Generative Models for Materials Discovery

AI-driven crystal structure generation

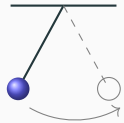
Astrid Klipfel

IA^2 2026 – Institut d'Automne en Intelligence Artificielle
Le Touquet-Paris-Plage, 22 October 2026

CRIL, UMR 8188 – Université d'Artois & CNRS

Conservation laws in physics

Energy



$E = K + U = \text{const}$

kinetic $\leftrightarrow$ potential

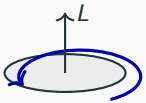
Momentum



$p = \sum_i m_i v_i = \text{const}$

collisions, recoil

Angular momentum



$L = \sum_i r_i \times p_i = \text{const}$

spinning skater

Why are these quantities conserved?

Emmy Noether (1882–1935)

1882 Born in Erlangen, Germany

1907 PhD in Erlangen, on invariant theory

1915 Invited to Göttingen by Hilbert and Klein (general relativity). As a woman, not allowed to teach: lectures under Hilbert's name

1918 *Invariante Variationsprobleme*: her two theorems

1919 Finally allowed to teach (habilitation)

1920s Founds modern abstract algebra

1933 Dismissed by the Nazis, emigrates to the USA

1935 Dies in Bryn Mawr



"The most significant creative mathematical genius thus far produced since the higher education of women began." — Einstein, 1935

Noether's theorems (1918)

First theorem: every continuous symmetry of a physical system gives a conserved quantity.

Symmetry		**Conserved quantity**
Translation in time	$\Rightarrow$	Energy
Translation in space	$\Rightarrow$	Momentum
Rotation	$\Rightarrow$	Angular momentum

Second theorem: explained why energy behaves strangely in Einstein's general relativity, the problem Hilbert and Klein had invited her to solve.

Symmetries determine the physics.

Build the symmetries into the model, without losing information.

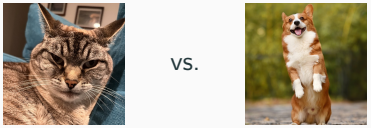
Equivalence: $x \sim y \Rightarrow f(x) \sim f(y)$



f: cat $\qquad$ f: cat ✓

A rotated cat is still a cat.

Expressivity: $f(x) \sim f(y) \Rightarrow x \sim y$



f: cat $\qquad$ f: cat ✗

Too much invariance: having fur is not enough to be a cat.

Goal: both at once, $f(x) \sim f(y) \iff x \sim y$.

Outline

Materials and Their Properties

GNNs for Geometric Machine Learning

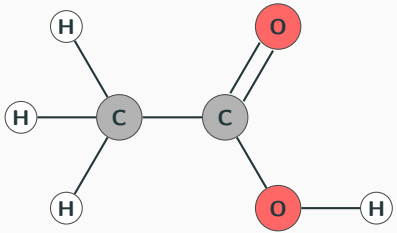
GNN-Based Diffusion Models

Materials and Their Properties

Molecules

A molecule is a graph $G = (V, E)$:

- **Nodes**: atoms, labelled by their element (C, O, H, ...)
- **Edges**: bonds, labelled by their type (single, double, triple, aromatic)
- **Valence rules**: each element forms a fixed number of bonds, counted with bond order (H: 1, O: 2, C: 4)



Acetic acid, CH_3COOH

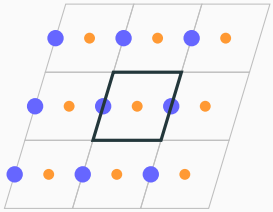
Crystals

A crystal is a unit cell repeated in space, $\mathcal{M} = (A, X, L)$:

- $A \in \{1, \ldots, Z_{\max}\}^N$: atom types
- $X \in [0, 1)^{3\times N}$: fractional coordinates
- $L \in \mathrm{GL}(3, \mathbb{R})$: lattice vectors as columns

L acts on the left, mapping fractional to Cartesian coordinates:

$$\{L(x_i + k) : k \in \mathbb{Z}^3\}$$



Molecules vs. inorganic crystals

	Molecules	**Inorganic crystals**
Extent	Finite set of atoms	Infinite
Periodicity	None	Periodic in 3 directions: a unit cell repeated by the lattice L
Elements	Mostly H, C, N, O, F, S, P	Almost the whole periodic table, including metals
Bonds	Covalent, directional, well-defined bond orders	Ionic, metallic or covalent; no unique bond list
Chemistry	Valence rules	Charge balance through oxidation states

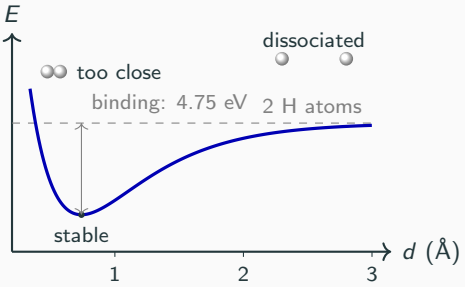
Total energy: definition

For fixed atom positions, the total energy E is the lowest solution of the static Schrödinger equation:

$$\hat{H}\psi = E\psi$$

E depends only on the geometry.

Example: H_2, E as a function of the bond length d



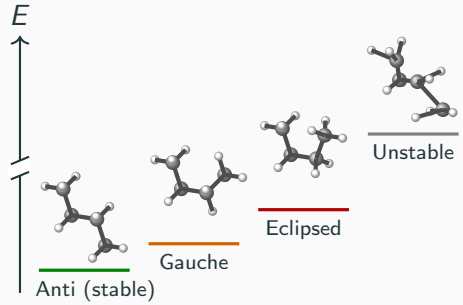
Minimum at $d_0 = 0.74$ Å: the stable geometry

Key property: total energy

One composition, several structures, each with its own total energy E: the lowest is stable. Distorted geometries have a much higher E: unstable.

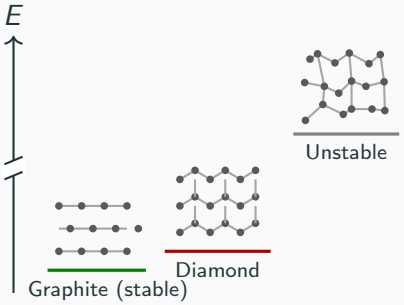
Molecules: conformers of butane

Same graph, rotation around the central C–C bond



Crystals: polymorphs of carbon

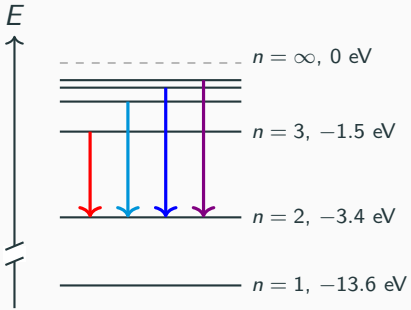
Same composition, different lattice and bonding



E is computed with density functional theory (DFT) (Kohn and Sham 1965). Energy gaps not to scale.

Energy levels: the hydrogen atom

Energies are the eigenvalues of the Hamiltonian, $\hat{H}\psi_n = E_n\psi_n$. For hydrogen, $E_n = -13.6\,\text{eV}/n^2$. An electron falling from level n to m emits a photon of energy $h\nu = E_n - E_m$.



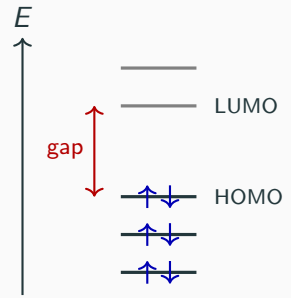
Emission spectrum (visible part)



Discrete levels ⇒ discrete lines
(Balmer series: $n \to 2$)

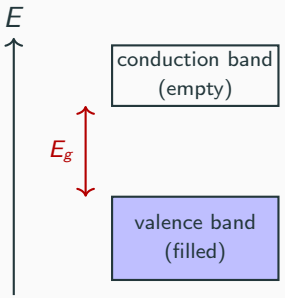
Key property: electronic gap

Molecules: HOMO–LUMO gap



Discrete orbitals, filled up to the HOMO.
Gap → reactivity, colour.

Crystals: band gap



Orbitals merge into continuous bands.
Metal: no gap. Semiconductor: Si, 1.1 eV.
Insulator: diamond, 5.5 eV.

What impacts these properties?

TL;DR: the energy

Structure → energy levels

- Composition: which elements
- Local environment: neighbours, bond lengths and angles
- Geometry: the 3D arrangement, not only the bonds

Energy levels → properties

- Stability: lowest level, E
- Formation energy: difference of E
- Elastic moduli: curvature E''
- Band gap: gap between levels

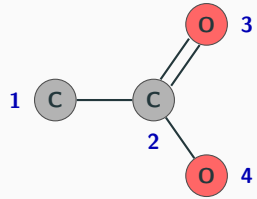
GNNs for Geometric Machine Learning

Graph Neural Networks

Graphs as data

A graph $G = (V, E)$:

- **Nodes**: atoms, with features h_v (element as a one-hot vector, ...)
- **Edges**: bonds, or neighbours within a cutoff for crystals, with features e_{uv} (bond type, distance)



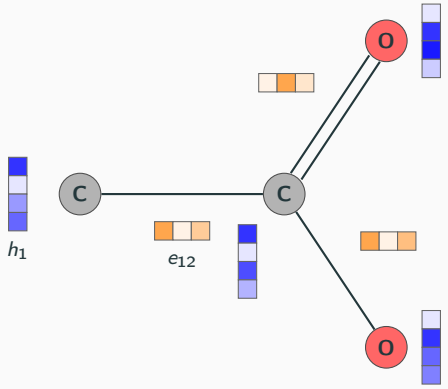
$$X = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{matrix} \mathrm{C} \\ \mathrm{C} \\ \mathrm{O} \\ \mathrm{O} \end{matrix} \qquad A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Acetic acid, hydrogens implicit. Blue: atom index. A holds the bond orders.

Node and edge features

Attach a vector to every node and every edge:

- **Node** v: $h_v \in \mathbb{R}^d$
  element (one-hot), electronegativity, charge, ...
- **Edge** (u, v): $e_{uv} \in \mathbb{R}^{d'}$
  bond order (one-hot), interatomic distance, ...



Blue: node vectors h_v. Orange: edge vectors e_{uv}.

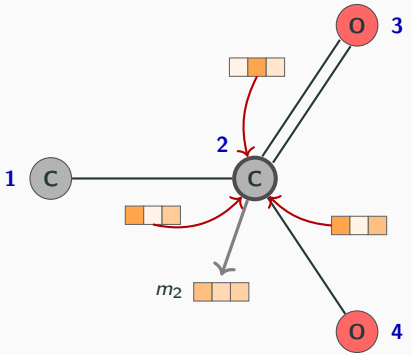
Aggregation: from edges to nodes

Each node gathers the vectors of its edges:

$$m_v = \square_{u \in \mathcal{N}(v)} e_{uv}$$

- $\mathcal{N}(v)$: neighbours of v
- $\square$: sum, mean or attention
- Independent of the order and number of neighbours

Example: $m_2 = \square(e_{12}, e_{32}, e_{42})$



Edge vectors aggregated onto node 2 (here: mean)

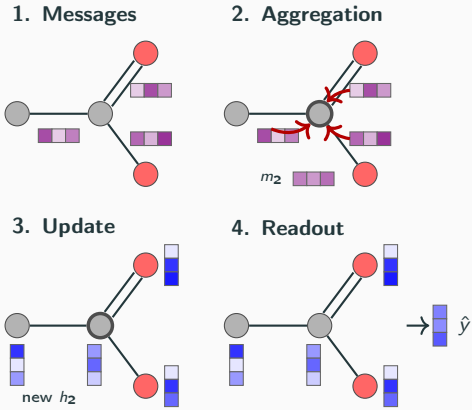
Message passing

General framework (Gilmer et al. 2017):

1. **Message formation**: each edge builds a message from its vector and its two nodes
2. **Aggregation**: each node combines its incoming messages
3. **Node update**: each node gets a new vector

Steps 1–3 are repeated T times.

4. **Readout**: combine all node vectors to predict the property



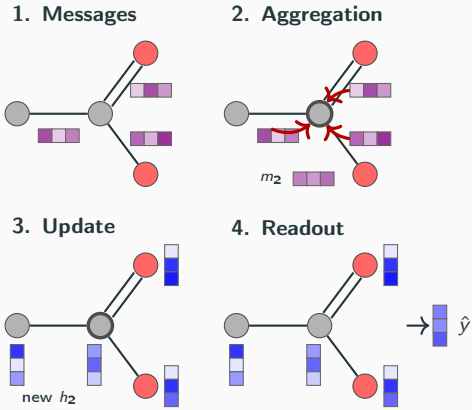
Purple: messages. Blue: node vectors.

Message passing neural network (MPNN)

For $t = 0, \ldots, T-1$:

1. $m_{uv}^{(t)} = M_t\big(h_u^{(t)}, h_v^{(t)}, e_{uv}\big)$
2. $m_v^{(t)} = \square_{u \in \mathcal{N}(v)}\, m_{uv}^{(t)}$
3. $h_v^{(t+1)} = U_t\big(h_v^{(t)}, m_v^{(t)}\big)$
4. $\hat{y} = R\big(\{h_v^{(T)} : v \in V\}\big)$

M_t, U_t: neural networks (e.g. MLPs)
R: permutation invariant (e.g. sum, then MLP)



Purple: messages. Blue: node vectors.

Types of GNN

MPNNs are one family of GNNs. The models below belong to it and differ by their choice of message, aggregation and update.

Model	**Message** m_{uv}	**Aggregation** $\square$	**Update** $h_v^{(t+1)}$
GCN (Kipf and Welling 2017)	$\frac{h_u}{\sqrt{d_u d_v}}$	sum	$\sigma(W m_v)$
GraphSAGE (Hamilton et al. 2017)	h_u	mean	$\sigma\big(W[h_v \,\|\, m_v]\big)$
GAT (Veličković et al. 2018)	$W h_u$	attention	$\sigma(m_v)$
GIN (Xu et al. 2019)	h_u	sum	$\mathrm{MLP}\big((1+\epsilon)\, h_v + m_v\big)$
CGCNN (Xie and Grossman 2018)	$\sigma(W_f z_{uv}) \odot g(W_s z_{uv})$	sum	$g(h_v + m_v)$

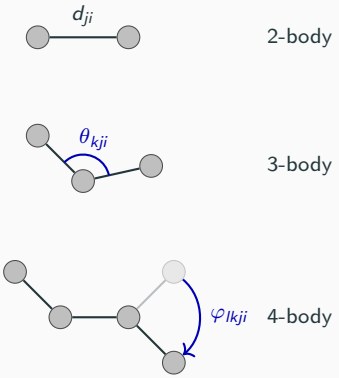
W: learned weights, σ, g: non-linearities, d_v: degree of v, $\|$: concatenation, $z_{uv} = h_u \,\|\, h_v \,\|\, e_{uv}$. GCN also includes v in its own neighbourhood. CGCNN: one of the first GNNs for crystals.

Triplets and quadruplets

Interactions involving more and more atoms:

- **Pairs** (j, i): distance d_{ji}
- **Triplets** (k, j, i): angle θ_{kji}
- **Quadruplets** (l, k, j, i): dihedral φ_{lkji}, e.g. butane conformers

With k neighbours per atom, the number of pairs, triplets and quadruplets grows as Nk, Nk^2, Nk^3.

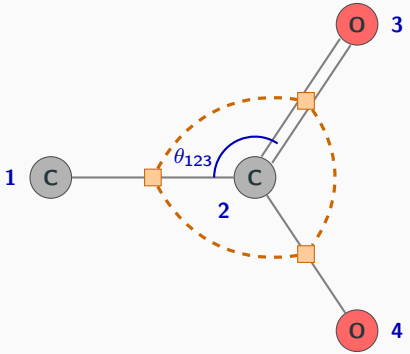


Higher-order information: the line graph

The graph G only sees pairs of atoms. Its line graph $L(G)$:

- **Nodes**: the bonds of G
- **Edges**: pairs of bonds sharing an atom, i.e. triplets of atoms (k, j, i)
- **Edge feature**: the bond angle θ_{kji}

Message passing on $L(G)$ updates the bond vectors with angle information.



Grey: G. Orange: $L(G)$, one node per bond.

Higher-order GNNs

Modern architectures add angles and dihedrals to message passing.

Model	**Order**	**How**	**Applied to**
DimeNet (Gasteiger, Groß, et al. 2020)	triplets	angles in bond messages	molecules
ALIGNN (Choudhary and DeCost 2021)	triplets	message passing on G and $L(G)$	molecules, crystals
GemNet (Gasteiger, Becker, et al. 2021)	quadruplets	adds dihedrals	molecules
M3GNet (Chen and Ong 2022)	triplets	3-body terms	crystals

Higher order: more geometric detail, but a higher cost.

GNNs for Geometric Machine Learning

Geometric Machine Learning for Material Science

Groups and group actions

A **group** G: a set of transformations that can be combined and undone.

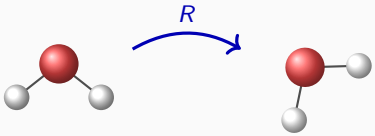
- Rotations $\mathrm{SO}(3)$
- Translations $(\mathbb{R}^3, +)$
- Permutations S_N

An **action** of G on a set S applies these transformations to the objects of S:

$$\alpha : G \times S \to S$$

Doing h then g is the same as doing $g \cdot h$.

Example: rotating a molecule



$G = \mathrm{SO}(3)$, S: atom positions

$$P = (p_1, \ldots, p_N)$$

$$\alpha(R, P) = (R\,p_1, \ldots, R\,p_N)$$

Equivalence under a group action

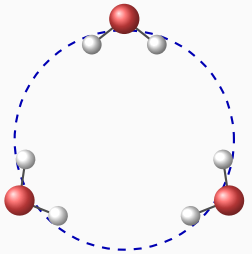
Two elements are **equivalent** if a group element maps one to the other:

$$s \sim s' \iff \exists g \in G,\ s' = \alpha(g, s)$$

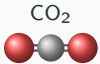
An equivalence relation, thanks to the group axioms:

- Reflexive: $s = \alpha(e, s)$
- Symmetric: $s = \alpha(g^{-1}, s')$
- Transitive: $\alpha(h, \alpha(g, s)) = \alpha(h \cdot g, s)$

Orbit: $[s] = \{\alpha(g, s) : g \in G\}$



orbit: same molecule, rotated



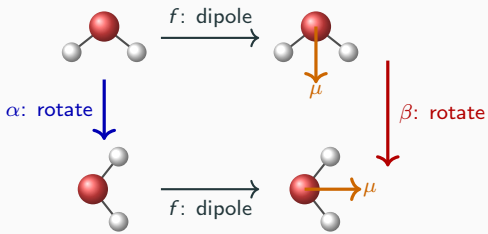
not equivalent

Rotations of a molecule are equivalent; a different molecule lies in another orbit.

Equivariance and invariance

Equivariance: vectors rotate too

$$f(RP) = R\,f(P)$$

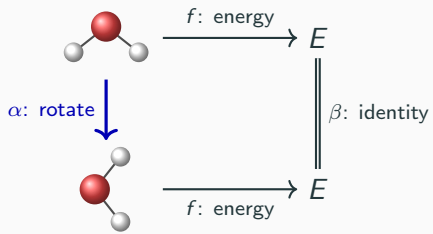


Rotate the molecule: the dipole μ rotates with it.

Also: forces, $F(RP) = R\,F(P)$

Invariance: scalars do not change

$$f(RP) = f(P)$$



Rotate the molecule: its energy does not change.

Also: band gap, formation energy, $\|\mu\|$

Both: $x \sim y \Rightarrow f(x) \sim f(y)$, with $\sim$ on outputs given by β (same g on both sides).

For scalars β is the identity, so $\sim$ is just $=$: invariance.

Which group?

Which group G should define when two structures are the same?

Choose G so that the total energy is conserved, $E\big(\alpha(g,P)\big) = E(P)$ for all $g \in G$.
E is the lowest eigenvalue of the Hamiltonian, nuclei (Z_I, p_I), electrons r_i:

$$\hat{H}_P = -\frac{1}{2}\sum_i \Delta_{r_i} - \sum_{i,I} \frac{Z_I}{\|r_i - p_I\|} + \sum_{i<j} \frac{1}{\|r_i - r_j\|} + \sum_{I<J} \frac{Z_I Z_J}{\|p_I - p_J\|}$$

The nuclei only enter through **distances** and **sums over atoms**:

- Rotations, reflections, translations keep all distances: same spectrum, same E
- Relabelling the atoms only reorders the sums: same $\hat{H}$

$$\boxed{G = \mathrm{E}(3) \times S_N}$$ rigid motions and permutations of the atoms

Geometric invariants

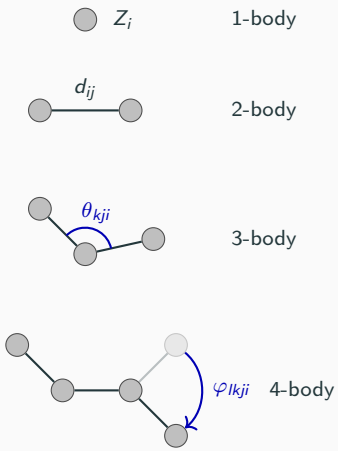
Quantities unchanged by rotations, translations and reflections:

- **1-body**: element Z_i
- **2-body**: distance $d_{ij} = \|p_i - p_j\|$
- **3-body**: unoriented angle $\theta_{kji} \in [0, \pi]$

$$\cos\theta_{kji} = \frac{(p_k - p_j)\cdot(p_i - p_j)}{d_{jk}\, d_{ji}}$$

- **4-body**: dihedral angle φ_{lkji}, between the planes (l, k, j) and (k, j, i)

The sign of φ flips under a reflection: only $|\varphi|$ is reflection invariant.



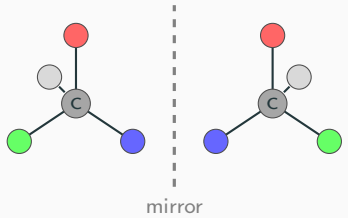
Note on chirality for biology

Mirror images (enantiomers) have the same energy: only the sign of φ differs.

But **life prefers one hand**:

- Proteins use only L-amino acids
- DNA, RNA use only D-sugars (right-handed helix)
- Receptors tell them apart: R-carvone smells of spearmint, S-carvone of caraway
- Drugs: often only one enantiomer is active

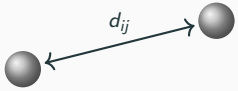
$\Rightarrow$ For biological properties, use $\mathrm{SE}(3)$ and keep the sign of φ.



Four different groups on one carbon: the two forms cannot be superimposed by a rotation.

Geometric GNNs: invariant vs. equivariant

Invariant models



Predict a **scalar**:
energy, band gap, formation energy

Equivariant models



Predict **vectors**:
forces, and positions for generation

Back to the introduction: **equivalence**

$$x \sim y \Rightarrow f(x) \sim f(y)$$



f: cat f: cat ✓

GNNs for Geometric Machine Learning

Expressivity of GNNs

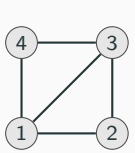
The graph isomorphism problem

$G_1 \cong G_2$ if there is a bijection $\pi : V_1 \to V_2$ with

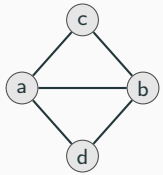
$$(u, v) \in E_1 \iff \big(\pi(u), \pi(v)\big) \in E_2$$

Adjacency matrices: $A_2 = PA_1P^\top$, P a permutation matrix.

- Same orbit under the action of S_N
- No polynomial algorithm known (quasi-polynomial: Babai 2016)



G_1



G_2

π: $1 \mapsto a$, $2 \mapsto c$, $3 \mapsto b$, $4 \mapsto d$

A permutation invariant GNN gives $f(G_1) = f(G_2)$ when $G_1 \cong G_2$.
Expressivity: does $f(G_1) \neq f(G_2)$ when $G_1 \ncong G_2$?

Isomorphism tests and distinguishability

A **test** computes a graph invariant $T(G)$:
$G_1 \cong G_2 \Rightarrow T(G_1) = T(G_2)$.

Distinguishability: T distinguishes G_1 and G_2 if $T(G_1) \neq T(G_2)$.

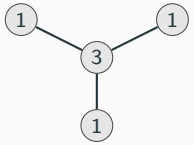
- Distinguished: $G_1 \ncong G_2$, for sure
- Not distinguished: maybe isomorphic

Example: degree test

$T_{\deg}(G) = \{\{\deg(v) : v \in V\}\}$



G_1



G_2

Nodes labelled by their degree:
$\{\{1,1,2,2\}\} \neq \{\{1,1,1,3\}\}$
Distinguished: $G_1 \ncong G_2$

Graph colouring (1-WL)

Weisfeiler–Leman colour refinement (Weisfeiler and Leman 1968):

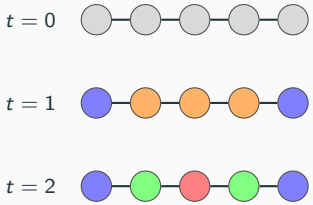
1. **Start**: all nodes get the same colour
2. **Refine**: new colour from the node's colour and the multiset of its neighbours' colours

$$c^{(t+1)}(v) = \text{hash}\Big(c^{(t)}(v),\ \{\{c^{(t)}(u) : u \in \mathcal{N}(v)\}\}\Big)$$

3. **Repeat** until the colours stop changing

Test: compare the colour histograms of G_1 and G_2.

After one step, the colours are the node degrees.



Path on 5 nodes: stable after $t = 2$.

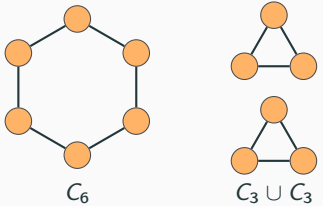
A counterexample: C_6 vs. $C_3 \cup C_3$

- $t = 0$: all nodes get the same colour
- $t = 1$: every node has two neighbours of the same colour, so all nodes get the same new colour
- Nothing ever splits: stable at once

Same histogram: 6 nodes, one colour.
1-WL does **not** distinguish them.

But C_6 is connected and $C_3 \cup C_3$ is not:
$C_6 \ncong C_3 \cup C_3$.

1-WL cannot separate regular graphs with the same number of nodes and the same degree.



Final 1-WL colours: identical.

Message passing GNNs vs. 1-WL

Without geometric information, with messages only along edges:

$$\textbf{1-WL:} \quad c_v^{(t+1)} = \text{hash}\Big(c_v^{(t)},\ \{\{(c_u^{(t)}, e_{uv}) : u \in \mathcal{N}(v)\}\}\Big)$$

$$\textbf{MPNN:} \quad h_v^{(t+1)} = U\Big(h_v^{(t)},\ \square_{u \in \mathcal{N}(v)}\ M\big(h_u^{(t)}, h_v^{(t)}, e_{uv}\big)\Big)$$

Theorem (Xu et al. 2019; Morris et al. 2019): if 1-WL does not distinguish G_1 and G_2, no MPNN does.

- Same information as 1-WL, but M, $\square$, U may not be injective
- Equality when they are: GIN (sum aggregation, MLP update)
- No MPNN distinguishes C_6 from $C_3 \cup C_3$

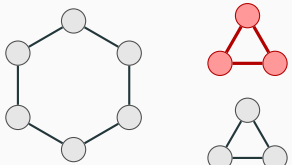
Higher-order tests: folklore k-WL

Folklore k-WL (k-FWL): colour k-tuples of nodes.

- **Start**: colour of a tuple = which of its nodes are equal or adjacent
- **Refine**: for each node w, look *jointly* at the tuples where w replaces one entry. For $k = 2$:

$$c(v_1, v_2) \leftarrow \text{hash}\big(c(v_1, v_2),\ m\big)$$
$$m = \{\!\{\big(c(w, v_2),\ c(v_1, w)\big) : w \in V\}\!\}$$

Hierarchy: 1-WL $<$ 2-FWL $<$ 3-FWL $< \ldots$
No fixed k is complete (Cai et al. 1992). Cost: n^k tuples.



2-FWL: an edge (v_1, v_2) and a node w adjacent to both form a triangle. It separates C_6 from $C_3 \cup C_3$.

Higher-order message passing vs. k-FWL

Pass messages between tuples of nodes, not single nodes:

Model	Messages between	Power	Cost
MPNN, GIN (Xu et al. 2019)	nodes, along edges	$=$ 1-WL	$O(\lvert E\rvert)$
k-GNN (Morris et al. 2019)	k-tuples	$=(k-1)$-FWL	$O(n^k)$
PPGN (Maron et al. 2019)	pairs, through every node w	$=$ 2-FWL	$O(n^3)$
Line graph, triplets	bonds sharing an atom	$\geq$ 1-WL*	$O(\lvert E\rvert\, d)$

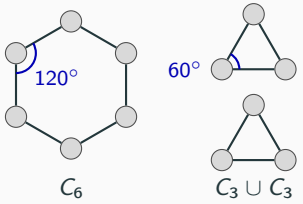
*Without geometry, still fails on C_6 vs. $C_3 \cup C_3$.

- 2-FWL looks *jointly* at $\big(c(w, v_2), c(v_1, w)\big)$, i.e. at triangles (v_1, w, v_2). Triplet message passing without geometry does not
- More power, but the cost grows as n^k (d: number of neighbours)

Expressivity of geometric GNNs

Graphs with positions: same structure up to permutations and $\mathrm{E}(3)$.

- **Geometric WL** (Joshi, Bodnar, et al. 2023): colour refinement that also sees the geometry of neighbours, an upper bound for geometric GNNs
- **Completeness with dihedral angles**: distances alone fail (Li et al. 2023); with distances, angles and dihedrals over two hops, invariant models are complete for connected structures, up to $\mathrm{SE}(3)$ and permutations (Gasteiger, Becker, et al. 2021; Wang et al. 2022)



Same bond lengths: distances along edges fail. Angles (3-body) separate them.

Remark: completeness is a property of the model family. A large enough GNN with injective updates separates any two structures, but a given trained network need not be injective.

Back to the introduction: **expressivity**

$$f(x) \sim f(y) \Rightarrow x \sim y$$



f: cat

vs.



f: dog ✓

GNN-Based Diffusion Models

Datasets

Dataset	Structures	Atoms per cell	Content
Perov-5 (Xie, Fu, et al. 2022)	18 928	5	perovskites ABX_3
Carbon-24 (Xie, Fu, et al. 2022)	10 153	6–24	carbon only
MP-20 (Xie, Fu, et al. 2022)	45 231	≤ 20	Materials Project
MPTS-52 (Jiao et al. 2023)	40 476	≤ 52	Materials Project, split by date
Alex-MP-20 (Zeni et al. 2025)	$\approx$ 608k	≤ 20	Materials Project + Alexandria

- Structures and energies computed with DFT, mostly from the Materials Project (Jain et al. 2013)
- Small benchmarks (Perov-5, Carbon-24, MP-20) to compare models; large datasets (Alex-MP-20) to train foundation models

Metrics for generation

No reference structure: judge the generated set itself.

- **Stable**: energy above the convex hull $< 0.1\,\mathrm{eV/atom}$, after DFT relaxation
- **Unique**: not a duplicate of another generated structure
- **Novel**: not in the training set
- **S.U.N.**: fraction that is stable, unique and novel (Zeni et al. 2025)

Why not a VAE or a GAN?

VAE: encode, then decode in one shot

(Noh et al. 2019; Ren et al. 2022)

- Decoder $z \mapsto (A, X, L)$ in a single pass
- Reconstruction loss needs a fixed atom order and unit cell
- Gaussian decoder averages the possible answers: overlapping atoms, low symmetry

Results

- MagGen: 94% of structures in P1 (Mal et al. 2025)
- FTCP: far behind CDVAE and MatterGen in S.U.N. (Zeni et al. 2025)

GAN: generator vs. discriminator

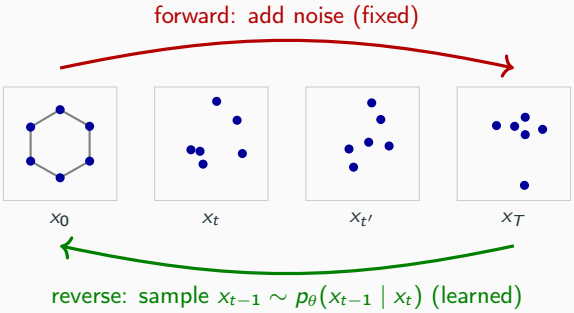
(Zhao, Al-Fahdi, et al. 2021; Zhao, Siriwardane, et al. 2023)

- No likelihood: nothing to tell how well the data is covered
- Unstable training, mode collapse: few prototypes
- Fixed-size arrays of atoms, a few space groups only

Results

- CubicGAN: cubic crystals only
- PGCGM: only 5.3% within 0.25 eV/atom of the hull

Diffusion: forward and reverse process



- **Forward**: $q(x_t \mid x_{t-1}) = \mathcal{N}\left(\sqrt{1-\beta_t}\, x_{t-1},\, \beta_t I\right)$, until $x_T \sim \mathcal{N}(0, I)$. In closed form: $x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, \epsilon$
- **Reverse**: the model estimates the distribution of the previous state, $p_\theta(x_{t-1} \mid x_t)$. Generation: sample $x_T \sim \mathcal{N}(0, I)$, then $x_{t-1} \sim p_\theta(x_{t-1} \mid x_t)$ down to x_0

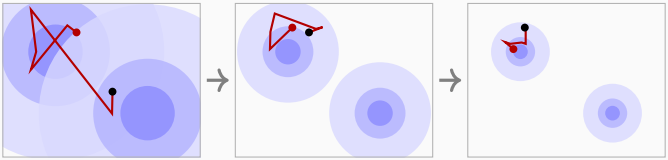
Annealed Langevin dynamics (Song and Ermon 2019)

Train one network on data blurred at noise levels $\sigma_1 > \cdots > \sigma_K$: $s_\theta(x, \sigma) \approx \nabla_x \log p_\sigma(x)$.

Sampling: start from noise, then for each σ_i, a few steps of

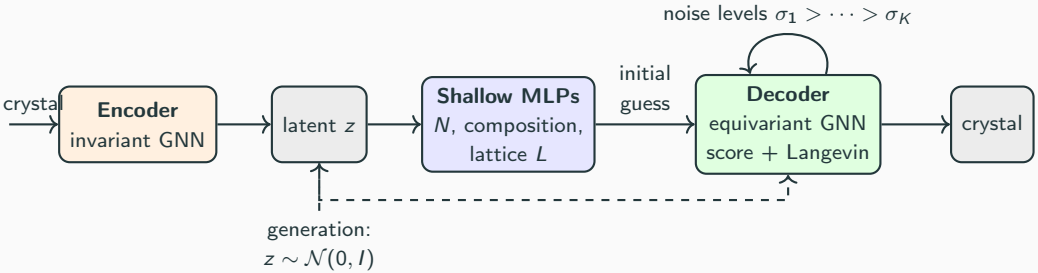
$$x \leftarrow x + \frac{\eta_i}{2} s_\theta(x, \sigma_i) + \sqrt{\eta_i}\, z, \qquad z \sim \mathcal{N}(0, I), \quad \eta_i \propto \sigma_i^2$$

- **Score term**: climbs the density. **Noise term**: keeps exploring
- **Large** σ: smooth density, reliable score, moves between modes
- **Small** σ: sharp density, refines the details



σ_1 large: explore σ_2 σ_K small: refine

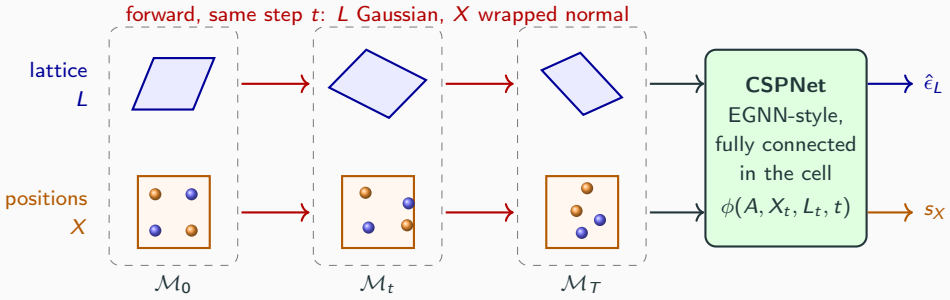
CDVAE (Xie, Fu, et al. 2022)



- **Invariant GNN encoder**: crystal $\to$ latent z, independent of rotations and atom order
- **Shallow networks**: initial guess from z: number of atoms, composition, lattice parameters
- **Equivariant GNN decoder**: score on coordinates (and atom types), conditioned on z; annealed Langevin refines the initial guess

DiffCSP (Jiao et al. 2023)

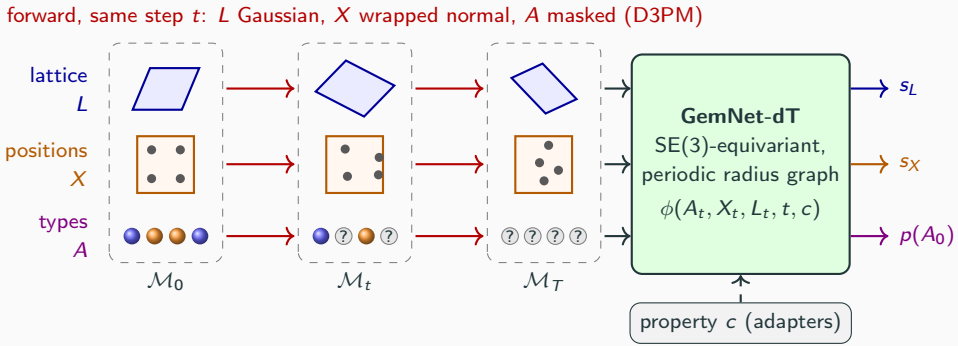
Task: crystal structure prediction, $p_\theta(X, L \mid A)$, atom types A given.



- **Joint**: one network sees L_t and X_t and predicts both (s_X: score on X); generation samples $\mathcal{M}_T \to \cdots \to \mathcal{M}_0$
- **Symmetry**: rotations only change L; the network sees $L^\top L$ and Fourier features of $x_j - x_i$

MatterGen (Zeni et al. 2025)

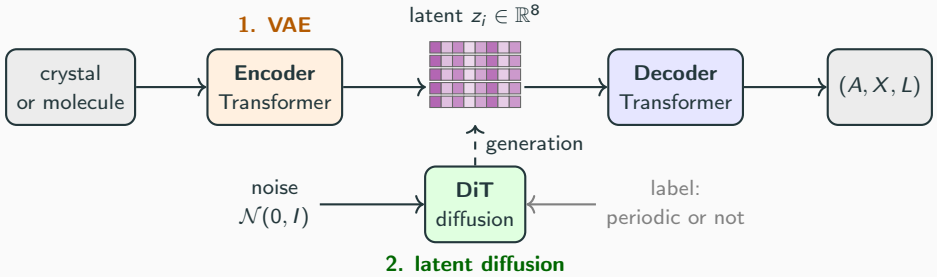
Task: de novo generation, $p_\theta(A, X, L)$, optionally conditioned on a property c.



- **Joint**: one network predicts scores on L, X and clean types A_0
- **Fine-tuning**: adapters and classifier-free guidance target a property c

ADiT (Joshi, Fu, et al. 2025)

Task: de novo generation of crystals *and* molecules with one model.



- **VAE**: one latent per atom, shared by molecules and crystals; diffusion generates, not the decoder
- **No equivariance built in**: plain Transformers, random rotations and translations as augmentation
- **Fast**: 10 000 samples in < 20 min on one GPU, $\approx 10\times$ faster than the baselines
- **S.U.N. on MP-20**: 5.3%, vs. 3.3% for DiffCSP and 2.8% for FlowMM

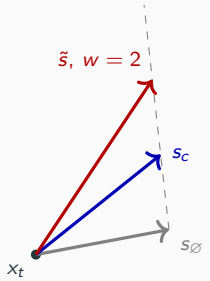
Classifier-free guidance (Ho and Salimans 2022)

Training: one network; c is randomly replaced by $\varnothing$, so it learns both $s_\theta(x, t, c)$ and $s_\theta(x, t, \varnothing)$.

Sampling: extrapolate from the unconditional score

$$\tilde{s} = s_\theta(x, t, \varnothing) + w\big[s_\theta(x, t, c) - s_\theta(x, t, \varnothing)\big]$$

- $w = 1$: conditional model; $w > 1$: stronger condition, less diversity
- By Bayes, $s_c - s_\varnothing \approx \nabla_x \log p_t(c \mid x)$: an implicit classifier, no extra model to train



guided score: further in the direction of c

MatterGen: c = chemical system, space group, magnetic density, band gap, ...

Limitations

Data and physics

- Labels come from DFT: its errors are learned too (e.g. graphite vs. diamond depends on the functional)
- Small cells (≤ 20 atoms), ordered crystals: no disorder, defects, surfaces
- Stability at 0 K: energy above hull is not synthesisability, temperature and kinetics are ignored
- Biased towards known chemistries

Inverse design

- One or two properties at a time
- Extreme targets lie outside the training data: poor extrapolation
- Pushing the property (guidance w) lowers stability and novelty

Summary

- **Molecules and crystals**: graphs with geometry; their properties follow from the total energy, solution of the Schrödinger equation
- **GNNs**: message formation, aggregation, node update, readout
- **Geometric ML**: the energy is invariant under rigid motions, permutations and change of unit cell; models build these symmetries in
- **Expressivity**: message passing is bounded by 1-WL; angles and dihedrals make geometric GNNs complete
- **Diffusion**: learn to sample the previous noisy state; joint diffusion on types, positions and lattice (CDVAE, DiffCSP, MatterGen), steered by classifier-free guidance

Thank you!

Questions?

Which group?

Which group G should define when two structures are the same?

Choose G so that the total energy is conserved:

$$E\big(\alpha(g, P)\big) = E(P) \qquad \forall g \in G$$

Group action on the Hamiltonian: conjugation

Nuclei at $P = (p_1, \ldots, p_M)$ with charges Z_I, electrons at $r_1, \ldots, r_n$:

$$\hat{H}_P = -\frac{1}{2}\sum_i \Delta_{r_i} - \sum_{i,I} \frac{Z_I}{\|r_i - p_I\|} + \sum_{i<j} \frac{1}{\|r_i - r_j\|} + \sum_{I<J} \frac{Z_I Z_J}{\|p_I - p_J\|}$$

Move the nuclei, $p_I \mapsto g\, p_I$ with $g \in \mathrm{E}(3)$:

- $\|r_i - g\, p_I\| = \|g^{-1} r_i - p_I\|$ and $\|g\, p_I - g\, p_J\| = \|p_I - p_J\|$
- The Laplacian Δ is unchanged by isometries

So $\hat{H}_{\alpha(g,P)}$ is $\hat{H}_P$ after the change of variables $r \mapsto g^{-1} r$: a **conjugation**

$$\hat{H}_{\alpha(g,P)} = U_g\, \hat{H}_P\, U_g^{-1}, \qquad U_g f = f \circ g^{-1}$$

Consequence: the energy is unchanged

Static Schrödinger equation: $\hat{H}_P \, \psi = E(P) \, \psi$.

Apply the group action, U_g on both sides:

$$\underbrace{U_g \, \hat{H}_P \, U_g^{-1}}_{\hat{H}_{\alpha(g,P)}} (U_g \psi) = E(P) \, (U_g \psi)$$

$U_g \psi$ solves the equation for the moved nuclei, with the same eigenvalue:

$$E\big(\alpha(g, P)\big) = E(P)$$

Permutations leave the Hamiltonian unchanged

Relabel the atoms with $\sigma \in S_M$, moving each charge with its position:

$$(Z_I, p_I) \mapsto \left(Z_{\sigma(I)}, p_{\sigma(I)}\right)$$

Nuclei only appear in sums over all atoms and over all pairs of atoms:

$$\sum_I \frac{Z_{\sigma(I)}}{\|r_i - p_{\sigma(I)}\|} = \sum_I \frac{Z_I}{\|r_i - p_I\|}, \qquad \sum_{I<J} \frac{Z_{\sigma(I)} Z_{\sigma(J)}}{\|p_{\sigma(I)} - p_{\sigma(J)}\|} = \sum_{I<J} \frac{Z_I Z_J}{\|p_I - p_J\|}$$

The order of the terms changes, not the sum:

$$\hat{H}_{\sigma\cdot(Z,P)} = \hat{H}_{(Z,P)} \quad \Rightarrow \quad E\big(\sigma \cdot (Z,P)\big) = E(Z,P)$$

Swapping positions alone only works between atoms of the same element.

Which group defines equivalence?

Two structures are the same if they lie in the same orbit: the group must contain every change of description.

Change	**Group**	**Molecules** (A, P)	**Crystals** (A, X, L)
Atom ordering	S_N	reorder atoms	reorder atoms
Rotation	$\mathrm{SO}(3)$	$P \mapsto RP$	$L \mapsto RL$
Translation	$\mathbb{R}^3$	$p_i \mapsto p_i + t$	$x_i \mapsto x_i + t \bmod 1$
Choice of unit cell	$\mathrm{SL}(3, \mathbb{Z})$	—	$(X, L) \mapsto (U^{-1}X \bmod 1,\ LU)$

- Molecules: $S_N \times \mathrm{SE}(3)$. Crystals: add $\mathrm{SL}(3, \mathbb{Z})$
- $\mathrm{E}(3)$ also includes reflections: it identifies a chiral structure with its mirror image

Pure VAEs and GANs in practice

Pure VAEs and GANs

Model	Type	Result
MagGen	VAE	94% in P1 (Mal et al. 2025)
PGCGM	GAN	5.3% within 0.25 eV/atom of hull (Zhao, Siriwardane, et al. 2023)
FTCP	VAE	far behind in S.U.N. (Zeni et al. 2025)
CDVAE	VAE + score	MatterGen: 1.8× more S.U.N.

S.U.N. on MP-20

same DFT setup (Joshi, Fu, et al. 2025)

Model	Type	S.U.N.
DiffCSP	diffusion	3.3%
FlowMM	flow matching	2.8%
FlowLLM	LLM + flow	4.7%
ADiT	VAE + latent diffusion	5.3%

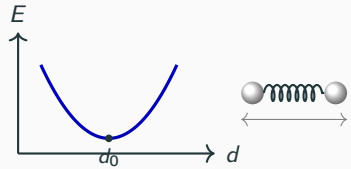
- VAEs are competitive only when diffusion generates: CDVAE, ADiT, Chemeleon2 (Park and Walsh 2025)
- A pure VAE (FTCP) still beats ion exchange on novelty (Szymanski and Bartel 2025)
- S.U.N. depends on hull, DFT settings, training set: compare within one table only

Key property: elastic moduli

How stiff is the structure? The curvature of the total energy around its minimum.

Molecules: bond stiffness

$$k = E''(d_0), \qquad \omega = \sqrt{k/\mu}$$

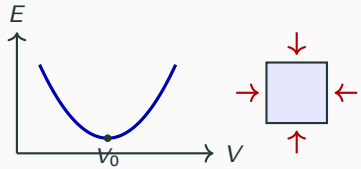


H_2: $k \approx 575$ N/m, vibration at 4400 cm^{-1}

(μ: reduced mass, ω: vibration frequency)

Crystals: elastic moduli

$$B = V_0\, E''(V_0), \qquad \sigma = C\,\varepsilon$$



Bulk modulus: diamond 443 GPa, NaCl 24 GPa

Elastic tensor C: stress σ vs. strain ε

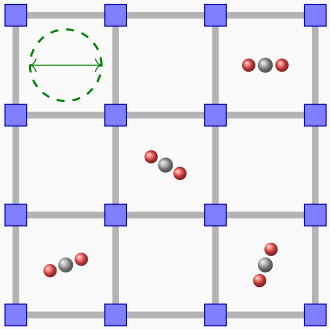
Key property: porosity

Some crystals contain empty channels and cages: zeolites, metal–organic frameworks (MOFs).

- **Void fraction**: $\phi = V_{\text{pores}}/V_{\text{cell}}$
- **Pore size**: largest sphere that fits, or can pass through
- **Surface area**: often > 3000 m^2/g for MOFs (record ≈ 7000)

Computed from the geometry (A, X, L) and atomic radii.

Uses: gas storage, CO_2 capture, catalysis.



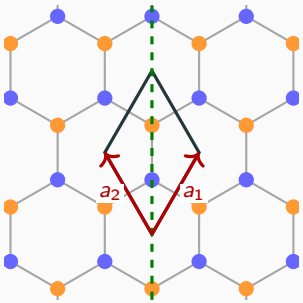
Metal nodes (blue), linkers (grey), CO_2 guests.
Dashed: pore size.

Space group symmetry

A **space group** is a discrete subgroup $G \subset \mathrm{E}(3)$.

Crystallographic restriction theorem. A rotation in G has order $n \in \{1, 2, 3, 4, 6\}$.

- No 5-fold symmetry
- 230 space groups



Mirror (dashed) swapping a_1 and a_2,

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{O}(2).$$

MatterGen: released models

Checkpoint	Condition c	Type
`mattergen_base`	none (Alex-MP-20)	base
`mp_20_base`	none (MP-20, for benchmarks)	base
`chemical_system`	set of elements, e.g. Li–Co–O	categorical
`space_group`	space group, 1 to 230	categorical
`dft_band_gap`	band gap (DFT)	continuous
`dft_mag_density`	magnetic density (DFT)	continuous
`ml_bulk_modulus`	bulk modulus (ML predicted)	continuous
`dft_mag_density_hhi_score`	magnetic density + HHI (supply risk)	two properties
`chemical_system_energy_above_hull`	elements + energy above hull	two properties

Fine-tuned from `mattergen_base` with adapters, sampled with classifier-free guidance. Several targets are the key properties seen earlier: band gap, elastic moduli, stability.

Two approaches

Probability based (Ho, Jain, et al. 2020)

Learn the reverse transitions

$$p_\theta(x_{t-1} \mid x_t) = \mathcal{N}\big(\mu_\theta(x_t, t),\, \sigma_t^2 I\big)$$

by predicting the added noise:

$$\mathcal{L} = \mathbb{E}\left\|\epsilon - \epsilon_\theta(x_t, t)\right\|^2$$

Score based (Song, Sohl-Dickstein, et al. 2021)

Learn the score $s_\theta(x, t) \approx \nabla_x \log p_t(x)$.
Forward SDE: $dx = f\,dt + g\,dw$
Reverse SDE:

$$dx = \left[f - g^2\,\nabla_x \log p_t(x)\right] dt + g\,d\bar{w}$$

Two views of the same model: $\epsilon_\theta(x_t, t) = -\sqrt{1-\bar{\alpha}_t}\; s_\theta(x_t, t)$.

Crystal structure prediction vs. generation

A crystal $\mathcal{M} = (A, X, L)$: atom types, fractional coordinates, lattice.

Crystal structure prediction (CSP)

$$p_\theta(X, L \mid A)$$

- Composition given (e.g. NaCl)
- Find the stable positions and lattice
- Compared to the known structure: match rate, RMSE
- DiffCSP (Jiao et al. 2023)

Generation (de novo)

$$p_\theta(A, X, L)$$

- Composition generated too
- New materials, possibly conditioned on a property
- No reference: valid, stable, unique, novel
- CDVAE (Xie, Fu, et al. 2022), MatterGen (Zeni et al. 2025)

Key property: formation energy

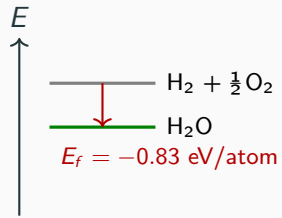
Energy change when forming a compound from its elements:

$$E_f = E(\text{compound}) - \sum_i n_i \, E(\text{element } i)$$

$E_f < 0$: the compound is more stable than its separated elements.

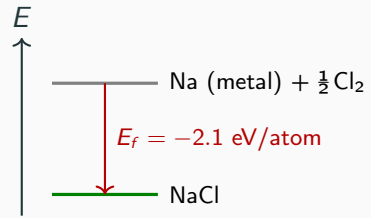
Molecules: water

$E_f = \frac{1}{3}\left[E(H_2O) - E(H_2) - \frac{1}{2}E(O_2)\right]$



Crystals: rock salt

$E_f = \frac{1}{2}\left[E(NaCl) - E(Na) - \frac{1}{2}E(Cl_2)\right]$



References i

- Babai, László (2016). **"Graph Isomorphism in Quasipolynomial Time".** In: *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 684–697.
- Cai, Jin-Yi, Martin Fürer, and Neil Immerman (1992). **"An optimal lower bound on the number of variables for graph identification".** In: *Combinatorica* 12.4, pp. 389–410.
- Chen, Chi and Shyue Ping Ong (2022). **"A universal graph deep learning interatomic potential for the periodic table".** In: *Nature Computational Science* 2, pp. 718–728.
- Choudhary, Kamal and Brian DeCost (2021). **"Atomistic Line Graph Neural Network for improved materials property predictions".** In: *npj Computational Materials* 7, p. 185.
- Gasteiger, Johannes, Florian Becker, and Stephan Günnemann (2021). **"GemNet: Universal Directional Graph Neural Networks for Molecules".** In: *Advances in Neural Information Processing Systems*.

References ii

- Gasteiger, Johannes, Janek Groß, and Stephan Günnemann (2020). **"Directional Message Passing for Molecular Graphs".** In: *International Conference on Learning Representations*.
- Gilmer, Justin et al. (2017). **"Neural Message Passing for Quantum Chemistry".** In: *International Conference on Machine Learning*, pp. 1263–1272.
- Hamilton, William L., Rex Ying, and Jure Leskovec (2017). **"Inductive Representation Learning on Large Graphs".** In: *Advances in Neural Information Processing Systems*.
- Ho, Jonathan, Ajay Jain, and Pieter Abbeel (2020). **"Denoising Diffusion Probabilistic Models".** In: *Advances in Neural Information Processing Systems*.
- Ho, Jonathan and Tim Salimans (2022). **"Classifier-Free Diffusion Guidance".** In: *arXiv preprint arXiv:2207.12598*.
- Jain, Anubhav et al. (2013). **"Commentary: The Materials Project: A materials genome approach to accelerating materials innovation".** In: *APL Materials* 1.1, p. 011002.

References iii

- Jiao, Rui et al. (2023). **"Crystal Structure Prediction by Joint Equivariant Diffusion".** In: *Advances in Neural Information Processing Systems*.
- Joshi, Chaitanya K., Cristian Bodnar, et al. (2023). **"On the Expressive Power of Geometric Graph Neural Networks".** In: *International Conference on Machine Learning*.
- Joshi, Chaitanya K., Xiang Fu, et al. (2025). **"All-atom Diffusion Transformers: Unified generative modelling of molecules and materials".** In: *International Conference on Machine Learning*.
- Kipf, Thomas N. and Max Welling (2017). **"Semi-Supervised Classification with Graph Convolutional Networks".** In: *International Conference on Learning Representations*.
- Kohn, W. and L. J. Sham (1965). **"Self-Consistent Equations Including Exchange and Correlation Effects".** In: *Physical Review* 140.4A, A1133–A1138.
- Li, Zian et al. (2023). **"Is Distance Matrix Enough for Geometric Deep Learning?"** In: *Advances in Neural Information Processing Systems*.

References iv

- Mal, Sourav et al. (2025). **"DiffCrysGen: A Generative Diffusion Model for Accelerated Design of Inorganic Crystalline Materials".** In: *arXiv preprint arXiv:2510.12329*.
- Maron, Haggai et al. (2019). **"Provably Powerful Graph Networks".** In: *Advances in Neural Information Processing Systems*.
- Morris, Christopher et al. (2019). **"Weisfeiler and Leman Go Neural: Higher-Order Graph Neural Networks".** In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33, pp. 4602–4609.
- Noh, Juhwan et al. (2019). **"Inverse Design of Solid-State Materials via a Continuous Representation".** In: *Matter* 1.5, pp. 1370–1384.
- Park, Hyunsoo and Aron Walsh (2025). **"Guiding Generative Models to Uncover Diverse and Novel Crystals via Reinforcement Learning".** In: *arXiv preprint arXiv:2511.07158*.

References

- Ren, Zekun et al. (2022). **"An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties".** In: *Matter* 5.1, pp. 314–335.
- Song, Yang and Stefano Ermon (2019). **"Generative Modeling by Estimating Gradients of the Data Distribution".** In: *Advances in Neural Information Processing Systems*.
- Song, Yang, Jascha Sohl-Dickstein, et al. (2021). **"Score-Based Generative Modeling through Stochastic Differential Equations".** In: *International Conference on Learning Representations*.
- Szymanski, Nathan J. and Christopher J. Bartel (2025). **"Establishing baselines for generative discovery of inorganic crystals".** In: *Materials Horizons* 12, pp. 8000–8011.
- Veličković, Petar et al. (2018). **"Graph Attention Networks".** In: *International Conference on Learning Representations*.

References vi

- Wang, Limei et al. (2022). **"ComENet: Towards Complete and Efficient Message Passing for 3D Molecular Graphs".** In: *Advances in Neural Information Processing Systems*.
- Weisfeiler, Boris and Andrei Leman (1968). **"The reduction of a graph to canonical form and the algebra which appears therein".** In: *Nauchno-Technicheskaya Informatsia* 2.9, pp. 12–16.
- Xie, Tian, Xiang Fu, et al. (2022). **"Crystal Diffusion Variational Autoencoder for Periodic Material Generation".** In: *International Conference on Learning Representations*.
- Xie, Tian and Jeffrey C. Grossman (2018). **"Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties".** In: *Physical Review Letters* 120.14, p. 145301.
- Xu, Keyulu et al. (2019). **"How Powerful are Graph Neural Networks?"** In: *International Conference on Learning Representations*.

- Zeni, Claudio et al. (2025). **"A generative model for inorganic materials design".** In: *Nature* 639, pp. 624–632.
- Zhao, Yong, Mohammed Al-Fahdi, et al. (2021). **"High-Throughput Discovery of Novel Cubic Crystal Materials Using Deep Generative Neural Networks".** In: *Advanced Science* 8.20, p. 2100566.
- Zhao, Yong, Edirisuriya M. Dilanga Siriwardane, et al. (2023). **"Physics guided deep learning for generative design of crystal materials with symmetry constraints".** In: *npj Computational Materials* 9, p. 38.