

Symmetry reduction from D_2 tetramer to C_2 dimer: group-theoretic and thermodynamic modeling of enzyme catalytic efficiency

Tianyi Qiu

Department of Mathematics, Bethany Lutheran College, Mankato, USA

vqiu@blc.edu

Abstract. A mathematical framework is presented to quantify the relationship between quaternary-structure symmetry, free energy, and catalytic efficiency during the transition from a D_2 -symmetric tetramer to a C_2 -symmetric dimer, exemplified with LDHA. The approach constructs explicit D_2 representations on subunit and interface feature spaces, derives projection operators to decompose operators and data into irreducible-representation components, and computes symmetry-resolved free-energy differences via Gaussian/statistical and harmonic/Hessian methods. Connections to kinetics are made through transition state theory with channel degeneracy. Reproducible algorithms and a workflow for mapping FoldX outputs into irrep-resolved diagnostics and efficiency predictions are provided.

Keywords: symmetry reduction, group representation, enzyme catalytic efficiency, free energy, LDHA

1. Introduction

Enzyme quaternary structure often exhibits symmetry, and this structural organization can significantly affect catalytic efficiency. Understanding how symmetry reduction impacts enzymatic function is a fundamental question in structural biology and mathematical modeling. A natural mathematical framework to address this problem is group theory, which formalizes symmetry and its action on suitable data spaces.

Not all enzymes are suitable for controlled structural perturbations. To study symmetry-dependent effects concretely, tetramer-to-dimer transitions provide a classical example of symmetry reduction, where experimental and structural data are available. In this work, lactate dehydrogenase A (LDHA) is selected as the model system because it forms a D_2 -symmetric tetramer, has documented mutational studies in the literature, and structural data are available from the Protein Data Bank.

The research proceeds in the following steps:

First, to define group actions on data spaces. Structural coordinates such as collective variables, interface energies, or density fields are formalized with an inner product structure, allowing symmetry operations to act linearly and quantitatively on the space.

Second, to construct explicit representations and projectors. Irreducible representations (irreps) of D_2 (tetramer) and C_2 (dimer) are constructed, and Reynolds projection operators are derived to decompose data and operators into irrep components, isolating the contributions of each symmetry mode.

Third, to compute symmetry-resolved free-energy differences. Gaussian/statistical (covariance-based) and harmonic/Hessian (normal-mode-based) approximations quantify the free-energy change associated with symmetry reduction, with explicit contributions from each irrep.

Fourth, to bridge to catalytic efficiency. Transition state theory (TST) connects free-energy differences to predicted catalytic efficiencies, accounting for channel degeneracy and transition-state probabilities.

Fifth, to implement a computational workflow. Energy terms and structural features from FoldX outputs are mapped into the symmetry-adapted framework, and diagnostic scores identify which symmetry components dominate the observed transition.

This approach results in a set of explicit, symmetry-resolved expressions for projected covariance matrices and free-energy differences. These formulas represent a novel mathematical formalism linking group-theoretic symmetry, structural transformation, and catalytic efficiency. The framework allows prediction of which symmetry modes most strongly drive the tetramer-to-dimer transition and provides a direct quantitative bridge from structural symmetry to enzymatic function.

2. Symmetry groups, homomorphism, and action spaces

2.1. Groups and homomorphism

Consider tetramers a finite group:

$$D_2 = \{e, \alpha, \beta, \alpha\beta\}, \alpha^2 = \beta^2 = e, \alpha\beta = \beta\alpha \quad (1)$$

and consider dimers a finite group:

$$C_2 = \{e, \gamma\}, \gamma^2 = e \quad (2)$$

By construction D_2 is abelian of order 4 (isomorphic to the Klein four group V_4), and C_2 is the cyclic group of order 2.

2.1.1. Group homomorphism

A map $\phi: G \rightarrow H$ between groups is a homomorphism iff $\phi(g_1 g_2) = \phi(g_1) \phi(g_2)$ for all $g_1, g_2 \in G$. Define

$$\phi: D_2 \rightarrow C_2, \phi(\alpha) = \gamma, \phi(\beta) = e \quad (3)$$

and extend multiplicatively to all of D_2 . Concretely,

$$\phi(e) = e, \phi(\alpha) = \gamma, \phi(\beta) = e, \phi(\alpha\beta) = \gamma \quad (4)$$

ϕ as defined above is a surjective group homomorphism.

Proof. Since D_2 is generated by α, β with relations $\alpha^2 = \beta^2 = e$ and $\alpha\beta = \beta\alpha$, it suffices to check that the relations are preserved:

$$\phi(\alpha^2) = \phi(\alpha)^2 = \gamma^2 = e = \phi(e), \phi(\beta^2) = \phi(\beta)^2 = e \quad (5)$$

and

$$\phi(\alpha\beta) = \phi(\alpha)\phi(\beta) = \gamma \cdot e = \gamma = \phi(\beta)\phi(\alpha) = \phi(\beta\alpha) \quad (6)$$

Hence ϕ respects the defining relations and extends to a homomorphism on D_2 . Its image contains γ and e , hence $\text{Im } \phi = C_2$, so ϕ is surjective.

2.1.2. Kernel and quotient

Recall $\ker \phi = \{g \in D_2 : \phi(g) = e\}$. Here

$$\ker \phi = e, \beta \quad (7)$$

which is a subgroup of order 2. Since ϕ is a homomorphism, $\ker \phi$ is a normal subgroup. By the First Isomorphism Theorem,

$$D_2/\ker\phi \cong \text{Im } \phi = C_2 \quad (8)$$

The quotient map $\pi: D_2 \rightarrow D_2/\ker\phi$ followed by the isomorphism from $D_2/\ker\phi$ to C_2 equals ϕ .

2.1.3. Functions factoring through the quotient

If X is any set on which D_2 acts and $\mathcal{H}$ is a function space on X , then the subspace of functions invariant under $\ker\phi$ [1],

$$\mathcal{H}^{\ker\phi} := \{f \in \mathcal{H} : U(k)f = f, \forall k \in \ker\phi\} \quad (9)$$

is canonically identified with functions on the quotient $X/\ker\phi$ or, algebraically, with functions on $D_2/\ker\phi$: every $f \in \mathcal{H}^{\ker\phi}$ is constant on $\ker\phi$ -cosets and thus descends to a function on the quotient; conversely, a function on the quotient pulls back to a $\ker\phi$ -invariant function.

2.2. Representations, invariant projectors and factoring through quotients

2.2.1. Representations and unitarity (definitions)

Let $(H, \langle \cdot, \cdot \rangle)$ be a (complex) Hilbert space. A representation of a group G on $\mathcal{H}$ is a group homomorphism $U: G \rightarrow GL(H)$ satisfying $U(g_1 g_2) = U(g_1)U(g_2)$. The representation U is unitary (or orthogonal in the real case) if [2]

$$\langle U(g)f, U(g)h \rangle = \langle f, h \rangle \quad \forall f, h \in H, g \in G \quad (10)$$

2.2.2. Averaging projector onto $\ker\phi$ -invariants

Let $K := \ker\phi$ be a finite subgroup. Define the averaging operator

$$P_K := \frac{1}{|K|} \sum_{k \in K} U(k) \quad (11)$$

Assume U is unitary. Then P_K is the orthogonal projection onto the closed subspace $\mathcal{H}^K := \{v \in H : U(k)v = v \quad \forall k \in K\}$. Moreover P_K commutes with $U(g)$ for every $g \in G$, hence the representation leaves $\mathcal{H}^K$ invariant.

Proof. First compute P_K^2 :

$$P_K^2 = \frac{1}{|K|^2} \sum_{k, k' \in K} U(k)U(k') = \frac{1}{|K|^2} \sum_{h \in K} \left(\sum_{k \in K} 1 \right) U(h) = \frac{1}{|K|} \sum_{h \in K} U(h) = P_K \quad (12)$$

so P_K is idempotent. Since each $U(k)$ is unitary, $U(k)^* = U(k)^{-1} = U(k^{-1})$, and taking adjoint yields $P_K^* = P_K$, hence P_K is self-adjoint. Idempotent self-adjoint operators are orthogonal projections; their range equals $\{v : P_K v = v\}$. But $P_K v = v$ iff $\frac{1}{|K|} \sum_{k \in K} U(k)v = v$, which is equivalent to $U(k)v = v$ for all $k \in K$ (apply $U(k_0)$ and use group averaging). Thus the range is $\mathcal{H}^K$.

Finally, for any $g \in G$,

$$U(g)P_K U(g)^{-1} = \frac{1}{|K|} \sum_{k \in K} U(g)U(k)U(g)^{-1} = \frac{1}{|K|} \sum_{k \in K} U(gkg^{-1}) \quad (13)$$

If K is normal then $gKg^{-1}=K$ and the right-hand side equals P_K . In our setting $K=\ker\phi$ is normal, so P_K commutes with $U(g)$. Hence $\mathcal{H}^K$ is G -invariant.

2.2.3. Factoring a representation through the quotient

Let $K=\ker\phi$ and suppose $U:G\rightarrow U(H)$ is a unitary representation. Consider the restriction of U to $\mathcal{H}^K$. For $g, g'\in G$ with $gK=g'K$, we have $g'=gk$ for some $k\in K$ and, for $v\in\mathcal{H}^K$,

$$U(g')v=U(g)U(k)v=U(g)v \quad (14)$$

Thus the map $\tilde{U}:G/K\rightarrow U(\mathcal{H}^K)$ given by $\tilde{U}(gK):=U(g)|_{\mathcal{H}^K}$ is well-defined and a representation. In other words, the representation on the K -fixed subspace factors through the quotient group $G/K\cong C_2$.

2.3. Action spaces: L^2 constructions and unitarity proofs

2.3.1. Why L^2 and which inner product

Let $\mathcal{H}=L^2(\mathbb{R}^3)$ (complex-valued functions) with the usual inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^3} \overline{f(r)} g(r) dr \quad (15)$$

L^2 is a Hilbert space: it is complete and admits orthogonal projections, spectral theorem for bounded self-adjoint operators, and a well-behaved theory of quadratic forms. These properties are prerequisites for using orthogonal (projector) decompositions, logdet formulas for Gaussian integrals, and spectral block-diagonalization of self-adjoint Hessians [3].

2.3.2. Density-field representation on $L^2(\mathbb{R}^3)$

Definition.

Let G act on $\mathbb{R}^3$ by orthogonal linear maps $R(g)\in O(3)$ (or by rigid motions with zero translations for point groups). Define [2]

$$(U(g)\rho)(r):=\rho(R(g)^{-1}r) \quad (16)$$

We check that U is a unitary representation.

$U:G\rightarrow U(L^2(\mathbb{R}^3))$ defined above is a unitary representation.

Proof. (Representation property.) For $g_1, g_2\in G$,

$$U(g_1)U(g_2)\rho(r)=U(g_1)\left(\rho(R(g_2)^{-1}r)\right)=\rho\left(R(g_1)^{-1}R(g_2)^{-1}r\right)=\rho\left((R(g_2)R(g_1))^{-1}r\right)=U(g_1g_2)\rho(r) \quad (17)$$

(Unitarity.) For $\rho_1, \rho_2\in L^2(\mathbb{R}^3)$,

$$\langle U(g)\rho_1, U(g)\rho_2 \rangle = \int_{\mathbb{R}^3} \overline{\rho_1(R(g)^{-1}r)} \rho_2(R(g)^{-1}r) dr \quad (18)$$

Change variables $s:=R(g)^{-1}r$. Since $R(g)$ is orthogonal, $\det R(g)=\pm 1$ and $dr=|\det R(g)| ds=ds$. Thus

$$\langle U(g)\rho_1, U(g)\rho_2 \rangle = \int_{\mathbb{R}^3} \overline{\rho_1(s)} \rho_2(s) ds = \langle \rho_1, \rho_2 \rangle \quad (19)$$

Therefore each $U(g)$ is unitary and U is a unitary representation.

2.3.3. CV (collective-variable) representation on $L^2(Y, \mu)$

Setup and unitarity criterion.

Let Φ be a map from full coordinates to CVs, $x \mapsto y = \Phi(x) \in Y \subset \mathbb{R}^m$, and let μ be a probability measure on Y . Consider $\mathcal{H} := L^2(Y, \mu)$ with inner product $\langle A, B \rangle = \int_Y \overline{A(y)} B(y) d\mu(y)$. Assume the group G acts on Y linearly or by permutations via $\Gamma: G \rightarrow GL(Y)$, and define

$$(U(g)A)(y) := A(\Gamma(g)^{-1}y) \quad (20)$$

U is unitary on $L^2(Y, \mu)$ iff μ is Γ -invariant, i.e. $\mu = \mu \circ \Gamma(g)^{-1}$ for all $g \in G$.

Proof. If μ is Γ -invariant, then

$$\langle U(g)A, U(g)B \rangle = \int_Y \overline{A(\Gamma(g)^{-1}y)} B(\Gamma(g)^{-1}y) d\mu(y) \quad (21)$$

Change variables $z = \Gamma(g)^{-1}y$; by invariance $d\mu(y) = d\mu(z)$ and the domain is unchanged, so the integral equals $\langle A, B \rangle$. Conversely, unitarity for all A, B forces invariance of μ by testing against indicator functions.

2.4. Intertwining operators and boundedness

2.4.1. A linear sampling operator

Suppose we construct a linear sampling operator $S: L^2(\mathbb{R}^3) \rightarrow L^2(Y, \mu)$ of the integral-kernel type

$$(S\rho)(y) = \int_{\mathbb{R}^3} K(y, r) \rho(r) dr \quad (22)$$

where $K: Y \times \mathbb{R}^3 \rightarrow \mathbb{C}$ is measurable. Typical choices of K are localized feature kernels (e.g. Gaussian windows integrated against coordinates).

If for μ -almost every y the function $r \mapsto K(y, r)$ lies in $L^2(\mathbb{R}^3)$ and the function $y \mapsto \|K(y, \cdot)\|_{L^2(\mathbb{R}^3)}^2$ is integrable w.r.t. μ , then S is a bounded linear operator $L^2(\mathbb{R}^3) \rightarrow L^2(Y, \mu)$.

Proof. By Cauchy–Schwarz,

$$|(S\rho)(y)|^2 = \left| \int K(y, r) \rho(r) dr \right|^2 \leq \|K(y, \cdot)\|_{L^2(\mathbb{R}^3)}^2 \|\rho\|_{L^2(\mathbb{R}^3)}^2 \quad (23)$$

Integrate both sides w.r.t. y and use the integrability assumption to obtain $\|S\rho\|_{L^2(Y, \mu)}^2 \leq C \|\rho\|_{L^2(\mathbb{R}^3)}^2$ with $C = \int_Y \|K(y, \cdot)\|_{L^2}^2 d\mu(y) < \infty$. Thus S is bounded.

2.4.2. Equivariance (intertwining) condition

An equivariance condition

$$S \circ U(g) = \Gamma(g) \circ S \quad \forall g \in G \quad (24)$$

is equivalent to a symmetry constraint on the kernel K :

$$K(\Gamma(g)^{-1}y, r) = K(y, R(g)r) \quad (\text{for a.e. } y, r) \quad (25)$$

Indeed,

$$(S(U(g)\rho))(y) = \int K(y, r) \rho(R(g)^{-1}r) \, dr = \int K(y, R(g)s) \rho(s) \, ds \quad (26)$$

while

$$(\Gamma(g)S\rho)(y) = (S\rho)(\Gamma(g)^{-1}y) = \int K(\Gamma(g)^{-1}y, s) \rho(s) \, ds \quad (27)$$

Equality for all ρ forces the kernel relation above. Under that relation, S is an intertwiner and thus maps symmetry-adapted subspaces into symmetry-adapted subspaces.

2.5. Decomposition into irreducibles and projector formulas

2.5.1. Complete reducibility for finite group

For a finite group G and a unitary representation U on a complex Hilbert space $\mathcal{H}$, Maschke's theorem implies $\mathcal{H}$ decomposes as a (Hilbert) direct sum of finite-dimensional G -invariant subspaces each of which splits into irreducible representations (isotypic decomposition) [2]:

$$\mathcal{H} \cong \bigoplus_{\lambda \in \hat{G}} \mathcal{H}_\lambda \quad (28)$$

where $\hat{G}$ denotes the set of (equivalence classes of) irreducible representations and $\mathcal{H}_\lambda$ is the λ -isotypic component.

2.5.2. Character projector (explicit orthogonal projector)

Let χ be the character of an irreducible unitary representation V_χ of dimension d_χ . Then the orthogonal projector onto the χ -isotypic component of $\mathcal{H}$ is

$$P_\chi = \frac{d_\chi}{|G|} \sum_{g \in G} \overline{\chi(g)} U(g) \quad (29)$$

One may verify $P_\chi^2 = P_\chi$, $P_\chi^* = P_\chi$, and $P_\chi U(h) = U(h) P_\chi$ for all $h \in G$; hence P_χ projects orthogonally onto an invariant subspace isomorphic to a direct sum of copies of V_χ .

3. Irreducible representations and projectors

3.1. Irreps and character table of D_2

3.1.1. Why all irreps are one-dimensional

The group D_2 considered here is the dihedral group of order 4, which is isomorphic to the direct product

$$D_2 \cong C_2 \times C_2 = \{ e, \alpha, \beta, \alpha\beta \}, \alpha^2 = \beta^2 = e, \alpha\beta = \beta\alpha \quad (30)$$

Since D_2 is abelian, each element commutes with every other. A standard theorem in representation theory states:

> For a finite abelian group G , every irreducible representation over $\mathbb{C}$ is one-dimensional.

The intuition is that the group algebra $\mathbb{C}[G]$ of an abelian group splits as a direct sum of one-dimensional eigenspaces, and thus irreps can only be 1D characters.

Therefore, all irreps of D_2 are homomorphisms

$$\chi: D_2 \rightarrow \{+1, -1\} \quad (31)$$

3.1.2. Conjugacy classes and the number of irreps

For any group, the number of irreducible representations equals the number of conjugacy classes. Since D_2 is abelian, each element forms a conjugacy class by itself:

$$e, \alpha, \beta, \alpha\beta \quad (32)$$

Thus D_2 has 4 inequivalent irreps in total.

3.1.3. Explicit construction of irreps

Each one-dimensional character χ is determined by its values on the generators α, β . There are 2 choices for $\chi(\alpha)$ (+1 or -1), and 2 choices for $\chi(\beta)$, hence $2 \times 2 = 4$ distinct irreps in total.

Explicitly:

$$(\chi(\alpha), \chi(\beta)) \in (1, 1), (1, -1), (-1, 1), (-1, -1) \quad (33)$$

The corresponding value on $\alpha\beta$ is then determined multiplicatively:

$$\chi(\alpha\beta) = \chi(\alpha)\chi(\beta) \quad (34)$$

3.1.4. The character table

It is conventional to label these four irreps as A_1, B_1, B_2, B_3 , giving the table [2]:

Table 1. Four irrep ρ

Irrep ρ	$\chi_\rho(e)$	$\chi_\rho(\alpha)$	$\chi_\rho(\beta)$	$\chi_\rho(\alpha\beta)$
A_1	1	1	1	1
B_1	1	1	-1	-1
B_2	1	-1	1	-1
B_3	1	-1	-1	1

3.1.5. Orthogonality sanity check

Characters of irreps are orthogonal with respect to the inner product

$$\langle \chi_\rho, \chi_\sigma \rangle = \frac{1}{|D_2|} \sum_{g \in D_2} \overline{\chi_\rho(g)} \chi_\sigma(g) \quad (35)$$

Since all values are real (± 1), orthogonality reduces to checking that each pair of rows in the table has dot product zero in $\mathbb{R}^4$, which is immediate. Thus the table is consistent.

3.2. Reynolds/character projectors: derivation and properties

3.2.1. General formula

Let $U: D_2 \rightarrow GL(V)$ be a unitary representation with character $\chi_U(g) = \text{Tr } U(g)$. For each irrep ρ with character χ_ρ , the associated projector is [2]

$$P_\rho = \frac{\text{dimp}}{|D_2|} \sum_{g \in D_2} \overline{\chi_\rho(g)} U(g) \quad (36)$$

3.2.2. Simplification for D_2

Here $\text{dimp}=1$ and $\chi_\rho(g) \in \{\pm 1\}$ is real, so

$$\begin{aligned} P_{A_1} &= 1/4 (I + U(\alpha) + U(\beta) + U(\alpha\beta)), \\ P_{B_1} &= 1/4 (I + U(\alpha) - U(\beta) - U(\alpha\beta)), \\ P_{B_2} &= 1/4 (I - U(\alpha) + U(\beta) - U(\alpha\beta)), \\ P_{B_3} &= 1/4 (I - U(\alpha) - U(\beta) + U(\alpha\beta)). \end{aligned} \quad (37)$$

Each projector is simply a linear combination of the four representation matrices, with coefficients $\pm 1/4$.

3.2.3. Why these are projectors

We check the key properties:

Idempotence: $P_\rho^2 = P_\rho$. This follows from Schur orthogonality of characters:

$$\frac{1}{|D_2|} \sum_{g \in D_2} \chi_\rho(g^{-1}) \chi_\sigma(hg) = \frac{\delta_{\rho\sigma}}{\dim \rho} \chi_\rho(h) \quad (38)$$

A direct substitution shows $P_\rho^2 = P_\rho$.

Mutual orthogonality: $P_\rho P_\sigma = 0$ if $\rho \neq \sigma$. This again comes from the orthogonality of characters.

Resolution of identity:

$$\sum_\rho P_\rho = I \quad (39)$$

Thus the representation space splits as a direct sum of the four invariant subspaces.

Commutation: $U(h)P_\rho = P_\rho U(h)$ for all $h \in D_2$. Hence each P_ρ is an intertwiner and respects the group symmetry.

3.2.4. Self-adjointness

If U is unitary, then

$$P_\rho^* = \frac{1}{|D_2|} \sum_g \chi_\rho(g) U(g)^* = \frac{1}{|D_2|} \sum_g \chi_\rho(g) U(g^{-1}) \quad (40)$$

Since $\chi_\rho(g^{-1}) = \overline{\chi_\rho(g)}$ for one-dimensional real characters, this equals P_ρ . Hence P_ρ is Hermitian and therefore an orthogonal projector.

3.2.5. Multiplicity formula

The multiplicity m_ρ of ρ in U is

$$m_\rho = \frac{1}{|D_2|} \sum_{g \in D_2} \overline{\chi_\rho(g)} \chi_U(g) \quad (41)$$

Equivalently, $\text{rank}(P_\rho) = m_\rho$. This gives a direct computational method to decompose U .

3.3. Concrete D_2 representations in LDHA: two working examples

3.3.1. Example A: chain representation on $\mathbb{R}^4$

Group action. Label four chains (A, B, C, D) as standard basis vectors. Define: $\alpha: A \leftrightarrow B, C \leftrightarrow D, \beta: A \leftrightarrow C, B \leftrightarrow D$ so that $\alpha\beta$ swaps $A \leftrightarrow D$ and $B \leftrightarrow C$. The $U(g)$ are 4×4 permutation matrices.

Character computation. The trace of each permutation matrix equals the number of fixed points: $\chi_U(e) = 4, \chi_U(\alpha) = \chi_U(\beta) = \chi_U(\alpha\beta) = 0$.

Decomposition. Using multiplicities,

$$m_\rho = 1/4 \sum_g \chi_\rho(g) \chi_U(g) \quad (42)$$

we obtain

$$m_{A_1} = m_{B_1} = m_{B_2} = m_{B_3} = 1 \quad (43)$$

Hence

$$\mathbf{R}^4 \cong \mathbf{A}_1 \oplus \mathbf{B}_1 \oplus \mathbf{B}_2 \oplus \mathbf{B}_3 \quad (44)$$

Explicit basis. An orthogonal eigenbasis realizing the decomposition is

$$\begin{aligned} \mathbf{A}_1: & (1,1,1,1), & \mathbf{B}_1: & (1,1,-1,-1), \\ \mathbf{B}_2: & (1,-1,1,-1), & \mathbf{B}_3: & (1,-1,-1,1). \end{aligned} \quad (45)$$

Applying the projectors and extracts each component.

3.3.2. Example B: interface representation on $\mathbf{R}^6$

Group action. List six interfaces:

$$(\mathbf{AB}, \mathbf{AC}, \mathbf{AD}, \mathbf{BC}, \mathbf{BD}, \mathbf{CD}) \quad (46)$$

The action of $\alpha, \beta, \alpha\beta$ permutes these as:

$$\begin{aligned} \alpha: & \mathbf{AB} \mapsto \mathbf{AB}, \mathbf{AC} \leftrightarrow \mathbf{BD}, \mathbf{AD} \leftrightarrow \mathbf{BC}, \mathbf{CD} \mapsto \mathbf{CD}, \\ \beta: & \mathbf{AB} \leftrightarrow \mathbf{CD}, \mathbf{AC} \mapsto \mathbf{AC}, \mathbf{AD} \leftrightarrow \mathbf{BC}, \mathbf{BD} \mapsto \mathbf{BD}, \\ \alpha\beta: & \mathbf{AB} \leftrightarrow \mathbf{CD}, \mathbf{AC} \leftrightarrow \mathbf{BD}, \mathbf{AD} \mapsto \mathbf{AD}, \mathbf{BC} \mapsto \mathbf{BC}. \end{aligned} \quad (47)$$

Characters. Counting fixed points:

$$\chi_U(\mathbf{e})=6, \chi_U(\alpha)=\chi_U(\beta)=\chi_U(\alpha\beta)=2 \quad (48)$$

Decomposition. Multiplicities:

$$\begin{aligned} m_{\mathbf{A}_1} &= 1/4(6+2+2+2)=3, \\ m_{\mathbf{B}_1} &= 1/4(6+2-2-2)=1, \\ m_{\mathbf{B}_2} &= 1/4(6-2+2-2)=1, \\ m_{\mathbf{B}_3} &= 1/4(6-2-2+2)=1. \end{aligned} \quad (49)$$

So

$$\mathbf{R}^6 \cong 3\mathbf{A}_1 \oplus \mathbf{B}_1 \oplus \mathbf{B}_2 \oplus \mathbf{B}_3 \quad (50)$$

Orbit structure and basis. Interfaces split into three orbits:

$$\mathbf{AB}, \mathbf{CD}, \mathbf{AC}, \mathbf{BD}, \mathbf{AD}, \mathbf{BC} \quad (51)$$

This yields:

$$\begin{aligned}
A_1 \text{ basis: } s_1 &= AB+CD, \quad s_2=AC+BD, \quad s_3=AD+BC, \\
B_1 \text{ basis: } b_1 &= AB-CD, \\
B_2 \text{ basis: } b_2 &= AC-BD, \\
B_3 \text{ basis: } b_3 &= AD-BC.
\end{aligned} \tag{52}$$

Thus the space naturally decomposes into symmetric orbit sums (A_1) and antisymmetric differences (B_i).
 Projector action. For any $w \in \mathbb{R}^6$,

$$w^{(p)} = P_p w \tag{53}$$

where $w^{(A_1)}$ averages over each orbit and $w^{(B_i)}$ extracts the signed difference. This provides an explicit computational recipe.

3.4. Applications of the projectors

3.4.1. Block diagonalization of equivariant operators

If K is D_2 -equivariant ($U(g)K=KU(g)$), then

$$K = \sum_p P_p K P_p \simeq \bigoplus_p K_p, \quad K_p = P_p K P_p|_{\text{Im} P_p} \tag{54}$$

Thus K block-diagonalizes into irreducible sectors.

3.4.2. Numerical recipe

Given $U(\alpha), U(\beta)$, construct the four projectors. Then compute

$$w^{(p)} = P_p w, \quad C_p = P_p C P_p^\top, \quad K_p = P_p K P_p^\top \tag{55}$$

Subsequent log-determinant and free energy formulas decompose accordingly.

3.4.3. Stability

Because P_p are orthogonal projectors with rational coefficients, the decomposition is numerically robust: the projections are exact up to floating-point error. For D_2 , all coefficients are $\pm 1/4$, ensuring excellent conditioning.

4. Explicit representations on subunit and interface spaces

4.1. Subunit (4D) representation: construction, checks, and decomposition

4.1.1. Basis and ordering

Let $V_{\text{sub}} = \mathbb{R}^4$ with the standard basis corresponding to chains (A,B,C,D) :

$$e_A = (1,0,0,0)^\top, \quad e_B = (0,1,0,0)^\top, \quad e_C = (0,0,1,0)^\top, \quad e_D = (0,0,0,1)^\top \tag{56}$$

We represent a general vector as $v = (x_A, x_B, x_C, x_D)^\top$.

4.1.2. Group action by permuting chains

We define the D_2 action by permuting chain labels (orthogonal maps in the Euclidean inner product):

$$\begin{aligned}\alpha: A \leftrightarrow B, C \leftrightarrow D, \\ \beta: A \leftrightarrow D, B \leftrightarrow C, \\ \alpha\beta: A \leftrightarrow C, B \leftrightarrow D.\end{aligned}\tag{57}$$

Each action is a permutation of the basis and is therefore represented by a permutation matrix $U(g)$ satisfying $U(g)^T U(g) = I$.

4.1.3. Building the matrices entry-by-entry

A permutation matrix $U(g)$ is determined by where g sends each basis vector. For example, α swaps $A \leftrightarrow B$ and $C \leftrightarrow D$, hence

$$U(\alpha)e_A = e_B, U(\alpha)e_B = e_A, U(\alpha)e_C = e_D, U(\alpha)e_D = e_C\tag{58}$$

so the columns of $U(\alpha)$ are the images of the basis vectors:

$$U(\alpha) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\tag{59}$$

Proceeding identically for β and $\alpha\beta$ gives

$$U(\beta) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, U(\alpha\beta) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\tag{60}$$

4.1.4. Representation sanity checks

We verify the group relations on matrices:

$$U(\alpha)^2 = I, U(\beta)^2 = I, U(\alpha)U(\beta) = U(\beta)U(\alpha) = U(\alpha\beta)\tag{61}$$

These hold because applying the swaps twice returns each chain to itself, and α, β commute on labels, so their permutation matrices commute.

4.1.5. Characters and quick multiplicity count

The character of a permutation is the number of fixed basis vectors (its trace). For the identity, $\chi_U(e) = \text{Tr } I = 4$. For each non-identity above, no chain is fixed, so $\chi_U(\alpha) = \chi_U(\beta) = \chi_U(\alpha\beta) = 0$. Using the multiplicity formula

$$m_p = \frac{1}{|D_2|} \sum_{g \in D_2} \chi_p(g) \chi_U(g), |D_2|=4 \quad (62)$$

and the D_2 character table, one obtains $m_{A_1}=m_{B_1}=m_{B_2}=m_{B_3}=1$. Therefore

$$\boxed{V_{\text{sub}} \cong A_1 \oplus B_1 \oplus B_2 \oplus B_3.} \quad (63)$$

4.1.6. An explicit irrep eigenbasis (and why it works)

Define the four vectors

$$u_{A_1} = 1/2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, u_{B_1} = 1/2 \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, u_{B_2} = 1/2 \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, u_{B_3} = 1/2 \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \quad (64)$$

Each u_p is an eigenvector of every $U(g)$ with eigenvalue $\chi_p(g)$ (the one-dimensional irrep property). For instance,

$$U(\alpha) u_{B_2} = 1/2 \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} = 1/2 \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = -u_{B_2} \quad (65)$$

while $U(\beta) u_{B_2} = +u_{B_2}$ and $U(\alpha\beta) u_{B_2} = -u_{B_2}$, matching the character row $[1, -1, 1, -1]$ of B_2 . Orthogonality is immediate from the sign patterns, and the $1/2$ factor normalizes the vectors to unit length.

4.1.7. Projectors (closed form matrices)

Because D_2 is abelian with real characters,

$$\begin{aligned} P_{A_1} &= 1/4 (I + U(\alpha) + U(\beta) + U(\alpha\beta)) = 1/4 \mathbf{1} \mathbf{1}^\top, \\ P_{B_1} &= 1/4 (I + U(\alpha) - U(\beta) - U(\alpha\beta)), \\ P_{B_2} &= 1/4 (I - U(\alpha) + U(\beta) - U(\alpha\beta)), \\ P_{B_3} &= 1/4 (I - U(\alpha) - U(\beta) + U(\alpha\beta)), \end{aligned} \quad (66)$$

where $\mathbf{1} = (1, 1, 1, 1)^\top$. Explicitly,

$$\begin{aligned} P_{A_1} &= \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}, P_{B_1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}, \\ P_{B_2} &= \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}, P_{B_3} = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}. \end{aligned} \quad (67)$$

For any $v \in \mathbb{R}^4$, the four components $v^{(p)} = P_p v$ lie on the irrep lines and sum to v .

4.2. Interface (6D) representation: construction, checks, and decomposition

4.2.1. Basis and ordering

Let $V_{\text{int}} = \mathbb{R}^6$ with the basis listing the unordered interfaces in the fixed order

$$(AB, BC, CD, DA, AC, BD) \quad (68)$$

Denote a vector as

$$w = (w_{AB}, w_{BC}, w_{CD}, w_{DA}, w_{AC}, w_{BD})^T \quad (69)$$

4.2.2. How to build each permutation matrix

Given a group element g , map each unordered edge (e.g. AB) by acting on its endpoints (e.g. $A \mapsto g(A)$, $B \mapsto g(B)$), then relabel the resulting unordered edge in our fixed order. Place a 1 in row “image index” and column “original index”. Doing this for all six edges yields the 6×6 permutation matrix $U(g)$.

4.2.3. Explicit edge mappings

Using your subunit actions,

$$\alpha: A \leftrightarrow B, C \leftrightarrow D, \beta: A \leftrightarrow D, B \leftrightarrow C, \alpha\beta: A \leftrightarrow C, B \leftrightarrow D \quad (70)$$

we obtain the following maps (written as “original edge $\mapsto$ image edge”):

Table 2. Edge mappings

Edge	AB	BC	CD	DA	AC	BD
α	AB	DA	CD	BC	BD	AC
β	CD	BC	AB	DA	BD	AC
$\alpha\beta$	CD	DA	AB	BC	AC	BD

Reading column-by-column (image of basis vectors) gives

$$U(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, U(\beta) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, U(\alpha\beta) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (71)$$

4.2.4. Representation checks

Each $U(g)$ is orthogonal ($U(g)^T U(g) = I$) since it permutes coordinates. Moreover,

$$U(\alpha)^2 = U(\beta)^2 = I, U(\alpha)U(\beta) = U(\beta)U(\alpha) = U(\alpha\beta) \quad (72)$$

so we indeed have a (unitary) representation of D_2 .

4.2.5. Characters (why the traces are 6,2,2,2)

- e fixes all 6 edges: $\chi_U(e) = 6$.
- α fixes AB and CD, swaps $BC \leftrightarrow DA$ and $AC \leftrightarrow BD$: $\chi_U(\alpha) = 2$.
- β fixes BC and DA, swaps $AB \leftrightarrow CD$ and $AC \leftrightarrow BD$: $\chi_U(\beta) = 2$.
- $\alpha\beta$ fixes AC and BD, swaps $AB \leftrightarrow CD$ and $BC \leftrightarrow DA$: $\chi_U(\alpha\beta) = 2$.

4.2.6. Multiplicity calculation spelled out

With the D_2 character table $\{\chi_{A_1}, \chi_{B_1}, \chi_{B_2}, \chi_{B_3}\}$ and the above χ_U ,

$$m_p = \frac{1}{4} (\chi_p(e) \cdot 6 + \chi_p(\alpha) \cdot 2 + \chi_p(\beta) \cdot 2 + \chi_p(\alpha\beta) \cdot 2) \quad (73)$$

so

$$m_{A_1} = 3, m_{B_1} = 1, m_{B_2} = 1, m_{B_3} = 1 \quad (74)$$

Hence

$$\boxed{V_{\text{int}} \cong 3A_1 \oplus B_1 \oplus B_2 \oplus B_3.} \quad (75)$$

4.2.7. Orbit structure and immediate block basis

The six edges split into three D_2 -orbits of size 2:

$$AB, CD, BC, DA, AC, BD \quad (76)$$

For each orbit, form the symmetric (sum) and antisymmetric (difference) combinations:

$$\begin{aligned} s_1 &= AB + CD, b_1 = AB - CD, \\ s_2 &= BC + DA, b_2 = BC - DA, \\ s_3 &= AC + BD, b_3 = AC - BD. \end{aligned} \quad (77)$$

Then s_1, s_2, s_3 span the $3A_1$ subspace (fixed by all $U(g)$), and b_1, b_2, b_3 are one-dimensional eigenlines carrying B_1, B_2, B_3 respectively. Indeed,

$$\begin{array}{cccc}
& e & \alpha & \beta & \alpha\beta \\
b_1: & 1 & 1 & -1 & -1 \text{ (row } B_1) \\
b_2: & 1 & -1 & 1 & -1 \text{ (row } B_2) \\
b_3: & 1 & -1 & -1 & 1 \text{ (row } B_3)
\end{array} \tag{78}$$

because α fixes AB,CD while β and $\alpha\beta$ swap them, etc.

4.2.8. Projectors as explicit 6×6 matrices

Using

$$P_\rho = 1/4 (I + \chi_\rho(\alpha)U(\alpha) + \chi_\rho(\beta)U(\beta) + \chi_\rho(\alpha\beta)U(\alpha\beta)) \tag{79}$$

we obtain (block-diagonal over the three orbits):

$$P_{A_1} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, P_{B_1} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{80}$$

$$P_{B_2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, P_{B_3} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \tag{81}$$

These satisfy $P_\rho^2 = P_\rho$, $P_\rho P_\sigma = \delta_{\rho\sigma} P_\rho$, and $\sum_\rho P_\rho = I$.

4.2.9. Worked projection formulas

For $w = (w_{AB}, w_{BC}, w_{CD}, w_{DA}, w_{AC}, w_{BD})^\top$,

$$\begin{aligned}
P_{A_1} w &= (w_{AB} + w_{CD}/2, w_{BC} + w_{DA}/2, w_{AB} + w_{CD}/2, w_{BC} + w_{DA}/2, w_{AC} + w_{BD}/2, w_{AC} + w_{BD}/2)^\top, \\
P_{B_1} w &= (w_{AB} - w_{CD}/2, 0, -w_{AB} - w_{CD}/2, 0, 0, 0)^\top, \\
P_{B_2} w &= (0, w_{BC} - w_{DA}/2, 0, -w_{BC} - w_{DA}/2, 0, 0)^\top, \\
P_{B_3} w &= (0, 0, 0, 0, w_{AC} - w_{BD}/2, -w_{AC} - w_{BD}/2)^\top.
\end{aligned} \tag{82}$$

Thus $w = w^{(A_1)} + w^{(B_1)} + w^{(B_2)} + w^{(B_3)}$ with each $w^{(\rho)}$ living in the claimed irrep block.

4.2.10. Why orthogonality holds

Since each P_ρ is a polynomial in the commuting orthogonal matrices $U(g)$ with real coefficients, we have $P_\rho^\top = P_\rho$ and $P_\rho P_\sigma = \delta_{\rho\sigma} P_\rho$. Hence images of different P_ρ are mutually orthogonal subspaces, providing a numerically stable decomposition for downstream energy/covariance block-diagonalization.

5. Symmetry-resolved free energy

We present two equivalent routes (Hessian / Gaussian route and statistical / covariance route), prove their equivalence under the harmonic approximation, treat practical numerical issues (zero modes, regularization, finite-sample bias), and define diagnostic irrep scores that attribute the free-energy change to symmetry sectors.

5.1. Harmonic (Hessian) route

5.1.1. Local quadratic approximation

Let $\mathbf{x} \in \mathbb{R}^d$ denote local internal coordinates (collective coordinates, normal-mode coordinates, or small displacements) measured relative to a stable configuration μ_G that depends on the symmetry state $G \in \{D_2, C_2\}$. Taylor-expand the potential energy about μ_G up to second order [4]:

$$E_G(\mathbf{x}) = E_0(G) + 1/2(\mathbf{x} - \mu_G)^\top K_G (\mathbf{x} - \mu_G) + O(\|\mathbf{x} - \mu_G\|^3) \quad (83)$$

where $K_G = \nabla^2 E_G(\mu_G)$ is the symmetric positive-definite Hessian (we assume a local minimum so $K_G \succ 0$; see zero-mode handling below). $E_0(G)$ is the potential energy minimum (baseline).

5.1.2. Partition function in the quadratic approximation

The canonical partition function (restricting to coordinates $\mathbf{x}$) is

$$Z(G) = \int_{\mathbb{R}^d} e^{-\beta E_G(\mathbf{x})} d\mathbf{x} \approx e^{-\beta E_0(G)} \int_{\mathbb{R}^d} \exp\left(-\beta/2(\mathbf{x} - \mu_G)^\top K_G (\mathbf{x} - \mu_G)\right) d\mathbf{x} \quad (84)$$

The Gaussian integral is standard:

$$\int_{\mathbb{R}^d} \exp\left(-\beta/2 \mathbf{x}^\top K \mathbf{x}\right) d\mathbf{x} = (2\pi/\beta)^{d/2} (\det K)^{-1/2} = (2\pi k_B T)^{d/2} (\det K)^{-1/2} \quad (85)$$

Therefore

$$Z(G) \approx e^{-\beta E_0(G)} (2\pi k_B T)^{d/2} (\det K_G)^{-1/2} \quad (86)$$

5.1.3. Free energy and logdetk term

Take $F(G) = -k_B T \ln Z(G)$ to obtain (up to additive constants independent of G)

$$F(G) \approx E_0(G) + \frac{k_B T}{2} \ln \det K_G + \text{const} \quad (87)$$

Here “const” contains $(d/2)k_B T \ln(2\pi k_B T)$ and any Jacobian factors from coordinate choices; since we always compute differences between G states, such G -independent constants cancel.

5.1.4. Symmetry and block-diagonalization

Suppose $U(g)$ is the orthogonal/unitary representation of G acting on the coordinate space so that E_G is G -invariant in the sense $E_G(U(g)x) = E_G(x)$ whenever g is a symmetry of that configuration. If the Hessian itself is G -invariant (i.e. $U(g)^T K_G U(g) = K_G$ for all g in the symmetry group of that structure), then K_G commutes with the representation operators $U(g)$:

$$[K_G, U(g)] = 0, \forall g \in G \quad (88)$$

By Schur's lemma and the general theory of finite-group representations (Maschke's theorem), $V = \mathbb{R}^d$ decomposes into isotypic components corresponding to irreducible representations ρ :

$$V = \bigoplus_{\text{isotypic component } \rho} V_\rho \otimes \mathbb{C}^{m_\rho} \quad (89)$$

and any operator commuting with $U(g)$ is block-diagonal with respect to this decomposition. Concretely, let P_ρ be the orthogonal projector onto the ρ -isotypic subspace (constructed via the character projector). Then

$$K_G = \bigoplus_\rho K_{\rho,G}, \quad K_{\rho,G} := P_\rho^T K_G P_\rho \text{ acting on } \text{Im}(P_\rho) \quad (90)$$

Each block $K_{\rho,G}$ is itself symmetric positive-definite on its subspace.

5.1.5. Dimension bookkeeping

Let $n_\rho = \text{rank}(P_\rho)$ denote the effective dimensionality of the ρ -block (for abelian groups n_ρ equals the multiplicity). Then $\sum_\rho n_\rho = d$.

5.1.6. Factorization of the gaussian integral

Because K_G is block-diagonal in this orthogonal decomposition, the determinant factorizes:

$$\det K_G = \prod_\rho \det K_{\rho,G} \quad (91)$$

The symmetry-resolved free energy:

$$F(G) \approx E_0(G) + \frac{k_B T}{2} \sum_\rho \ln \det K_{\rho,G} + \text{const} \quad (92)$$

5.1.7. Symmetry-resolved free-energy difference

Subtracting the D_2 and C_2 expressions gives

$$\Delta F \equiv F(C_2) - F(D_2) = \frac{k_B T}{2} \sum_{\rho} [\text{ln det } K_{\rho, C_2} - \text{ln det } K_{\rho, D_2}] + \Delta E_0 \quad (93)$$

where $\Delta E_0 \equiv E_0(C_2) - E_0(D_2)$ is the baseline potential-energy difference (can be approximated from enthalpic terms such as FoldX).

5.2. Statistical (Covariance) route

5.2.1. Linear-response relation between hessian and covariance

Under harmonic fluctuations at temperature T , equipartition and Gaussian statistics give [4]

$$C_G = \langle (x - \mu_G)(x - \mu_G)^{\top} \rangle = k_B T K_G^{-1} \quad (94)$$

provided samples are drawn from the quadratic Boltzmann weight $\propto e^{-\beta x^{\top} K_G x / 2}$. This is the standard fluctuation–dissipation relation.

5.2.2. Express $\text{ln det } K$ via $\text{ln det } C$

Taking determinants and logarithms on each ρ -block yields

$$\text{ln det } K_{\rho, G} = -\text{ln det } C_{\rho, G} + n_{\rho} \ln(k_B T) \quad (95)$$

where $n_{\rho} = \text{rank}(P_{\rho})$.

Substituting and absorbing the $n_{\rho} \ln(k_B T)$ terms into the baseline gives the covariance form

$$\Delta F \approx -\frac{k_B T}{2} \sum_{\rho} [\text{ln det } C_{\rho, C_2} - \text{ln det } C_{\rho, D_2}] + \Delta E_0', \quad (96)$$

where $\Delta E_0'$ differs from ΔE_0 by the additive constant $k_B T / 2 \sum_{\rho} n_{\rho} \ln(k_B T)$ (which cancels in comparisons or can be included in the baseline).

Thus the Hessian and covariance routes are equivalent under the harmonic approximation and when C and K are invertible on the projected subspaces.

5.2.3. Eigenvalue (mode) representation

Let $\{\lambda_{\rho, i}^{(G)}\}_{i=1}^{n_{\rho}}$ denote the positive eigenvalues of $K_{\rho, G}$ (Hessian modes in the ρ channel). Then

$$\text{ln det } K_{\rho, G} = \sum_{i=1}^{n_{\rho}} \ln \lambda_{\rho, i}^{(G)} \quad (97)$$

Equivalently, for covariance eigenvalues $\{\sigma_{\rho, i}^{2, (G)}\}$ (so $\sigma^2 = k_B T / \lambda$),

$$\text{ln det } C_{\rho, G} = \sum_{i=1}^{n_{\rho}} \ln \sigma_{\rho, i}^{2, (G)} \quad (98)$$

Therefore each mode contributes additively to $F(G)$ and to ΔF ; this allows per-mode attribution.

5.3. Diagnostic irrep scores and linear-response approximation

5.3.1. Definition of diagnostic scores

We define two complementary irrep-resolved diagnostics:

$$\begin{aligned}\Delta_{\rho}^{(\text{spec})} &:= \text{ln det} \left(C_{\rho, C_2} C_{\rho, D_2}^{-1} \right) = \text{ln det} C_{\rho, C_2} - \text{ln det} C_{\rho, D_2}, \\ \Delta_{\rho}^{(\text{var})} &:= \text{tr} \left(C_{\rho, C_2} - C_{\rho, D_2} \right).\end{aligned}\quad (99)$$

The contribution of irrep ρ to ΔF (covariance form) is [4]

$$\Delta F_{\rho} = -\frac{k_B T}{2} \Delta_{\rho}^{(\text{spec})} \quad (100)$$

Large positive $\Delta_{\rho}^{(\text{spec})}$ (i.e. larger $\text{det} C$ in C_2) yields a negative contribution to ΔF (i.e. stabilization of the C_2 basin relative to D_2), and vice versa.

5.3.2. First-order (linear-response) sensitivity of ln det

If C is perturbed by a small symmetric δC , then

$$\text{ln det}(C + \delta C) = \text{ln det} C + \text{tr}(C^{-1} \delta C) - 1/2 \text{tr} \left(C^{-1} \delta C C^{-1} \delta C \right) + O(\|\delta C\|^3) \quad (101)$$

Thus, when differences between C_{ρ, C_2} and C_{ρ, D_2} are small, the leading contribution is

$$\Delta_{\rho}^{(\text{spec})} \approx \text{tr} \left(C_{\rho, D_2}^{-1} (C_{\rho, C_2} - C_{\rho, D_2}) \right) \quad (102)$$

This gives a practical linearized approximation to ΔF_{ρ} and shows the direct connection between $\Delta_{\rho}^{(\text{var})}$ (trace difference) and $\Delta_{\rho}^{(\text{spec})}$ via the inverse covariance weighting.

5.3.3. Interpretation of signs

If $C_{\rho, C_2} \succ C_{\rho, D_2}$ in a generalized sense (more variance in many directions), then $\text{ln det} C_{\rho, C_2} > \text{ln det} C_{\rho, D_2}$, so $\Delta_{\rho}^{(\text{spec})} > 0$ and $\Delta F_{\rho} < 0$: the C_2 basin gains entropic stabilization in channel ρ .

Conversely, reduced variance in C_2 relative to D_2 (tightening) gives positive ΔF_{ρ} (destabilization of C_2).

5.4. Practical computation: projection, eigen-decomposition, and numerics

5.4.1. Algorithmic recipe

1. Choose feature / CV vector $w \in \mathbb{R}^m$ that captures interface energies, contact counts, active-site geometries, etc., and collect N samples $w^{(1)}, \dots, w^{(N)}$ from MD/ENM/resampled FoldX ensembles for each state G .

2. Compute empirical covariance $\hat{C}_G = \frac{1}{N-1} \sum_{j=1}^N (w^{(j)} - \bar{w})(w^{(j)} - \bar{w})^{\top}$.

3. Construct representation matrices $U(g)$ on the CV space (permutation/sign or linear action) and form projectors $P_p = \frac{\dim p}{|G|} \sum_g \chi_p(g) U(g)$.
4. Compute projected covariance blocks $\hat{C}_{p,G} = P_p \hat{C}_G P_p^\top$.
5. Compute $\text{slogdet}(\hat{C}_{p,G})$ (numerically stable log-determinant via e.g. eigenvalues or Cholesky with regularization); accumulate

$$\widehat{\Delta F} = -\frac{k_B T}{2} \sum_p \left[\text{slogdet}(\hat{C}_{p,C_2}) - \text{slogdet}(\hat{C}_{p,D_2}) \right] + \widehat{\Delta E_0} \quad (103)$$

6. Estimate confidence intervals by bootstrap resampling of the sample set $\{w^{(i)}\}$ (resample replicates, recompute $\hat{C}_{p,G}$ and $\widehat{\Delta F}$).

5.4.2. Zero modes and coordinate gauge

Physical systems contain trivial zero modes (overall translations and rotations) that give zero eigenvalues in K and divergences in $\det K$. Remedies:

Project out rigid-body modes from w (work in internal/coarse-grained coordinates) or perform computations in internal coordinates where rigid motions are absent.

Remove near-zero eigenvalues before computing log-determinant, i.e. compute the product over nonzero eigenvalues only, or add a small regularizer ϵI and track the dependence on ϵ .

5.4.3. Regularization and finite-sample stability

Empirical $\hat{C}$ may be rank-deficient or ill-conditioned when N is not much larger than the projected dimension p_p . Use:

Ridge regularization: $\hat{C}_{p,G}^{(\epsilon)} = \hat{C}_{p,G} + \epsilon I$ with $\epsilon > 0$ small; compute $\ln \det$ of this regularized matrix. Choose ϵ by cross-validation or L-curve inspection.

Shrinkage estimators (Ledoit–Wolf): $\tilde{C} = (1-\lambda)\hat{C} + \lambda T$ with target T (e.g. diagonal), gives lower-variance $\ln \det$ estimates.

Dimensionality reduction: retain only principal components that capture a large fraction of variance within each p block; compute $\ln \det$ on the reduced block (adds a model selection step).

5.4.4. Stable evaluation of $\ln \det$

Compute the log-determinant via slogdet routines (Cholesky if positive-definite, or eigen-decomposition)

$$\ln \det C = \sum_{i=1}^p \ln \lambda_i \quad (104)$$

where $\{\lambda_i\}$ are eigenvalues. Use numerically stable libraries (e.g. LAPACK routines) and avoid forming full dense inverses.

5.5. Baseline energy ΔE_0 and mapping FoldX outputs

5.5.1. What ΔE_0 represents

ΔE_0 is the difference in basin minima energies $E_0(C_2) - E_0(D_2)$. In practice one often uses empirical or computed enthalpic proxies (FoldX total or interface energies) as an approximation:

$$\Delta E_0 \approx E_{\text{FoldX}}(C_2) - E_{\text{FoldX}}(D_2) \quad (105)$$

with the caveat that FoldX energies are not exact free energies (lack full entropy).

5.5.2. Character-weighted baseline separation

To retain symmetry attribution in the baseline enthalpy, decompose per-group-element energies $E(g)$ via character inner products:

$$\langle \chi_p, E \rangle_G = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_p(g)} E(g) \quad (106)$$

A simple model for baseline difference is

$$\Delta E_0 \approx \sum_p \frac{\text{dimp}}{|G|} \left(\langle \chi_p, E \rangle_{C_2} - \langle \chi_p, E \rangle_{D_2} \right) \quad (107)$$

i.e. project the FoldX energies into irrep channels and sum the channel differences. Treat this as an enthalpic proxy to be combined with the fluctuation-derived terms [5,6].

5.6. Final expressions and per-irrep contributions

5.6.1. Final covariance-based formula

$$\widehat{\Delta F} = -\frac{k_B T}{2} \sum_p \left[\ln \det \widehat{C}_{\rho, C_2} - \ln \det \widehat{C}_{\rho, D_2} \right] + \widehat{\Delta E_0}, \quad (108)$$

with $\widehat{C}_{\rho, G} = P_\rho \widehat{C}_G P_\rho^\top$ and $\widehat{\Delta E_0}$ the chosen baseline enthalpy difference.

5.6.2. Per-irrep free-energy contribution

Define

$$\Delta F_\rho = -\frac{k_B T}{2} \left[\ln \det C_{\rho, C_2} - \ln \det C_{\rho, D_2} \right] + \Delta E_{0, \rho} \quad (109)$$

so that $\Delta F = \sum_p \Delta F_\rho$. Here $\Delta E_{0, \rho}$ denotes the irrep-resolved baseline enthalpy term (from FoldX projection).

5.6.3. First-order attribution using the linear approximation

If $\delta C_\rho := C_{\rho, C_2} - C_{\rho, D_2}$ is small,

$$\Delta F_\rho \approx -\frac{k_B T}{2} \text{tr} \left(C_{\rho, D_2}^{-1} \delta C_\rho \right) + \Delta E_{0, \rho} \quad (110)$$

This linear form is useful to identify dominant directions using the eigenvectors of C_{ρ, D_2}^{-1} (i.e. high-sensitivity directions).

6. From free energy to efficiency: a Transition-State-Theory (TST) bridge

Notation and preliminaries. We denote by $G \in \{D_2, C_2\}$ the symmetry group of the enzyme assembly (tetramer vs dimer). For a given symmetry state G we write:

- $Q_R(G)$ (or $Z_R(G)$) for the reactant-basin partition function (reactant ensemble),

- $Q^\ddagger(G)$ for the transition-state (TS) partition function associated with the reactive dividing surface,
- $F_R(G) = -k_B T \ln Q_R(G)$ for the reactant free energy, and
- $F^\ddagger(G) = -k_B T \ln Q^\ddagger(G)$ for the TS free energy.

We also define the free-energy difference already used in the manuscript:

$$\Delta F = F_R(C_2) - F_R(D_2) \quad (111)$$

6.1. TST basic formula and species ratio

Transition-state theory gives (up to the usual prefactor and a transmission coefficient) [7]

$$k(G) = \frac{k_B T}{h} \kappa(G) \frac{Q^\ddagger(G)}{Q_R(G)} = \frac{k_B T}{h} \kappa(G) \exp(-\beta \Delta G^\ddagger(G)) \quad (112)$$

where $\kappa(G) \in (0, 1]$ is the transmission (or recrossing) coefficient for state G and

$$\Delta G^\ddagger(G) = F^\ddagger(G) - F_R(G) \quad (113)$$

is the activation free energy measured relative to the reactant basin.

Now compare dimer (C_2) and tetramer (D_2). Define also m_G as the number of equivalent catalytic channels (active sites) per oligomer: for a homotetramer $m_{D_2} = 4$, for a homodimer $m_{C_2} = 2$ (unless some sites are silent). The per-oligomer (or per-species) catalytic capability scales with m_G , so the ratio of specificity-like constants (k_{cat}/K_M) (under the rapid-equilibrium approximation for binding) may be written schematically as

$$\boxed{\frac{(k_{cat}/K_M)_{C_2}}{(k_{cat}/K_M)_{D_2}} \approx \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \frac{Q^\ddagger(C_2)/Q_R(C_2)}{Q^\ddagger(D_2)/Q_R(D_2)} = \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \exp\left[-\beta \left(\Delta G^\ddagger_{C_2} - \Delta G^\ddagger_{D_2}\right)\right]} \quad (114)$$

This is identical to the short boxed formula you gave; we now unpack and connect it to the ΔF expressions from the symmetry-resolved free-energy analysis.

6.2. Partition-function form and relation to ΔF

Using $F_R(G) = -k_B T \ln Q_R(G)$ and $F^\ddagger(G) = -k_B T \ln Q^\ddagger(G)$, expands to

$$\Delta G^\ddagger_{C_2} - \Delta G^\ddagger_{D_2} = (F^\ddagger(C_2) - F^\ddagger(D_2)) - (F_R(C_2) - F_R(D_2)) \quad (115)$$

Hence

$$\frac{Q^\ddagger(C_2)/Q_R(C_2)}{Q^\ddagger(D_2)/Q_R(D_2)} = \exp\left[-\beta \left(\Delta G^\ddagger_{C_2} - \Delta G^\ddagger_{D_2}\right)\right] = \exp\left[-\beta \left(F^\ddagger(C_2) - F^\ddagger(D_2)\right)\right] \exp\left[\beta \left(F_R(C_2) - F_R(D_2)\right)\right] \quad (116)$$

6.2.1. Interpretation

The rate ratio thus splits into two conceptually separate effects:

$$(\text{TS-shift}) \times (\text{reactant-baseline-shift}) \quad (117)$$

The reactant-baseline term contains the ΔF we computed in the symmetry-resolved analysis:

$$\exp [\beta (F_R(C_2) - F_R(D_2))] = \exp(\beta \Delta F) \quad (118)$$

Therefore the full ratio can be written as

$$\frac{(k_{\text{cat}}/K_M)_{C_2}}{(k_{\text{cat}}/K_M)_{D_2}} = \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \exp(\beta \Delta F) \cdot \exp [-\beta (F^\ddagger(C_2) - F^\ddagger(D_2))] \quad (119)$$

6.2.2. Two limiting cases

1. TS is invariant under symmetry change. If the transition-state free energy is essentially the same for the two assemblies (i.e. $F^\ddagger(C_2) \approx F^\ddagger(D_2)$), the TS-shift factor is unity and

$$\frac{(k_{\text{cat}}/K_M)_{C_2}}{(k_{\text{cat}}/K_M)_{D_2}} \approx \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \exp(\beta \Delta F) \quad (120)$$

In this scenario a higher reactant free energy $F_R(C_2)$ (i.e. $\Delta F > 0$) increases the rate of C_2 relative to D_2 because the barrier measured from the reactant basin is effectively lower.

2. Barrier shift parallels reactant shift (barrier measured in absolute energy). If $F^\ddagger(C_2) - F^\ddagger(D_2) \approx F_R(C_2) - F_R(D_2) = \Delta F$ (i.e. both TS and reactant basin shift by the same absolute energy so the absolute barrier is unchanged), then the exponential terms cancel and

$$\frac{(k_{\text{cat}}/K_M)_{C_2}}{(k_{\text{cat}}/K_M)_{D_2}} \approx \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \quad (121)$$

That is, only multiplicity and dynamical recrossing differences remain.

Important note on sign conventions: in earlier sections we defined $\Delta F = F_R(C_2) - F_R(D_2)$. When you see formulas of the form $\exp(-\beta \Delta F)$ in other parts of the manuscript, verify the context — sometimes authors report efficiency $\propto Q_R$ (not $\propto 1/Q_R$). The TST result above is unambiguous once $Q^\ddagger$ and Q_R are explicitly identified.

6.3. TS-probability surrogate: $p(\text{TS})$ and $\Delta G^\ddagger$

A convenient experimental / data-driven surrogate for the activation free energy is the TS-region occupancy. Define a TS-like region S_{TS} in CV space (geometric thresholds around the dividing surface). Its Boltzmann weight relative to the reactant basin is

$$p_{\text{TS}}(G) := \frac{\int_{S_{\text{TS}}} e^{-\beta E(x)} dx}{\int_{\text{reactant}} e^{-\beta E(x)} dx} \approx \frac{Q^\ddagger(G)}{Q_R(G)} \quad (122)$$

Hence [5]

$$\Delta G^\ddagger(G) = -k_B T \ln \frac{Q^\ddagger(G)}{Q_R(G)} \approx -k_B T \ln p_{\text{TS}}(G) \quad (123)$$

An immediately estimable ratio:

$$\frac{(k_{cat}/K_M)_{C_2}}{(k_{cat}/K_M)_{D_2}} \approx \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \frac{p_{TS}(C_2)}{p_{TS}(D_2)} \quad (124)$$

This is practical: compute or estimate p_{TS} from MD or by coarse-grained CV sampling (subject to careful definition of S_{TS}).

6.4. Speciation / oligomeric equilibrium and observed (bulk) kinetics

6.4.1. Equilibrium between dimer and tetramer

In a preparation where both dimers (D) and tetramers (T) can exist and interconvert via



define the (dissociation-like) equilibrium constant

$$K_d = \frac{[D]^2}{[T]} \quad (126)$$

Let C_{tot} denote the total subunit concentration (monomer equivalents)

$$C_{tot} = 4 [T] + 2 [D] \quad (127)$$

Solving for $[D]$ in terms of C_{tot} and K_d yields a quadratic in $d := [D]$:

$$4 \frac{d^2}{K_d} + 2d - C_{tot} = 0 \quad (128)$$

Hence

$$d = [D] = \frac{-K_d + \sqrt{K_d^2 + 4K_d C_{tot}}}{4} \quad (129)$$

taking the physically positive root. The tetramer concentration is then

$$[T] = \frac{d^2}{K_d} \quad (130)$$

6.4.2. Observed (bulk) catalytic rate under substrate-limited linear regime

At low substrate (initial-rate linear regime), each active site contributes approximately $(k_{cat}/K_M)[S]$ to the second-order rate, so the bulk initial rate per unit volume is

$$v_0 \approx [S] (m_T [T] (k_{cat}/K_M)_T + m_D [D] (k_{cat}/K_M)_D) \quad (131)$$

It is often convenient to normalize by total subunit concentration C_{tot} to obtain an observed efficiency per subunit:

$$\left(\frac{k_{\text{cat}}}{K_M}\right)_{\text{obs}} = \frac{m_T [T] (k_{\text{cat}}/K_M)_T + m_D [D] (k_{\text{cat}}/K_M)_D}{C_{\text{tot}}} \quad (132)$$

Using the solution for $[T], [D]$ above one can predict how measured bulk kinetics vary with total concentration and K_d .

6.4.3. Alternative normalization (per oligomer molecule)

If instead one reports rate per oligomer molecule (not per subunit), define $N_{\text{mol}} = [T] + [D]$, and the per-molecule observed efficiency is

$$\left(\frac{k_{\text{cat}}}{K_M}\right)_{\text{mol}} = \frac{m_T [T] (k_{\text{cat}}/K_M)_T + m_D [D] (k_{\text{cat}}/K_M)_D}{[T] + [D]} \quad (133)$$

Choose the normalization that matches how experimental data are reported.

6.5. Assumptions, caveats, and practical recommendations

6.5.1. Assumptions made in the bridge

Rapid oligomeric equilibration: we assumed that the $D \rightleftharpoons T$ interconversion is fast compared to catalysis (so the equilibrium distribution holds during initial-rate measurement). If not, a kinetic model including interconversion rates must be used.

Well-defined TS partition function: $Q^\ddagger$ must be meaningfully defined (requires a reasonable dividing surface in CV space).

Separable effects: we treated multiplicity m_G , transmission κ_G , and free-energy terms multiplicatively; in reality these can be coupled (e.g. interface changes may alter reaction coordinate friction and hence κ).

Harmonic / local approximations: when expressing Q_R via ΔF we typically used the Gaussian (log-det) approximation for fluctuation contributions; large anharmonic changes require more careful sampling.

7. Why the regular-character convolution fails

A tempting but incorrect formula is [8]

$$\Delta F = \frac{1}{|D_2|} \sum_{g \in D_2} (E(g) - E(\phi(g))) \chi_{\text{reg}}(g) \quad (134)$$

where χ_{reg} denotes the character of the regular representation of the group (here D_2). We unpack why this formula is mathematically and physically unsound, and we show the correct operator-based alternative.

7.1. Why the naive formula is algebraically trivial

Recall that for any finite group G , the regular character satisfies [2]

$$\chi_{\text{reg}}(g) = \begin{cases} |G|, & g=e, \\ 0, & g \neq e. \end{cases} \quad (135)$$

Substituting this

$$\frac{1}{|D_2|} \sum_{g \in D_2} (E(g) - E(\Phi(g))) \chi_{\text{reg}}(g) = \frac{1}{|D_2|} (E(e) - E(\Phi(e))) |D_2| = E(e) - E(\Phi(e)) \quad (136)$$

Because Φ is a homomorphism with $\Phi(e)=e$, the right-hand side reduces to $E(e)-E(e)=0$, so $\Delta F=0$ identically. Thus the formula cannot encode any nontrivial structural information — it is algebraically nullified by the properties of the regular character.

7.2. Conceptual reason: scalars vs operators

The underlying conceptual mistake is treating $E(g)$ as if it were the full object controlling the free-energy change under symmetry, while the free energy (in the Gaussian/harmonic approximation) is controlled by operators (Hessians or covariance matrices) whose spectra determine $\det K$ or $\det C$ and thus the entropic part of the free energy.

Concretely:

$E(g)$ is a scalar-valued function on the group elements — e.g. an interface enthalpy associated (by some choice) to the labeling induced by g .

The free energy in the harmonic regime is

$$F(G) \approx E_0(G) + \frac{k_B T}{2} \sum_p \ln \det K_{p,G} \quad (137)$$

so it depends on determinants of matrices $K_{p,G}$ (or equivalently the spectra of covariance blocks $C_{p,G}$). Spectral information cannot be recovered from a single scalar per group element.

A simple counterexample (illustrative). Consider two different Hessians $K^{(1)}$ and $K^{(2)}$ that, under some ad-hoc mapping, yield identical scalar lists $\{E(g)\}$ but have different eigenvalue spectra. Their $\ln \det K$ will differ, hence their F differ, while the scalar convolution returns zero (or the same trivial value) and misses the actual free-energy difference.

7.3. Correct object: projectors acting on operators

The correct symmetry-aware decomposition acts on operators, not scalars. Construct the character projectors

$$P_p = \frac{\dim p}{|G|} \sum_{g \in G} \overline{\chi_p(g)} U(g) \quad (138)$$

and apply them to the operator of interest (Hessian K or covariance C):

$$K_{p,G} = P_p K_G P_p^\top, C_{p,G} = P_p C_G P_p^\top \quad (139)$$

Then each p -block contains the full spectral information for that symmetry channel, and the free-energy difference is recovered from the $\log \det$ of those blocks:

$$\Delta F = -\frac{k_B T}{2} \sum_p [\ln \det C_{p,C_2} - \ln \det C_{p,D_2}] + \Delta E_0 \quad (140)$$

8. Free energy difference as a symbolic expression

8.1. Setup and notation

Let $G \in \{D_2, C_2\}$ denote the oligomeric symmetry state. The six inter-subunit interface variables are collected into

$$\mathbf{w} \in \mathbb{R}^6, \mu^{(G)} = \mathbb{E}[\mathbf{w}|G], \Sigma^{(G)} = \text{Cov}(\mathbf{w}|G) \quad (141)$$

We will denote by n the dimension of the fluctuating subspace under consideration (here $n \leq 6$ after removing any constrained/rigid modes). For each irreducible representation ρ of D_2 define the projection operator

$$P_\rho = \frac{\dim \rho}{|D_2|} \sum_{g \in D_2} \overline{\chi_\rho(g)} U(g) \quad (142)$$

where $U(g)$ are the permutation (orthogonal) matrices on interface space and χ_ρ the characters. These satisfy [2]

$$P_\rho P_\sigma = \delta_{\rho\sigma} P_\rho, P_\rho^\top = P_\rho, \sum_\rho P_\rho = I \quad (143)$$

We define the symmetry-resolved covariances and means by

$$\Sigma_\rho^{(G)} = P_\rho \Sigma^{(G)} P_\rho^\top, \mu_\rho^{(G)} = P_\rho \mu^{(G)} \quad (144)$$

A convenient parameterization of the mean vector is to assign one variable to each orbit of interfaces under D_2 :

$$\mu^{(G)} = \begin{bmatrix} s_1^{(G)} \\ s_2^{(G)} \\ s_1^{(G)} \\ s_2^{(G)} \\ s_3^{(G)} \\ s_3^{(G)} \end{bmatrix} \quad (145)$$

corresponding to the interfaces (AB, BC, CD, DA, AC, BD).

8.2. Gaussian (harmonic) partition function — full derivation

Let the microscopic conformational coordinate be $\mathbf{x} \in \mathbb{R}^d$. The energy function $E_G(\mathbf{x})$ respects the symmetry action $U(g)$ of group G . The constrained partition function (or $Z(G)$) is expressed as a Reynolds average to ensure only G -equivalent configurations contribute:

$$Z(G) = \frac{1}{|G|} \sum_{g \in G} \int_{\mathbb{R}^d} \exp(-\beta E_G(U(g)\mathbf{x})) d\mathbf{x}, \beta = (k_B T)^{-1} \quad (146)$$

Assume the reactant-basin energy is quadratic about its minimum $\mathbf{w}_0$ (or μ_G):

$$E(\mathbf{w}) \approx E_0(G) + \frac{1}{2}(\mathbf{w} - \mathbf{w}_0)^\top \mathbf{K}_G (\mathbf{w} - \mathbf{w}_0) \quad (150)$$

with symmetric positive-semidefinite Hessian $\mathbf{K}_G$ of size $n \times n$ (restricted to the fluctuation subspace). The reactant partition function in the harmonic approximation is

$$\begin{aligned} Q_R(G) &= Z(G) = \int_{\mathbb{R}^n} e^{-\beta E(w)} dw \approx e^{-\beta E_0(G)} \int_{\mathbb{R}^n} e^{-\frac{\beta}{2} y^\top K_G y} dy \\ &= e^{-\beta E_0(G)} (2\pi)^{n/2} (\det(\beta K_G))^{-1/2}, \end{aligned} \quad (151)$$

where we used the Gaussian integral identity $\int e^{-1/2 y^\top A y} dy = (2\pi)^{n/2} (\det A)^{-1/2}$ for $A \succ 0$. Taking the free energy $F(G) = -k_B T \ln Q_R(G)$ yields

$$F(G) = E_0(G) + \frac{k_B T}{2} \ln \det K_G + \frac{k_B T n}{2} \ln \beta - \frac{k_B T n}{2} \ln(2\pi) \quad (152)$$

The last two terms are G -independent constants (they depend only on n and T) and may be absorbed into “const” in what follows.

It is often convenient to express F in terms of the covariance matrix

$$\Sigma^{(G)} = \langle yy^\top \rangle = (\beta K_G)^{-1} \quad (153)$$

whence $\det K_G = \beta^{-n} \det(\Sigma^{(G)})^{-1}$ and

$$F(G) = E_0(G) - \frac{k_B T}{2} \ln \det \Sigma^{(G)} + \text{const} \quad (154)$$

This form makes the entropic role of the covariance explicit: larger covariance $\Rightarrow$ larger $\det \Sigma \Rightarrow$ lower F (more entropy).

8.3. Symmetry-resolved blocks and factorization

If the basin/hessian K_G is invariant under the group action (i.e. $U(g)K_GU(g)^\top = K_G$ for all g), then K_G commutes with every $U(g)$. By standard representation theory one may choose an orthonormal basis that simultaneously block-diagonalizes all $U(g)$ and K_G so that

$$K_G \cong \bigoplus_\rho K_{\rho,G}, \quad \Sigma^{(G)} \cong \bigoplus_\rho \Sigma_\rho^{(G)} \quad (155)$$

with

$$K_{\rho,G} = P_\rho K_G P_\rho^\top, \quad \Sigma_\rho^{(G)} = P_\rho \Sigma^{(G)} P_\rho^\top \quad (156)$$

Determinants factorize over blocks, we obtain the symmetry-resolved free energy

$$F(G) = E_0(G) - \frac{k_B T}{2} \sum_\rho \ln \det \Sigma_\rho^{(G)} + \text{const} \quad (157)$$

Subtracting the two symmetry states yields the central symbolic expression:

$$\boxed{\Delta F = F(C_2) - F(D_2) = \Delta E_0 - \frac{k_B T}{2} \sum_\rho \left(\ln \det \Sigma_\rho^{(C_2)} - \ln \det \Sigma_\rho^{(D_2)} \right)}, \quad (158)$$

where $\Delta E_0 = E_0(C_2) - E_0(D_2)$. (Note: For positive-definite matrices, $\ln \det \Sigma = \ln |\Sigma|$.)

Remark on zero modes / pseudo-determinants. If K_G has zero eigenvalues (rigid translations/rotations or constrained directions), the Gaussian integral is formally divergent. Practically one removes those rigid modes (restrict to fluctuation subspace) and uses the pseudodeterminant

$$\text{pdet } A = \prod_{\lambda_i > 0} \lambda_i \quad (159)$$

replacing $\det$. Equivalently, fix gauges or integrate out rigid coordinates—this affects only the absorbed “const” and not the difference between symmetry states if the zero-mode count is the same.

8.4. Orbit-based parameterization and explicit closed form

Label the six unordered edges as

$$(AB, BC, CD, DA, AC, BD) \quad (160)$$

which form three D_2 -orbits of size two:

$$O_1 = \{AB, CD\}, \quad O_2 = \{BC, DA\}, \quad O_3 = \{AC, BD\} \quad (161)$$

Assume the covariance is block-diagonal by orbit (no cross-orbit covariances):

$$\Sigma^{(G)} = \text{diag} \left(\Sigma_1^{(G)}, \Sigma_2^{(G)}, \Sigma_3^{(G)} \right), \quad \Sigma_i^{(G)} = \begin{bmatrix} v_i^{(G)} & c_i^{(G)} \\ c_i^{(G)} & v_i^{(G)} \end{bmatrix}, \quad i=1,2,3 \quad (162)$$

$$\Sigma^{(G)} = \begin{bmatrix} v_1^{(G)} & c_1^{(G)} & 0 & 0 & 0 & 0 \\ c_1^{(G)} & v_1^{(G)} & 0 & 0 & 0 & 0 \\ 0 & 0 & v_2^{(G)} & c_2^{(G)} & 0 & 0 \\ 0 & 0 & c_2^{(G)} & v_2^{(G)} & 0 & 0 \\ 0 & 0 & 0 & 0 & v_3^{(G)} & c_3^{(G)} \\ 0 & 0 & 0 & 0 & c_3^{(G)} & v_3^{(G)} \end{bmatrix} \quad (163)$$

Introduce normalized symmetric / antisymmetric orbit coordinates for orbit i :

$$u_{s,i} = \frac{e_{i,1} + e_{i,2}}{\sqrt{2}}, \quad u_{b,i} = \frac{e_{i,1} - e_{i,2}}{\sqrt{2}} \quad (164)$$

where $e_{i,1}, e_{i,2}$ are the standard basis vectors for the two edges in orbit i (e.g. for O_1 , $e_{1,1} = e_{AB}, e_{1,2} = e_{CD}$). Compute the variances in these coordinates:

$$\text{Var}(u_{s,i}) = v_i^{(G)} + c_i^{(G)}, \quad \text{Var}(u_{b,i}) = v_i^{(G)} - c_i^{(G)} \quad (165)$$

Under the orbit-decoupling assumption the projection onto irreps yields

$$\Sigma_{A_1}^{(G)} = \text{diag} \left(v_1^{(G)} + c_1^{(G)}, v_2^{(G)} + c_2^{(G)}, v_3^{(G)} + c_3^{(G)} \right) \quad (166)$$

and the nontrivial one-dimensional irreps give

$$\Sigma_{B_1}^{(G)} = \left(v_1^{(G)} - c_1^{(G)} \right), \Sigma_{B_2}^{(G)} = \left(v_2^{(G)} - c_2^{(G)} \right), \Sigma_{B_3}^{(G)} = \left(v_3^{(G)} - c_3^{(G)} \right) \quad (167)$$

Hence the determinants are products of these scalar entries, and substituting into explicit orbit form

$$\Delta F = \Delta E_0 - \frac{k_B T}{2} \sum_{i=1}^3 \ln \frac{\left(v_i^{(C_2)} + c_i^{(C_2)} \right) \left(v_i^{(C_2)} - c_i^{(C_2)} \right)}{\left(v_i^{(D_2)} + c_i^{(D_2)} \right) \left(v_i^{(D_2)} - c_i^{(D_2)} \right)}. \quad (168)$$

This is the closed-form symbolic expression in terms of the orbit-parameters $\{s_i^{(G)}, v_i^{(G)}, c_i^{(G)}\}$ and the enthalpic baseline difference ΔE_0 , enabling sensitivity analysis without specific numerical values.

Positivity constraint. For physical positive-definiteness one requires

$$v_i^{(G)} > |c_i^{(G)}|, \quad i=1,2,3, \quad (169)$$

so that each orbit-block is SPD and the logarithms are well-defined.

8.5. Sensitivity (partial derivatives) — how each parameter affects ΔF

Differentiating gives closed-form sensitivities. For a fixed orbit i and varying the C_2 parameters,

$$\begin{aligned} \frac{\partial \Delta F}{\partial v_i^{(C_2)}} &= -\frac{k_B T}{2} \left(\frac{1}{v_i^{(C_2)} + c_i^{(C_2)}} + \frac{1}{v_i^{(C_2)} - c_i^{(C_2)}} \right) \\ &= -k_B T \frac{v_i^{(C_2)}}{(v_i^{(C_2)})^2 - (c_i^{(C_2)})^2}, \end{aligned} \quad (170)$$

and

$$\begin{aligned} \frac{\partial \Delta F}{\partial c_i^{(C_2)}} &= -\frac{k_B T}{2} \left(\frac{1}{v_i^{(C_2)} + c_i^{(C_2)}} - \frac{1}{v_i^{(C_2)} - c_i^{(C_2)}} \right) \\ &= k_B T \frac{c_i^{(C_2)}}{(v_i^{(C_2)})^2 - (c_i^{(C_2)})^2}. \end{aligned} \quad (171)$$

Analogous formulas (with opposite sign inside the big parentheses) hold for derivatives with respect to $v_i^{(D_2)}$ and $c_i^{(D_2)}$.

8.5.1. Interpretation

- Increasing $v_i^{(C_2)}$ (more variance on orbit i in the dimer) decreases ΔF if $v_i^{(C_2)} > 0$ (entropy stabilizes C_2 relative to D_2).
- Increasing positive covariance $c_i^{(C_2)} > 0$ increases ΔF .
- The sensitivity scales as $1/(v^2 - c^2)$ and therefore grows if the block becomes nearly singular; this indicates directions where small structural changes produce large free-energy effects.

8.5.2. Numerical stability and regularization

When $(v_i^2 - c_i^2)$ is very small, add a small Tikhonov regularizer $\epsilon > 0$ to Σ (i.e. replace $\Sigma \mapsto \Sigma + \epsilon I$) to stabilize logarithms and derivatives; carry this through the algebra if needed for numerical work.

The formulas above provide a fully symbolic, algebraically explicit map

$$\{ \Delta E_0, s_i^{(G)}, v_i^{(G)}, c_i^{(G)} \} \mapsto \Delta F \quad (172)$$

suitable for sensitivity analysis, uncertainty propagation, and for guiding mutational design that targets particular orbit variances or correlations.

9. Conclusion

9.1. Summary of symbolic results

Starting from a harmonic basin approximation and symmetry-resolved block diagonalization we derived

$$F(G) = E_0(G) - \frac{k_B T}{2} \sum_p \ln \det \Sigma_p^{(G)} + \text{const} \quad (173)$$

and therefore

$$\Delta F = \Delta E_0 - \frac{k_B T}{2} \sum_p \left(\ln \det \Sigma_p^{(C_2)} - \ln \det \Sigma_p^{(D_2)} \right). \quad (174)$$

Under orbit-decoupling this reduces to the explicit orbit expression in terms of $\{s_i^{(G)}, v_i^{(G)}, c_i^{(G)}\}$.

9.2. Bridge to transition-state theory and efficiency

In the TST approximation the catalytic-efficiency ratio can be expressed schematically as

$$\frac{(k_{\text{cat}}/K_M)_{C_2}}{(k_{\text{cat}}/K_M)_{D_2}} \approx \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \exp \left[-\beta (\Delta G_{C_2}^\ddagger - \Delta G_{D_2}^\ddagger) \right] \quad (175)$$

Using the identity

$$\Delta G_{C_2}^\ddagger - \Delta G_{D_2}^\ddagger = (F^\ddagger(C_2) - F^\ddagger(D_2)) - (F_R(C_2) - F_R(D_2)) \quad (176)$$

and inserting the symmetry-resolved expression for $F_R(G)$ (i.e. ΔF), we isolate the reactant-baseline contribution:

$$\frac{(k_{\text{cat}}/K_M)_{C_2}}{(k_{\text{cat}}/K_M)_{D_2}} = \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \exp(\beta \Delta F) \cdot \exp \left[-\beta (F^\ddagger(C_2) - F^\ddagger(D_2)) \right] \quad (177)$$

Two useful limiting scenarios are:

1. TS invariant: if $F^\ddagger(C_2) \approx F^\ddagger(D_2)$ then

$$\frac{(k_{cat}/K_M)_{C_2}}{(k_{cat}/K_M)_{D_2}} \approx \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \cdot \exp(\beta \Delta F) \quad (178)$$

2. Absolute-barrier preserved: if $F^\ddagger(C_2) - F^\ddagger(D_2) \approx \Delta F$ (i.e. both TS and reactant shift in parallel) then exponential terms cancel and only multiplicity and dynamical factors remain:

$$\frac{(k_{cat}/K_M)_{C_2}}{(k_{cat}/K_M)_{D_2}} \approx \frac{m_{C_2}}{m_{D_2}} \cdot \frac{\kappa_{C_2}}{\kappa_{D_2}} \quad (179)$$

Absorbing the main symmetry effect into ΔF , a simplified statistical-kinetic bridge for the efficiency ratio $\eta = k_{cat}/K_M$ is

$$\boxed{\frac{\eta_{C_2}}{\eta_{D_2}} \approx \left(\frac{|C_2|}{|D_2|} \right)^{1/2} \exp \left(-\frac{\Delta F}{k_B T} \right)} \quad (180)$$

The prefactor $\sqrt{|C_2|/|D_2|}$ reflects reduced symmetry volume, and the exponential encodes the thermodynamic penalty.

9.3. Connection to FoldX data

For each group element g , let $E(g)$ denote FoldX energy components. Define the character inner product:

$$\langle \chi_\rho, E \rangle_G = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_\rho(g)} E(g) \quad (181)$$

$$\boxed{\Delta E_0 \approx \sum_\rho \frac{\text{dimp}}{|G|} \left(\langle \chi_\rho, E \rangle_{C_2} - \langle \chi_\rho, E \rangle_{D_2} \right)} \quad (182)$$

Similarly, $\Sigma_{\rho, G} = P_\rho \Sigma^{(G)} P_\rho^\dagger$ with w the feature vector ensemble.

9.4. Final synthesis

$$\boxed{\begin{aligned} \text{(i)} \quad & P_\rho = \frac{\text{dimp}}{|G|} \sum_g \overline{\chi_\rho(g)} U(g), \Sigma_{\rho, G} = P_\rho \Sigma^{(G)} P_\rho^\dagger; \\ \text{(ii)} \quad & \Delta F = -\frac{k_B T}{2} \sum_\rho [\ln \det \Sigma_{\rho, C_2} - \ln \det \Sigma_{\rho, D_2}] + \Delta E_0; \\ \text{(iii)} \quad & \frac{\eta_{C_2}}{\eta_{D_2}} \approx \left(\frac{|C_2|}{|D_2|} \right)^{1/2} \exp(-\Delta F/k_B T). \end{aligned}} \quad (183)$$

Acknowledgements

First and foremost, I would like to thank my parents for their unconditional support at all times and for inspiring me in mathematics and making me decide to take this path. Standing on their shoulders, I was able to glimpse higher and farther into the future, and embrace more possibilities. It was my parents who, with their unconditional love and tolerance, embraced every moment of my confusion, exhaustion, and anxiety, mended and healed the cracks in my life, and transformed the hardships and sufferings of my growth into a warm spiritual journey. It was they who made me understand that I never walk alone. No matter where I go, no matter whether I succeed or fail, there is always a light waiting for me to return home, giving me the courage and confidence to accept my imperfections and move on to the next intersection in life.

I would like to thank the professors who have guided me over the past four years, Professor Espinoza and Professor Kjeer. They have led me into the vast and profound world of mathematics, allowing me to experience the warmth of the wisdom left by our predecessors. It was their trust and recognition that soothed my anxieties and allowed me to see my own strengths. They have always accompanied me on my academic journey, leaving each tender commentary on my immature and clumsy steps. Even amidst my own setbacks, they have taught me that "there is no need to fear the infinite truth; every step forward brings its own joy." That ever-burning light has illuminated my small corner of refuge and given me the resilience to pursue my studies.

I would also like to thank my high school friend Yang Xilin for his help in biology during my research and thesis.

References

- [1] Gallian, J. A. (1990). Contemporary abstract algebra. New Delhi: Cengage Learning.
- [2] Fulton, W., & Harris, J. (2013). Representation theory: a first course (Vol. 129). *Springer Science & Business Media*.
- [3] Conway, J. B. (2019). A course in functional analysis (Vol. 96). *Springer*.
- [4] Horn, R. A., & Johnson, C. R. (2012). Matrix analysis. Cambridge University Press.
- [5] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... & Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Research*, 28(1), 235–242.
- [6] Berman, H. M., Henrick, K., Kleywegt, G., Nakamura, H., & Markley, J. (2012). RCSB PDB. *Nucleic Acids Research*, 40(D1), D445–D452.
- [7] Truhlar, D. G., Garrett, B. C., & Klippenstein, S. J. (1996). Current status of transition-state theory. *The Journal of Physical Chemistry*, 100(31), 12771–12800.
- [8] Tinkham, M. (2003). Group theory and quantum mechanics. Courier Corporation.