Whitney-Form-Based Cellular Sheaves: Construction, Laplacians, and Flat-Transport Moduli

Yu Chen

2026.11.22

1 Motivation and Background

2 Construction

3 More Properties

4 Molecular generation

Cellular sheaves on a simplicial complex

A triangle and its boundary have the same vertices and edges. The filled triangle can also carry a face signal.

Definition

A real cellular sheaf is a covariant functor $\mathcal{F}\colon P(K) \to \mathrm{Vect}_{\mathbb{R}}$.

- **Domain.** $P(K)$ orders the nonempty simplices by face inclusion.
- **Stalk.** Each simplex σ carries a vector space $\mathcal{F}_\sigma$.
- **Restriction.** Each $\sigma \le \tau$ gives a linear map $\mathcal{F}_{\sigma,\tau}\colon \mathcal{F}_\sigma \to \mathcal{F}_\tau$.

For every $\rho \le \sigma \le \tau$,

$$\mathcal{F}_{\sigma,\sigma} = \mathrm{id}_{\mathcal{F}_\sigma}, \qquad \mathcal{F}_{\rho,\tau} = \mathcal{F}_{\sigma,\tau} \circ \mathcal{F}_{\rho,\sigma}.$$

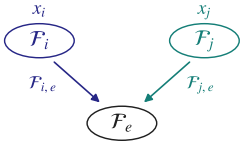
On a graph, the only strict face inclusions are $i \le e$ from vertices to incident edges.

Hansen & Ghrist, Toward a Spectral Theory of Cellular Sheaves

What is a graph sheaf neural network?

A GNN passes node features along edges. In a graph sheaf neural network (SNN), every vertex and edge has its own inner-product space (stalk). Restriction maps bring both endpoint features into the edge stalk for comparison. Their adjoints send the difference back to update the vertices.

- **Vertex signals.** $x \in C^0(G;\mathcal{F}) = \bigoplus_i \mathcal{F}_i$ is one column of X^ℓ.
- **Compare on $e = \{i, j\}$.** Form the difference $(\delta x)_e = \mathcal{F}_{i,e} x_i - \mathcal{F}_{j,e} x_j \in \mathcal{F}_e$.
- **Return to i.** Edge e contributes $\mathcal{F}_{i,e}^*(\delta x)_e \in \mathcal{F}_i$.



Compare in the shared edge space

One Laplacian-based layer. $L_{\mathcal{F}} = \delta^* \delta$ sums signed edge contributions at each vertex.

$$X^{\ell+1} = \rho_\ell\left(X^\ell W_0^\ell - L_{\mathcal{F}} X^\ell W_1^\ell\right).$$

W_0^ℓ, W_1^ℓ mix channels and ρ_ℓ is an activation. Stacking these trainable layers gives the SNN.

Hansen & Gebhart (2020) Bodnar et al. (2022)

Two limitations of common sheaf neural networks

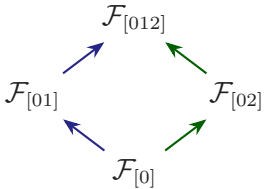
- **Domain.** Graph sheaves compare vertices with edges. Higher-dimensional complexes also require edge-to-face maps and higher incidences.
- **Source of the maps.** Common models learn restrictions from node features. How can geometry and topology specify them across dimensions?

Graph sheaves: Bodnar et al., 2022 Hypergraph extension: Duta et al., 2023 Physical prior: Hu et al., 2025

Which priors give valid restriction maps?

- **Geometry.** **Whitney forms** measure directional response on each simplex.
- **Topology.** **Flat transport** specifies how local frames compare along paths.
- **Compatibility.** Metric factors cancel at intermediate simplices, so restrictions compose.

On a filled triangle, two routes must agree



- Start with a vector at vertex 0.
- Restrict along either edge to face $[012]$. Both routes must agree for every input.

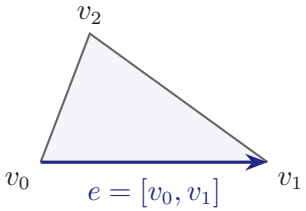
$$\mathcal{F}_{01,012}\mathcal{F}_{0,01} = \mathcal{F}_{02,012}\mathcal{F}_{0,02}.$$

Assigning a matrix to each arrow independently can violate this equality.

How can triangle shape enter a compatible restriction? Start with one edge measurement.

A Whitney 1-form for one oriented edge

On a triangle $T = [v_0, v_1, v_2]$, let λ_i be the barycentric coordinate of vertex v_i.



The basis 1-form on T

$$\mathcal{W}(e) = \lambda_0 \, d\lambda_1 - \lambda_1 \, d\lambda_0.$$

- Along e, set $\lambda_1 = t$, $\lambda_0 = 1 - t$. The form reduces to dt.
- Its integral is 1 on e and 0 on the other two edges.

Arnold, Falk & Winther (2006), § 4.3

Whitney forms on oriented p-simplices

Let $\sigma = [v_{i_0}, \ldots, v_{i_p}]$ be an oriented face of a simplex T.

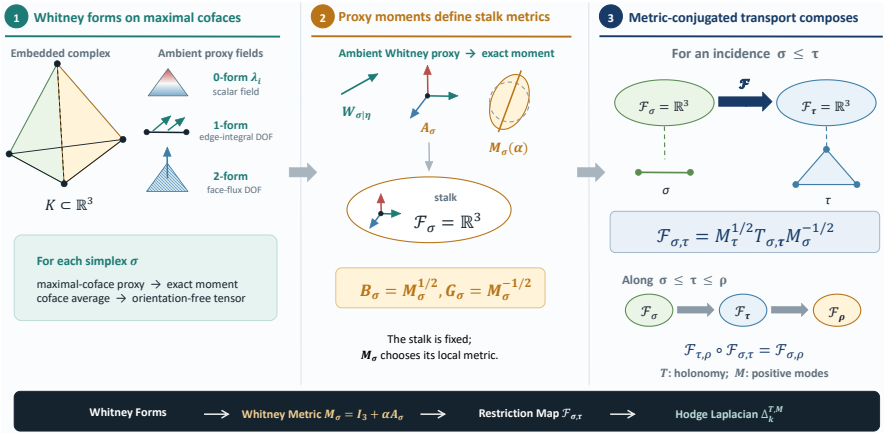
Whitney basis form on T

$$\mathcal{W}(\sigma) = p! \sum_{j=0}^{p} (-1)^j \lambda_{i_j} \, d\lambda_{i_0} \wedge \cdots \wedge \widehat{d\lambda_{i_j}} \wedge \cdots \wedge d\lambda_{i_p}.$$

- The hat omits one factor. For $p = 0$, this gives the usual nodal basis $\mathcal{W}(v_i) = \lambda_i$.
- With one orientation fixed on each p-face, $\int_{\sigma'} \mathcal{W}(\sigma) = \delta_{\sigma\sigma'}$.

Arnold, Falk & Winther (2006), § 4.3

Geometry and flat transport



On edge e: $M_e = I + \alpha A_e$

$T_{\sigma,\tau}$**: prescribed flat transport**

1 Motivation and Background

2 Construction

3 More Properties

4 Molecular generation

From the Whitney form to a directional response

For a face $\sigma \le \eta$, the barycentric functions λ_i and their differentials $d\lambda_i$ define the Whitney p-form $\mathcal{W}(\sigma)|_\eta$.

- **Normalize.** With median positive edge length ℓ_K, set $\widetilde{\omega} = \ell_K^p \, \mathcal{W}(\sigma)|_\eta$. The factor ℓ_K^p cancels the length scaling of its p differential factors.
- **Read an ambient direction.** Extend each $d\lambda_i$ by zero on normal vectors. Use the Euclidean Hodge star $\star_m$ on oriented $\mathbb{R}^m$ to select the lower degree.

$$\zeta = \begin{cases} \widetilde{\omega}, & p \le m/2, \\ \star_m \widetilde{\omega}, & p > m/2, \end{cases} \qquad r = \min(p, m-p).$$

For $r > 0$, $C(x)u = \iota_u \zeta(x) = \zeta(x)(u, \cdot\,)$ inserts u in the first slot.
For $r = 0$, set $C(x) = \zeta(x) I_m$.

From a Whitney form to a stalk metric

For a direction u, average its squared response over the coface η.

$$u^{\mathsf{T}} Q_{\sigma|\eta} u = \frac{1}{|\eta|} \int_{\eta} \|C(x)u\|^2 \, dx, \qquad Q_{\sigma|\eta} = \frac{1}{|\eta|} \int_{\eta} C(x)^* C(x) \, dx.$$

Here C^* is the Euclidean adjoint.

Average over all maximal cofaces containing σ

Let $\mathrm{MaxStar}(\sigma)$ be the maximal simplices containing σ, and set $N_\sigma = |\,\mathrm{MaxStar}(\sigma)|$.

$$\boxed{A_\sigma = \frac{1}{N_\sigma} \sum_{\eta \in \mathrm{MaxStar}(\sigma)} Q_{\sigma|\eta}}, \qquad M_\sigma = I_m + \alpha A_\sigma.$$

- For an edge in two maximal triangles, $A_e = \frac{1}{2}(Q_{e|\eta_1} + Q_{e|\eta_2})$.
- Every term is positive semidefinite, so $M_\sigma \succ 0$ for $\alpha \geq 0$.

Definition of the Whitney-form-based cellular sheaf

On a complex K realized in $\mathbb{R}^m$, fix $\alpha \geq 0$ and use the Whitney metrics $M_\sigma = I_m + \alpha A_\sigma$.

- **Stalks.** Assign $\mathcal{F}_\sigma^{W,T} = \mathbb{R}^m$ with its Euclidean inner product to each simplex.
- **Flat transport.** Choose $T_{\sigma,\tau} \in GL(m, \mathbb{R})$ with $T_{\sigma,\sigma} = I_m$ and $T_{\sigma,\tau} T_{\rho,\sigma} = T_{\rho,\tau}$ for $\rho \leq \sigma \leq \tau$.

Restriction for every face inclusion $\sigma \leq \tau$

$$\boxed{\mathcal{F}_{\sigma,\tau}^{W,T} = M_\tau^{1/2}\, T_{\sigma,\tau}\, M_\sigma^{-1/2}} \;:\; \mathcal{F}_\sigma^{W,T} \longrightarrow \mathcal{F}_\tau^{W,T}.$$

$$\underbrace{s_\sigma}_{\text{source stalk}} \xmapsto{M_\sigma^{-1/2}} \underbrace{u_\sigma}_{\text{source coefficients}} \xmapsto{T_{\sigma,\tau}} \underbrace{T_{\sigma,\tau} u_\sigma}_{\text{target coefficients}} \xmapsto{M_\tau^{1/2}} \underbrace{\mathcal{F}_{\sigma,\tau}^{W,T} s_\sigma}_{\text{target stalk}}.$$

Composition. Since $M_\sigma^{-1/2} M_\sigma^{1/2} = I_m$, we have $\mathcal{F}_{\sigma,\tau}^{W,T} \mathcal{F}_{\rho,\sigma}^{W,T} = \mathcal{F}_{\rho,\tau}^{W,T}$.

Barycentric gradients for the proxy field

To compute M_σ in $\mathbb{R}^3$, fix a coface $\eta = [v_0, \dots, v_q]$. Represent each $d\lambda_i$ by its tangent vector g_i.

$$d\lambda_i(u) = g_i \cdot u, \qquad g_i = \nabla_\eta \lambda_i.$$

Compute g_i from the vertices of η

For a nondegenerate coface with $q > 0$, row a of D_η is $(v_a - v_0)^\mathsf{T}$.

$$g_i = D_\eta^\mathsf{T}(D_\eta D_\eta^\mathsf{T})^{-1} e_i \quad (1 \le i \le q), \qquad g_0 = -\sum_{i=1}^{q} g_i.$$

Example. For an oriented edge $[ij] \le \eta$,

$$\underbrace{\lambda_i\, d\lambda_j - \lambda_j\, d\lambda_i}_{\text{Whitney 1-form on } \eta} \quad \overset{\sharp}{\longmapsto} \quad \underbrace{\lambda_i g_j - \lambda_j g_i}_{\text{vector proxy on } \eta}.$$

Whitney proxy fields in three dimensions

In oriented $\mathbb{R}^3$, use the gradients g_a of the same coface η. Write $D_{abc} = g_a \cdot (g_b \times g_c)$.

p	Simplex σ	Field	Whitney proxy $w_{\sigma\|\eta}(x)$
0	$[i]$	scalar	λ_i
1	$[ij]$	vector	$\lambda_i g_j - \lambda_j g_i$
2	$[ijk]$	vector	$2\big(\lambda_i(g_j \times g_k) - \lambda_j(g_i \times g_k) + \lambda_k(g_i \times g_j)\big)$
3	$[ijkl]$	scalar	$6\big(\lambda_i D_{jkl} - \lambda_j D_{ikl} + \lambda_k D_{ijl} - \lambda_l D_{ijk}\big)$

For $\widetilde{w} = \ell_K^p w_{\sigma|\eta}$, C^*C is $\widetilde{w}\widetilde{w}^\mathsf{T}$ (vector) or $|\widetilde{w}|^2 I_3$ (scalar).

Exact coface average. On a q-coface, set $C_a = C(v_a)$ and $S = \sum_{a=0}^{q} C_a$.

$$Q_{\sigma|\eta} = \frac{S^*S + \sum_{a=0}^{q} C_a^* C_a}{(q+1)(q+2)}.$$

Why include a flat transport?

Write $B_\sigma = M_\sigma^{1/2}$. Using the metrics alone gives

$$\mathcal{F}^{W,I}_{\sigma,\tau} = B_\tau B_\sigma^{-1}, \qquad B_\tau^{-1}\mathcal{F}^{W,I}_{\sigma,\tau} B_\sigma = I_m.$$

- In coordinates $y_\sigma = B_\sigma^{-1} x_\sigma$, every map is the identity.
- The underlying linear sheaf is isomorphic to the constant sheaf.

Flat transport allows nontrivial loop holonomy

With $\mathcal{F}^{W,T}_{\sigma,\tau} = B_\tau T_{\sigma,\tau} B_\sigma^{-1}$, the product around a loop γ based at b is

$$\mathcal{F}^{W,T}(\gamma) = B_b\, T(\gamma)\, B_b^{-1}.$$

A loop with $T(\gamma) \neq I_m$ prevents a global trivialization.

- The transport specifies the local system.
- The Whitney metrics weight directional discrepancies.

From restriction maps to the sheaf Laplacian

- **Cochains.** A k-cochain assigns a stalk vector to each oriented k-simplex.
- **Coboundary.** For each $(k+1)$-simplex, sum the restricted face vectors with their incidence signs.

Signed coboundary and Laplacian

$$(C_k)_{\tau,\sigma} = [\sigma : \tau]\mathcal{F}^{W,T}_{\sigma,\tau}, \qquad C_{k+1}C_k = 0.$$

$$L_k^{W,T} = C_k^T C_k + C_{k-1}C_{k-1}^T, \qquad L_0^{W,T} = C_0^T C_0.$$

For a codimension-two face, the two restriction paths agree and their incidence signs cancel.

1 Motivation and Background

2 Construction

3 More Properties

4 Molecular generation

The ANM Hessian describes soft protein motions



1URP, chain A. The inset enlarges the spring between residues 1 and 33.

- In the **anisotropic network model** (ANM), springs connect nearby residues.
- Let $E(u)$ be the spring energy for residue displacements $u = (u_i)_i$ from the native structure.
- At the native structure, the **Hessian** is the matrix of second derivatives of E,

$$H = \nabla_u^2 E(u)\big|_{u=0}.$$

A small displacement costs $\Delta E \approx \frac{1}{2} u^{\mathsf{T}} H u$. Eigenvectors with small positive eigenvalues give low-energy collective motions.

PDB 1URP Atilgan et al. (2001) Hu et al., 1URP modes

The Whitney graph limit recovers a weighted ANM Hessian

Pure graph, nonzero edge lengths, no isolated vertices, $T = I_m$

D_0 takes vertex displacements (degree 0) to differences on oriented edges (degree 1).

$$(\mathsf{D}_0 u)_{[i,j]} = u_j - u_i, \qquad H_\gamma = \mathsf{D}_0^{\mathsf{T}} \operatorname{diag}_e(\gamma_e P_e) \mathsf{D}_0.$$

$P_e = n_e n_e^{\mathsf{T}}$ projects onto the unit edge direction n_e. The weight $\gamma_e = 3(\ell_G / r_e)^2$ uses the edge length r_e and the median edge length ℓ_G.

Untwisted 0-Laplacian

$$\boxed{L_0^{W,I}(\alpha) = \big(1 - \theta(\alpha)\big)(L_G \otimes I_m) + \theta(\alpha) H_\gamma, \qquad \theta(\alpha) = \frac{\alpha}{\alpha + 3}}.$$

- At $\alpha = 0$, diffusion penalizes all components equally.
- In 3D, the limit as $\alpha \to \infty$ is the ANM Hessian with spring weights γ_e.

Flat transport and loop holonomy

Fix a finite connected simplicial complex K and a matrix group G.

- **Flat transports.** $\mathrm{Flat}_G(K)$ consists of assignments $T_{\sigma,\tau} \in G$ to face inclusions satisfying
$$T_{\sigma,\sigma} = I, \qquad T_{\tau,\eta}T_{\sigma,\tau} = T_{\sigma,\eta} \quad (\sigma \le \tau \le \eta).$$
- **Frame changes.** Choose $H_\sigma \in G$ at each simplex. The gauge group $\mathcal{G}_K = \prod_{\sigma\in K} G$ acts by $T_{\sigma,\tau} \mapsto H_\tau T_{\sigma,\tau} H_\sigma^{-1}$.

One point represents a flat transport up to local frame changes

$$\boxed{\mathcal{M}_{\mathrm{flat}}(K;G) := \mathrm{Flat}_G(K)/\mathcal{G}_K} \;\cong\; \mathrm{Hom}(\pi_1(|K|,b),G)/G.$$

At the basepoint b, the same H_b conjugates all loop matrices of $\rho_T\colon \pi_1(|K|,b) \to G$.

Equal homology can hide different loop relations

$X = (S^1 \vee S^1) \vee S^2$ $Y = T^2$



- Both spaces have integer homology $H_0 \cong \mathbb{Z},\ H_1 \cong \mathbb{Z}^2,\ H_2 \cong \mathbb{Z}$, with $H_k = 0$ for $k > 2$.
- Their fundamental groups are

$$\pi_1(X) = F_2 = \langle a, b \rangle, \qquad \pi_1(Y) = \mathbb{Z}^2 = \langle a, b \mid aba^{-1}b^{-1} = 1 \rangle.$$

On the torus, the two loop transports must commute.

Which transport parameters does a model need to learn?

Suppose a prediction depends only on the flat transport class. Changing the common frame identifies equivalent loop rotations $A, B \in SO(3)$.

$$(A, B) \sim (HAH^{-1}, HBH^{-1}), \qquad H \in SO(3).$$

- **Wedge.** $\mathcal{M}_{\text{flat}}(X; SO(3)) \cong SO(3)^2/SO(3)$.
- **Torus.** $\mathcal{M}_{\text{flat}}(Y; SO(3)) \cong \{(A, B) \in SO(3)^2 : AB = BA\}/SO(3)$.

Implication for learning

The moduli space identifies the independent transport parameters. Fitting these directly could simplify the learning task.

1 Motivation and Background

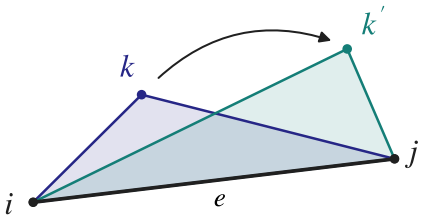
2 Construction

3 More Properties

4 Molecular generation

Can local triangle shape improve a generated conformation?

- We jointly predict atom types, charges, bonds, and 3D coordinates.
- A molecule can pass validity checks yet still release energy on relaxation.



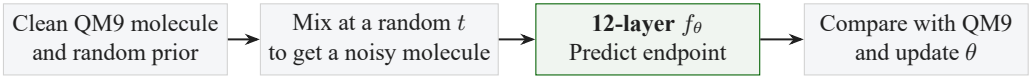
What a triangle adds to an edge

With i, j fixed, moving k to k' leaves $x_j - x_i$ unchanged but changes the triangle's shape.

Whitney Sheaf Attention Flow (WSA-Flow) uses this edge metric. We evaluate MMFF94 energy and strain per atom on QM9.

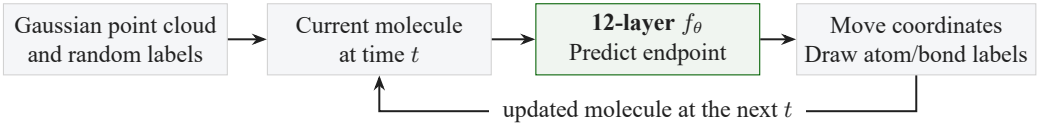
Flow matching for molecular generation with WSA-Flow

Training



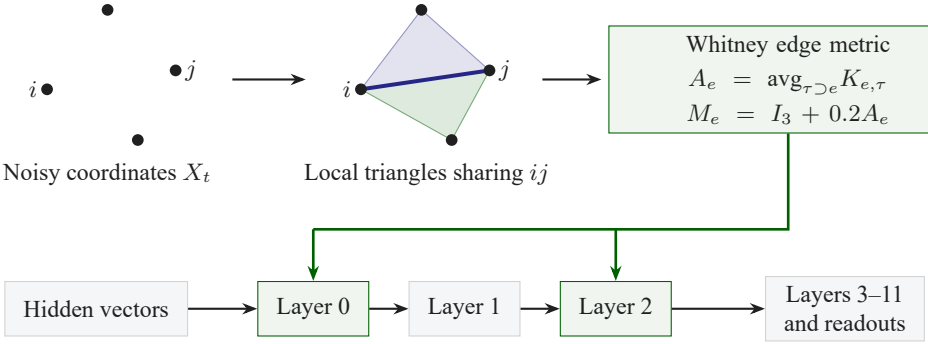
Fit predicted coordinates $\widehat{X}_1$ by MSE and atom, bond and charge labels by cross-entropy.

Sampling



Use $v_t = (\widehat{X}_1 - X_t)/(1-t)$ for coordinates. After 100 steps, return the last predicted molecule.

One forward pass through f_θ with Whitney metrics in layers 0 and 2



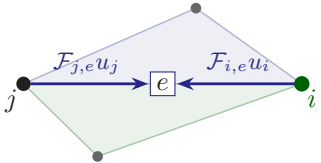
- At the molecular length scale ℓ, each triangle contributes $K_{e,\tau} = Q_{e|\tau}$.
- The two marked layers share M_e. Each learns its own gate to mix native and Whitney directions.

How the metric changes an attention direction

In one channel, center atom i's hidden vector $v_i \in \mathbb{R}^3$ over all n atoms to get $u_i = v_i - \frac{1}{n}\sum_{k=1}^{n} v_k$.

Local triangle and edge counts set $h_i > 0$ in the restriction $\mathcal{F}_{i,e} = h_i M_e^{1/2}$.

Compare on the edge



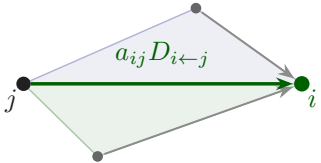
Difference in the edge stalk

$$\mathcal{F}_{i,e}u_i - \mathcal{F}_{j,e}u_j = M_e^{1/2}(h_i u_i - h_j u_j)$$

Map back to i

$$D^W_{i\leftarrow j} = \frac{\mathcal{F}_{i,e}^{\top}\left(\mathcal{F}_{i,e}u_i - \mathcal{F}_{j,e}u_j\right)}{\left\|\mathcal{F}_{i,e}u_i - \mathcal{F}_{j,e}u_j\right\|_2 + \varepsilon}$$

Weight messages at i



Edges outside triangles use the endpoint direction alone.

Mix endpoint and Whitney directions (layers 0 and 2)

$$D_{i\leftarrow j} = (1 - g_l(t))\,\frac{u_i - u_j}{\|u_i - u_j\|_2 + \varepsilon} + g_l(t)\,D^W_{i\leftarrow j}$$

Sum at i over its neighbors

$$m_i = \sum_j a_{ij} D_{i\leftarrow j}$$

What does MMFF94 relaxation measure?

For a small molecule with assigned force-field parameters, the Merck Molecular Force Field (MMFF94) defines an empirical potential $E(X)$ on its atomic positions.

$$X_{\text{gen}} \xrightarrow[\text{same atoms and bonds}]{\text{local MMFF94 minimization in RDKit}} X_{\text{rel}}.$$

Two energy measurements for the same molecule

$$\text{energy/atom} = \frac{E(X_{\text{gen}})}{N}, \qquad \text{strain/atom} = \frac{E(X_{\text{gen}}) - E(X_{\text{rel}})}{N}.$$

- Smaller strain means less energy released from that starting geometry.
- Different local minima can both give zero energy drop.

Halgren, MMFF94 RDKit MMFF optimization

QM9 molecular generation across five models

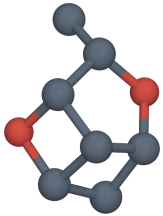
Mean ± population SD across three batches of 10,000 requests per model. **First** Second

Model	Atom stability (%) ↑	Molecule stability (%) ↑	Validity (%) ↑	Uniqueness (%) ↑	Energy/atom ↓	Strain/atom ↓
MiDi	99.4±0.0	90.2±0.1	88.0±0.3	**99.0**±0.0	2.877±0.007	1.814±0.013
EQGAT	99.7±0.0	95.6±0.2	95.0±0.2	98.0±0.1	4.108±0.021	2.955±0.025
FlowMol2	100.0±0.0	99.3±0.1	99.1±0.0	96.2±0.1	2.322±0.016	0.986±0.004
SemlaFlow	100.0±0.0	99.5±0.0	**99.2**±0.1	95.9±0.3	2.349±0.028	1.031±0.009
WSA-Flow	**100.0**±0.0	**99.5**±0.0	99.1±0.0	95.9±0.0	**2.295**±0.014	**0.984**±0.006

- Ranks use the means before rounding.
- MMFF94 energy and strain (kcal mol^{-1} atom^{-1}) use each model's computable molecules.

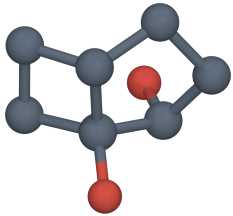
Three conformations generated by WSA-Flow

C02 $C_7H_{10}O_2$



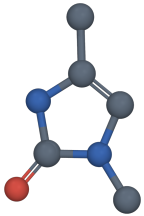
Energy 1.887 Strain 0.312

C05 $C_7H_{12}O_2$



Energy 2.100 Strain 0.200

C06 $C_5H_8N_2O$



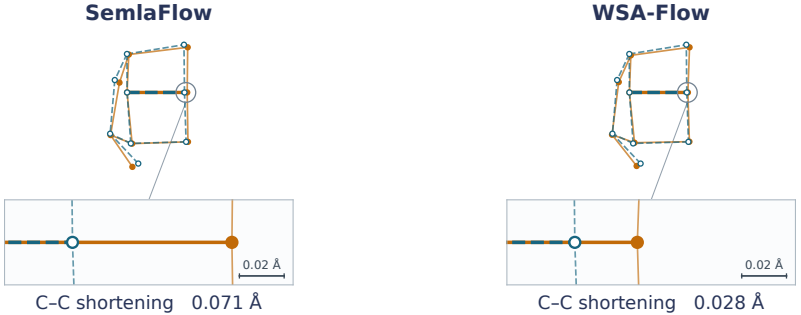
Energy -0.574 Strain 0.452

MMFF94 energy and strain per atom (kcal mol^{-1} atom^{-1}) · Heavy atoms shown

- Carbon is gray, oxygen red, and nitrogen blue. Scores include hydrogens.
- C05 releases the least energy per atom of the three conformations.

How far do the bonds move during relaxation?

C02 ($C_7H_{10}O_2$) ● Generated (solid) ○ After MMFF94 relaxation (dashed)



The insets share a scale and track the bond with the largest $|\Delta r|$ in the SemlaFlow conformer. Across all 11 heavy-atom bonds, mean $|\Delta r|$ is 0.025 Å (SemlaFlow) and 0.016 Å (WSA-Flow).

Thank you!