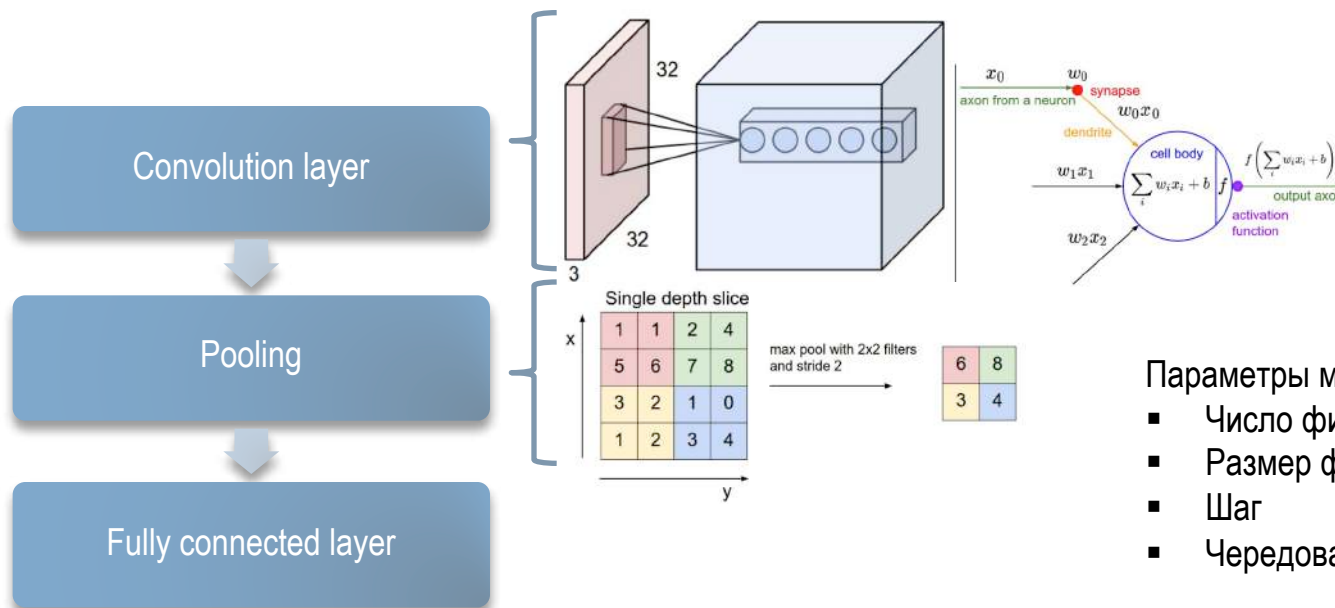
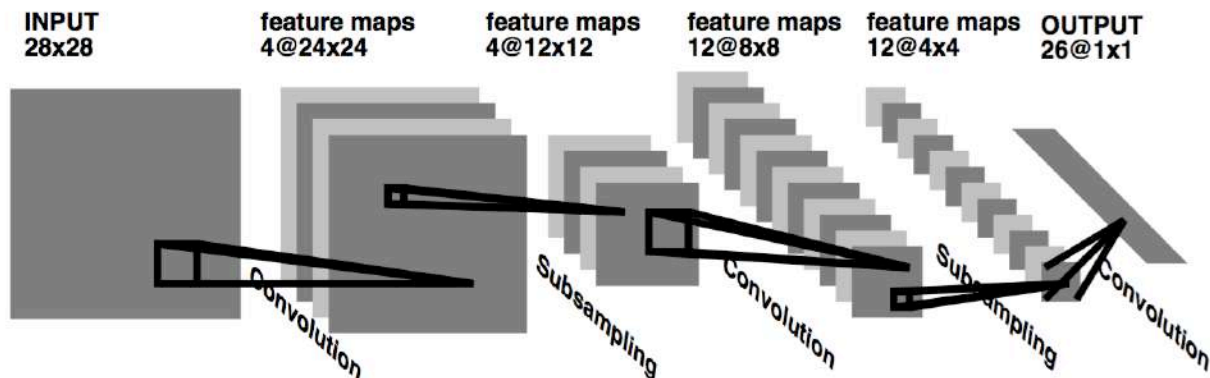
Discriminating vacancy type



CONVOLUTIONAL NEURAL NETWORKS

Specific characteristics:

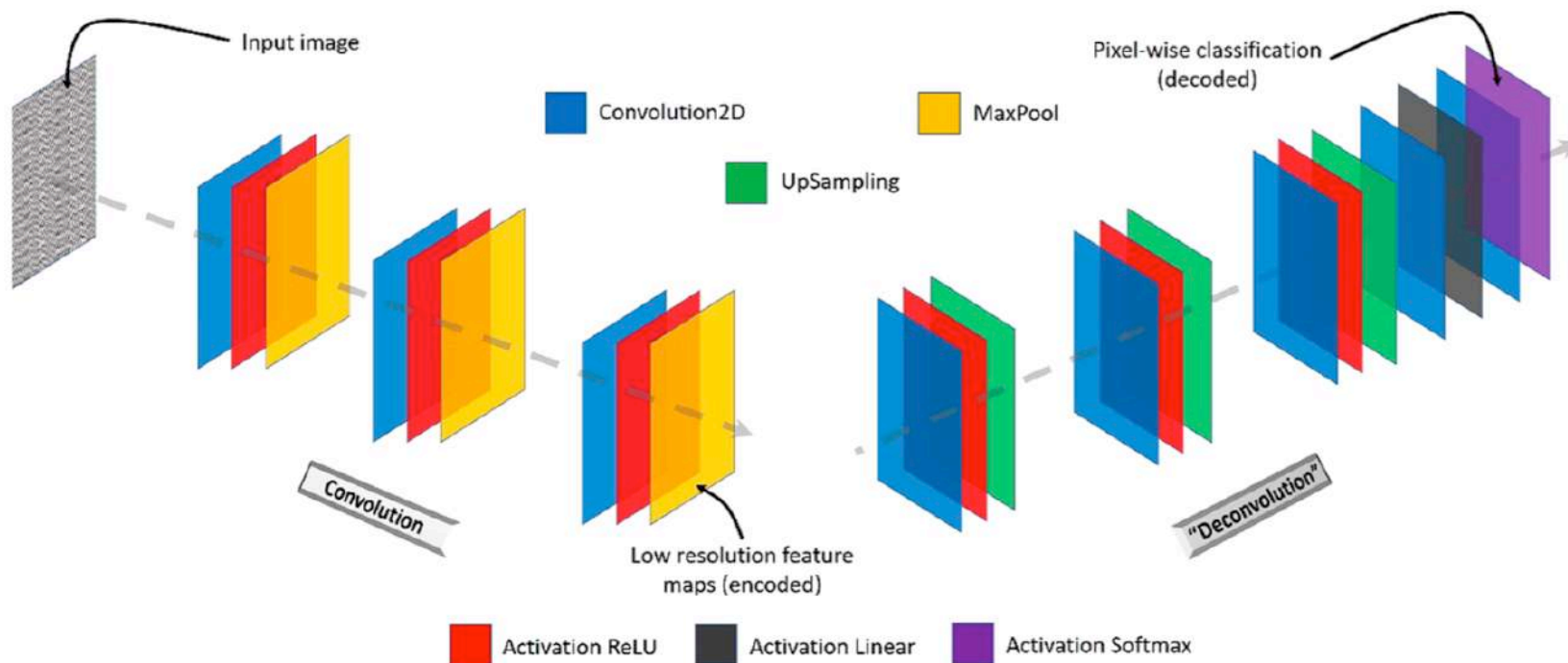
- local connections
- shared weights
- pooling
- using many layers



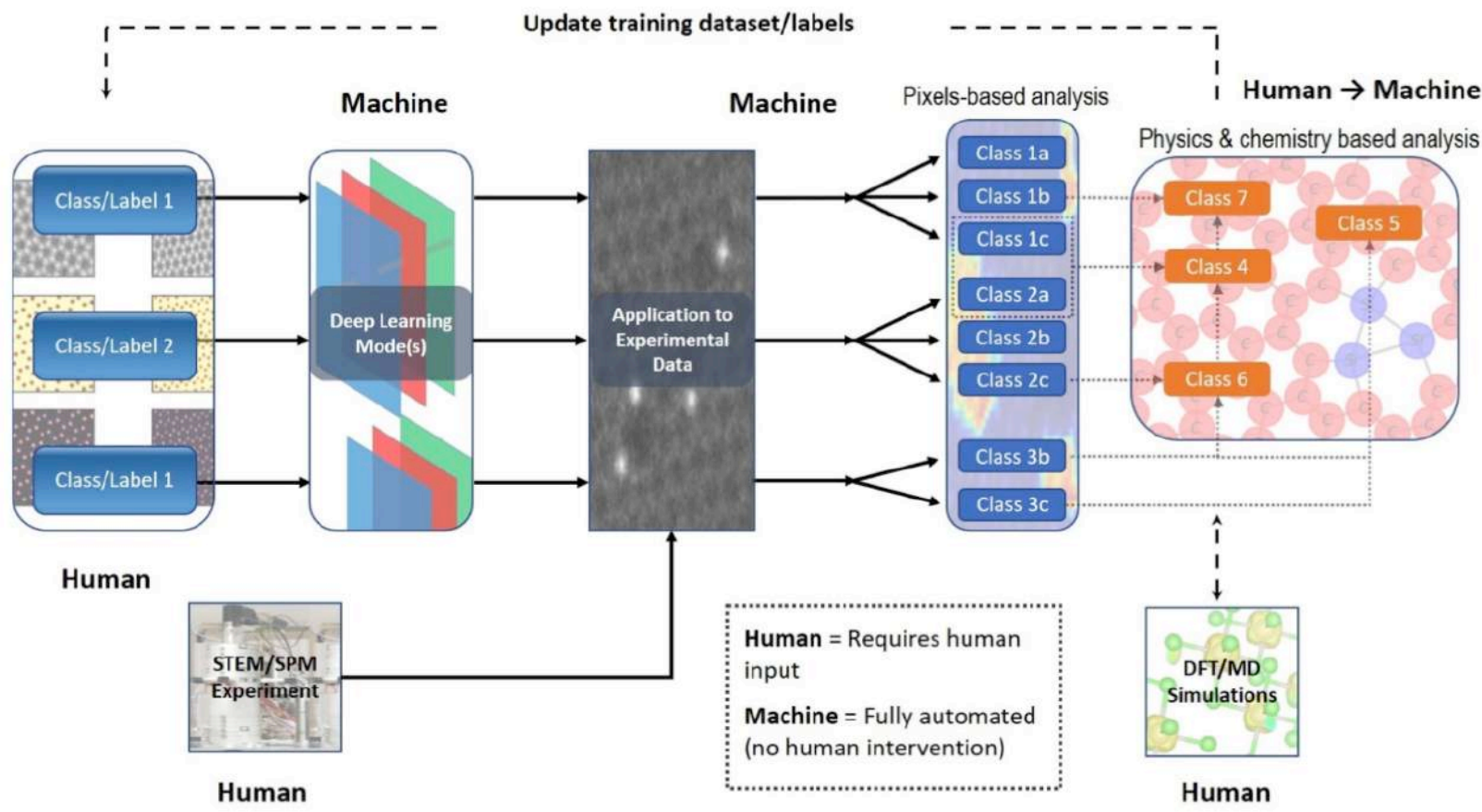
Параметры метода:

- Число фильтров
- Размер фильтров
- Шаг
- Чередование слоев

АНАЛИЗ ЛОКАЛЬНОЙ СТРУКТУРЫ МАТЕРИАЛА: ИСПОЛЬЗОВАНИЕ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

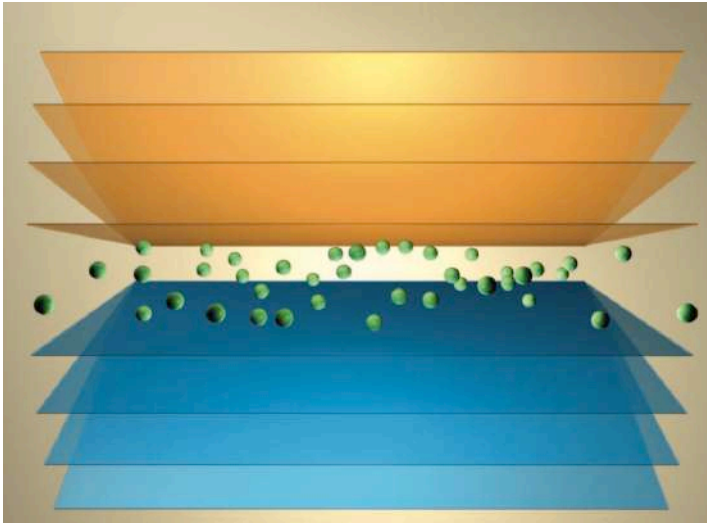


АНАЛИЗ ЛОКАЛЬНОЙ СТРУКТУРЫ МАТЕРИАЛА: ИСПОЛЬЗОВАНИЕ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ



АНАЛИЗ ПРОЦЕССОВ, ПРОИСХОДЯЩИХ НА ИНТЕРФЕЙСАХ:
МЕЖЗЁРЕННЫЕ ГРАНИЦЫ

PHASE BEHAVIOR OF INTERFACES: COMPLEXIONS



Grain boundary (GB) can be viewed as interface-stabilized “complexion” phase that can be formed between different crystals within a material and that are chemically and structurally distinct from any bulk phases and can be chemically and structurally inhomogeneous

Nanoscale interface structures are equilibrium phases that obey thermodynamic rules analogous to those of bulk phases

A number of unexplained observations concerned:

- abnormal grain growth
- deliberate incorporation of certain atomic elements into grain boundaries can drastically change the materials properties

THE IMPORTANCE OF GRAIN BOUNDARY COMPLEXIONS IN AFFECTING PHYSICAL PROPERTIES

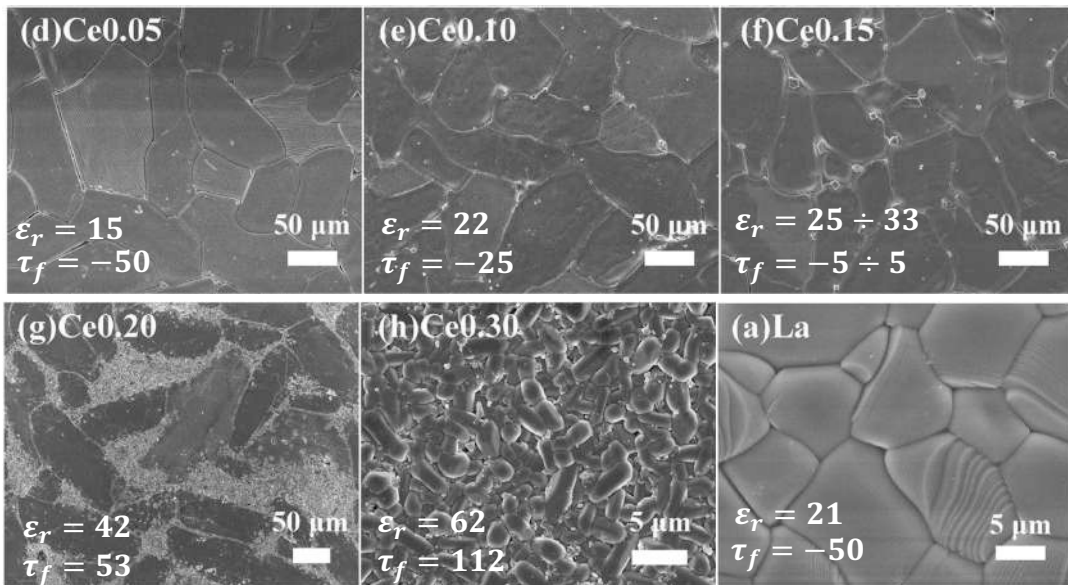
Some effects of complexions on diffusional transport and physical properties

- Non-Arrhenius enhancement of diffusional transport limited processes such as sintering and grain growth, unascertained impact on ion diffusion processes
- Activated sintering process
- Coarse grain structure obtained after prolonged annealing at high t° on average exhibit different complexions -> complex unexpected materials behavior
- The presence of localized phonon modes as a result of layering transitions may lead to modified localized interfacial vibrational modes

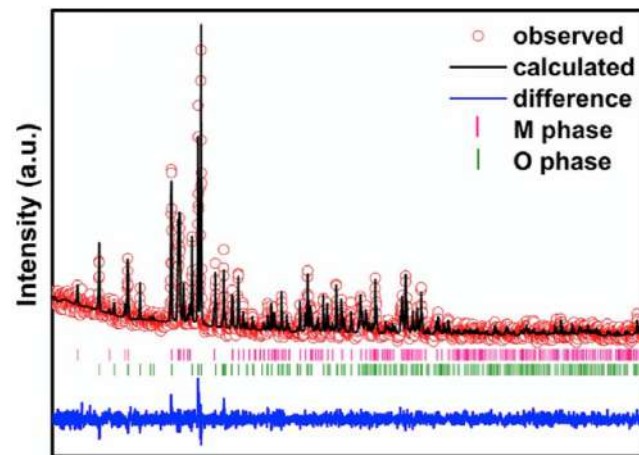
THE PHASE BEHAVIOR OF INTERFACES: MICROWAVE DIELECTRIC PROPERTIES

Aeschnites/Euxenites

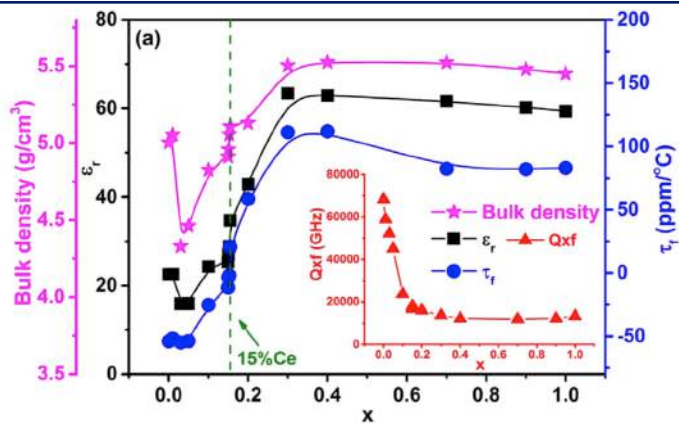
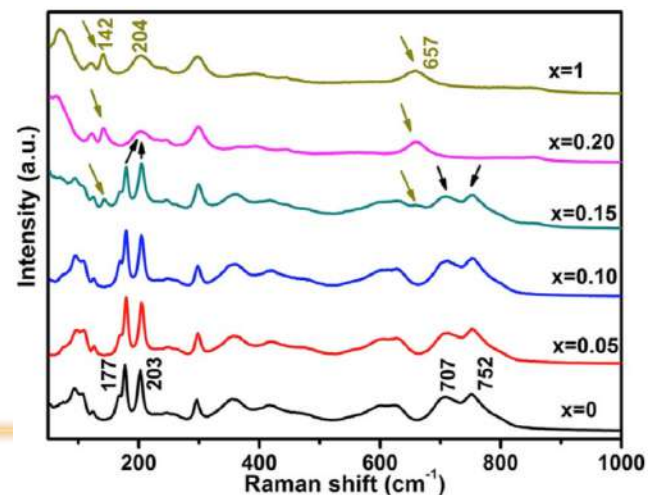
$\text{La}_{1-x}\text{Ce}_x\text{TiNbO}_6$



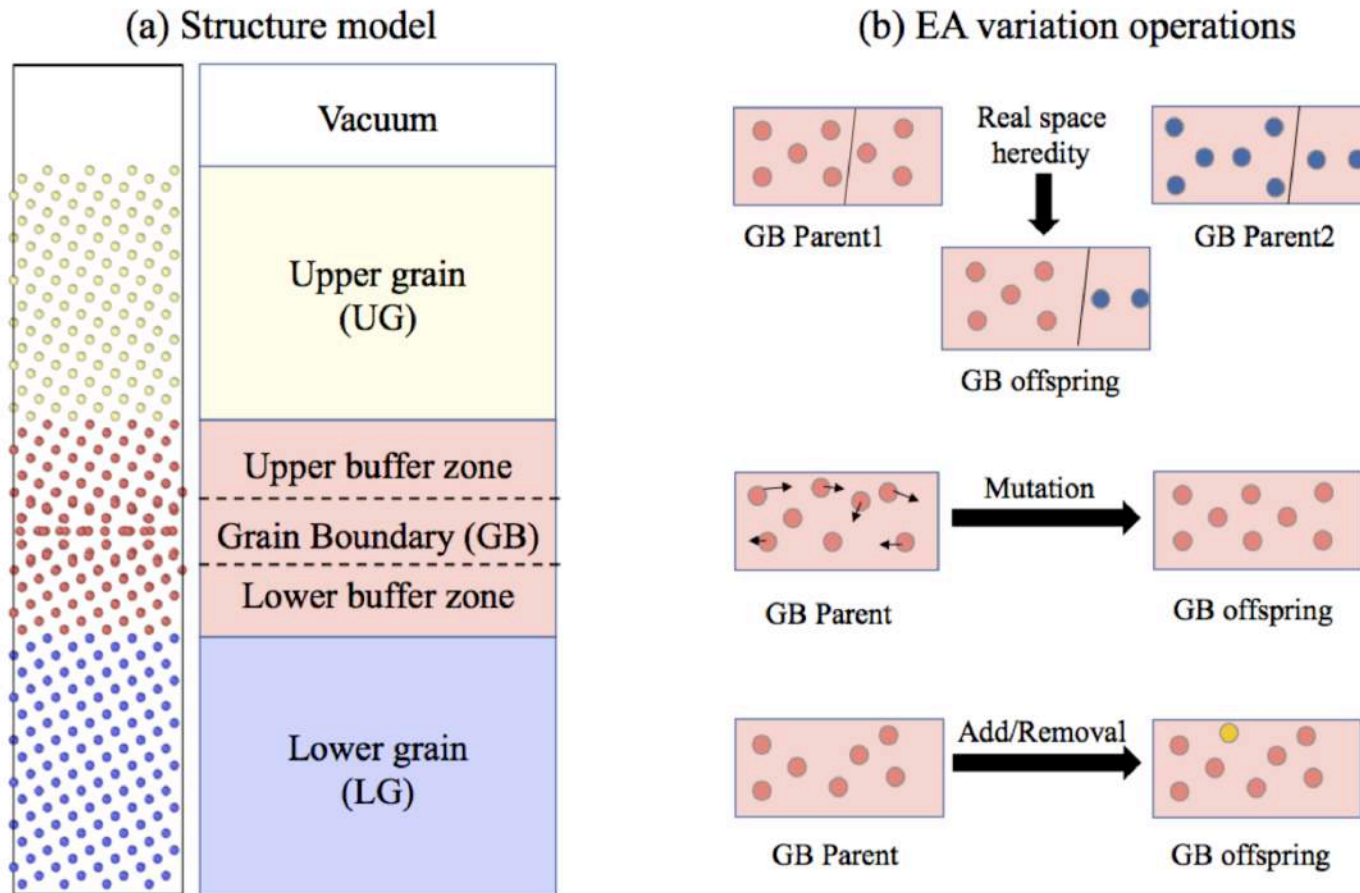
Rietveld refinement plots of $\text{La}_{0.85}\text{Ce}_{0.15}\text{TiNbO}_6$



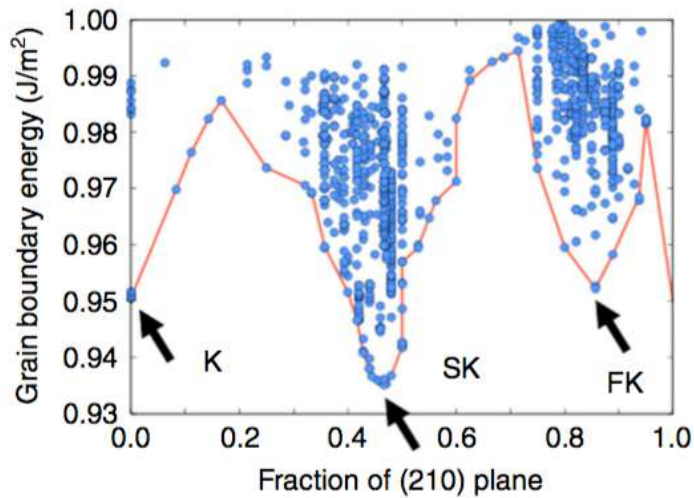
Raman spectra of $\text{La}_{1-x}\text{Ce}_x\text{TiNbO}_6$



ANALYSIS OF INTERFACE PROCESSES IN CERAMICS: GRAIN BOUNDARIES



ANALYSIS OF INTERFACE PROCESSES IN CERAMICS: GRAIN BOUNDARIES (GB)



Identifying the number and type of grain boundary phases by virtue of clustering algorithms

Descriptors

thermodynamics-based

structural features

excess volume per unit area

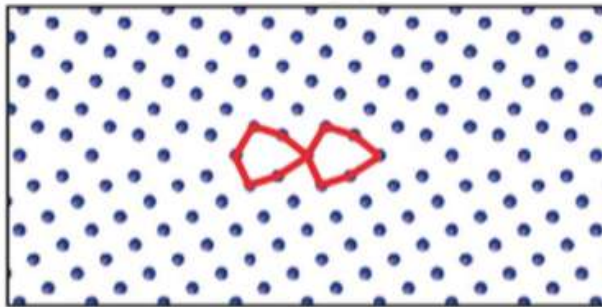
GB stress tensor

number of atoms

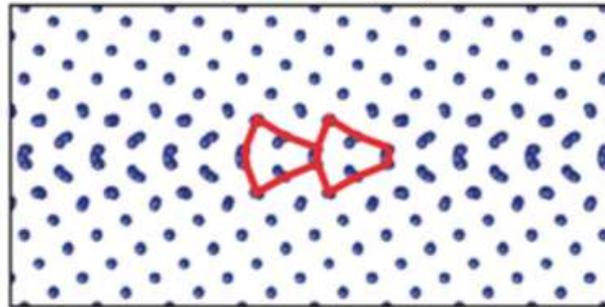
Steinhardt order parameters

GBs distinguished

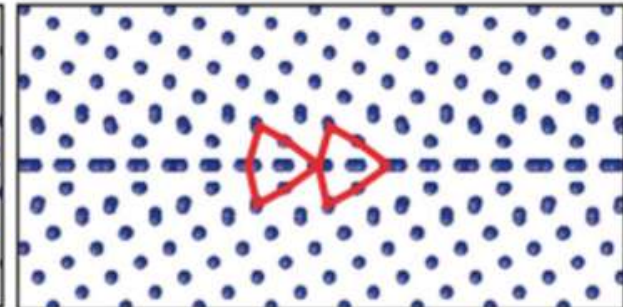
Kites (K)



Split Kites (SK)



Filled Kites (FK)



MATERIALS SYNTHESIS DETAILS AS PARAMETERS UNRAVELING
COMPOSITION-STRUCTURE-PROPERTY RELATIONSHIPS

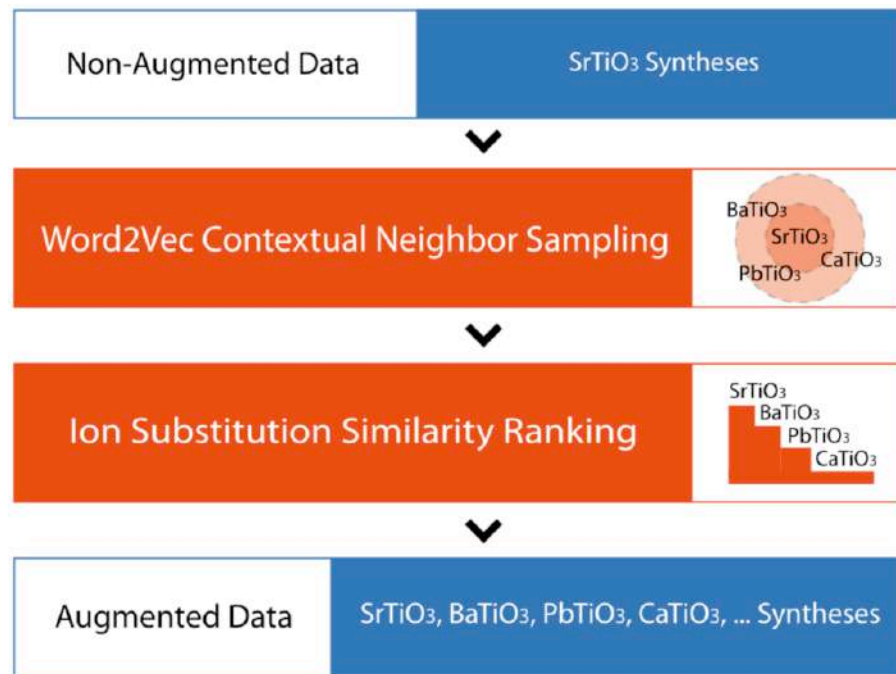
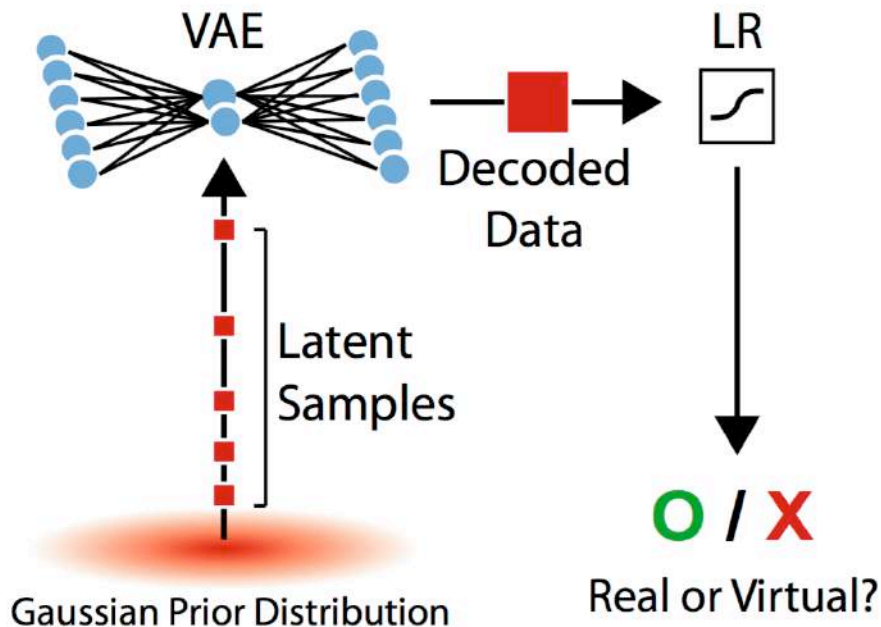
VIRTUAL SCREENING OF INORGANIC MATERIALS SYNTHESIS PARAMETERS WITH DEEP LEARNING

Descriptors:

processing (synthesis) information – sintering and calcination temperature and time, method of synthesis, solvent

Tasks:

- SrTiO₃/BaTiO₃ synthesis details discriminating
- MnO₂ polymorph elucidation



VIRTUAL SCREENING OF INORGANIC MATERIALS SYNTHESIS PARAMETERS WITH DEEP LEARNING

Loss function:

$$\mathcal{L}(x, x') + \sum_j KL(q_j(z|x) || p(z))$$

