

Sufficient and Ancillary Reductions of a Chrysene Cluster

Transverse Seeding as a Decision Rule and the Mitochondrial Code as Normal Form

Charles D. Schaper, Ph.D.

Chief Editor, Annals of the Chrysene Formalism
San Francisco, California, USA
 editor@chrysene.com

Abstract

A chrysene cluster with an adenine tether presents a sixty-four-element word space partitioned in two ways: twenty frame classes and sixteen codon blocks. Excision of a residue is a function of the class. In the joint experiment that records the word together with the residue, the pair (T, r) is minimal Neyman-sufficient for the unknown excision map; among functions of the word alone, T is a minimal sufficient reduction; and the within-class coordinate is conditionally ancillary for that map. The raw four-valued frame coordinate is not independent of the class, and Basu's theorem does not apply.

The two partitions are transverse. The class is therefore not a sufficient reduction for the block, and assignments measurable with respect to both partitions are constant on each incidence component and take at most two values. A seed is a section of incidence; propagation is the unique block-measurable extension of a non-conflicting seed. Refinement at the third position, pyrimidine against purine, produces only the block patterns (4) and (2, 2). The standard genetic code realises also (3, 1) and (2, 1, 1), and lies outside that range. Under the four matching hypotheses, with chain polarity used only to identify the domain, and with demand read from the standard table, one refined assignment equals the vertebrate mitochondrial table. Parity of class-functional maps excludes the standard table — methionine and tryptophan are singleton fibres — and is silent on the mitochondrial table. The refined assignment is not class-functional, so the obstruction does not apply to it. It does not restore a withdrawn stop hypothesis.

No axiom is added. The Identification Hypothesis is not used. The dual cipher of A_8 is recorded as a complementary experiment and is not run.

Keywords: Chrysene Formalism; sufficiency; ancillarity; transversality; seed map; normal form; mitochondrial code; parity obstruction; dual cipher.

1 Introduction

The seeded-assignment construction of [6] produces, from a chrysene cluster and an adenine-tethered complex, an assignment of residues to codon words. The output, under four named hypotheses, is the vertebrate mitochondrial table. That paper works in the language of seeds, demand, propagation, and audit. The present paper reads the same construction as a pair of reductions on a finite statistical experiment, and asks two questions that [6] does not ask.

The first: in what sense is the frame class a sufficient statistic for the residue formed at the cluster? The second: in what sense does that sufficiency fail to determine a codon table?

The answers are not interchangeable. Excision is a function of the class, by definition of excision. The class statistic

is therefore a sufficient reduction of the word for excision. In the joint experiment that records the word together with the residue, the pair (T, r) is minimal Neyman-sufficient for the unknown excision map, and among statistics that are functions of the word alone, T is a minimal sufficient reduction. The within-class coordinate is conditionally ancillary for that map. Those statements are Section 5.

A codon table is not a function of the class. The sixty-four words carry two partitions — frame classes, twenty of them, and codon blocks, sixteen of them — and the partitions are transverse: a class meets a block in at most one word. The class does not determine the block. Assignments measurable with respect to both partitions are constant on each of the two incidence components and take at most two values. Transporting a residue from a class onto a block therefore requires a choice the class does not make.

That choice is a seed, a section of the incidence graph. Propagation is the unique block-measurable extension of a non-conflicting seed. Those statements are Sections 6 and 7.

Refinement at the third position, pyrimidine against purine, then produces only two block patterns, (4) and (2, 2). The standard table realises also (3, 1) and (2, 1, 1), and lies outside the range. The vertebrate mitochondrial table realises only the two permitted patterns. Under the four matching hypotheses already named in [6], and with chain polarity required only to identify the domain with a sixty-four-codon table, one refined assignment equals that table. Parity of class-functional maps excludes the standard table, because methionine and tryptophan are singleton fibres, and is silent on the mitochondrial table. The refined assignment is not class-functional, so the obstruction does not apply to it. Those statements are Sections 8 and 10.

The paper introduces no axiom. It uses the twelve-axiom carrier of [1] and the encoded stack of [2, 3] exactly as [6] used them. The axioms that supply a tensor network, a crossed-product algebra, and a spectral triple remain unused in the matching, the sufficiency theorems, and the audit. They are available as ambient geometry of a complementary experiment, recorded in Section 9 and not run.

Two declarations are in force throughout. The standard table is used constructively: the demand graph is built from it. That is the opposite convention of [5], and is stated at Declaration 2.7 so that it can be checked. The Identification Hypothesis Φ of [7] is not used. An empty audit is an equality of two functions on a sixty-four-element set. It is not an identification of the excised residues with biological amino acids.

The logical vocabulary, the letter assignment, the imported objects, and the register of hypotheses occupy Section 2. The implications the constructions do not licence occupy Section 3, each with a witness. Targets left open — chain polarity, the chemical rule on the purine half, the environmental pair, cross-ancillarity, the commutant reading, and Φ — are listed in Section 11.

2 Conventions, Imported Objects, and the Register

This section is definitional. It fixes the logical vocabulary, the letter assignment, the objects imported from the twelve-axiom foundation and from the seeded-assignment construction, and the statistical experiment in which those objects will be read. No existence result, matching theorem, or audit is proved here. Later sections may cite the definitions below and may not alter them.

2.1 Logical status

Definition 2.1 (Status labels). Imported from [1] and used as in [6]:

- [A] axiomatic, or a finite enumeration labelled as such;
- [D] derived realization theorem;
- [T] target;
- [E] empirical instantiation claim.

An [E] claim is confined to a remark and enters no proof. A statement $P_{\mathfrak{R}}$ is a derived realization theorem when

$$\mathcal{A} \cup H_{\mathfrak{R}} \models P_{\mathfrak{R}} \quad (1)$$

for an explicitly named hypothesis family $H_{\mathfrak{R}}$, where $\mathcal{A} = \{A_1, \dots, A_{12}\}$ is the axiom set of [1]. If $H_{\mathfrak{R}}$ is used, one has $\mathcal{A} \not\models P_{\mathfrak{R}}$ unless a separate argument shows otherwise. The governing restraint of [7] remains in force: restoring a foundational object never licenses restoring an unproved implication of that object.

Remark 2.2 (No new axiom). [A] This paper introduces no axiom. Every structure below is imported from [1, 2, 3, 4, 5, 6], is definitional within $\mathcal{A}$, or is a realization hypothesis registered at Section 2.6. In particular A_3 , A_4 , A_5 , A_7 , A_9 , A_{10} and A_{11} are unused in the matching, sufficiency, and audit arguments of Sections 5–8. Where they appear later they supply ambient geometry of the experiment — a metric on assignments, a commutant representation, a physical motif space — and are not ingredients of those arguments.

Remark 2.3 (Two counting conventions). [A] Statements about a finite object — a set of words, a group acting on it, a published codon table — are proved by enumeration and labelled [A]. Each such statement is instantiated once, by a corollary labelled [D], as a statement about clusters under the hypotheses named there. The two are kept apart, as in [2, 6].

2.2 Provenance

Proposition 2.4 (Provenance). [A] *The site set and the stacking direction are supplied by A_6 ; the two attachment sets, hence the species alternation, by A_1 ; the non-collapse of the marker by A_2 ; the oriented perimeter cycle by A_8 ; and the constraint functionals by A_{12} . No constituent of the carrier used below is supplied by A_3 , A_4 , A_5 , A_7 , A_9 , A_{10} or A_{11} .*

Remark 2.5. A_7 being unused in the combinatorial arguments, nothing in Sections 5–8 concerns the order dependence of an encoded word, and [2, Thm. 4.17] is not invoked. Sequence-induced noncommutativity of $\mathfrak{A}_{\mathfrak{C}} = \mathfrak{Q}_{\mathfrak{C}} \rtimes_{\alpha} \mathbb{Z}^3$ is available as ambient structure from Section 9 onward and is not presupposed before then.

2.3 The assignment

Definition 2.6 (Letter assignment). Throughout,

$$FF = A, \quad FB = C, \quad BF = G, \quad BB = U, \quad (2)$$

so that

$$\text{reg}^{-1}(0) = \{A, U\}, \quad \text{reg}^{-1}(1) = \{C, G\} : \quad (3)$$

the registry partition is carried to the Watson–Crick partition. Write $\Sigma := \{A, C, G, U\}$ for the resulting four-letter alphabet, and $W := \Sigma^3$ for its word space of length three. Thus $|W| = 64$.

Remark 2.7 (A re-choice, not a relabelling). [4, Def. 2.16] adopts an assignment carrying the registry partition to the amino/keto partition. Definition 2.6 does not. The eight assignments compatible with a given bipartition of Σ differ by an element of a copy of $\mathbb{Z}/2\mathbb{Z}$ acting on the source, so all orbit counts are unchanged — but the codon carried by each class changes, and the coding table and both block operators of [4] are rebuilt under the present assignment. This is a re-choice of the datum that [3, Target 7.7] leaves open, and is recorded as one, not as a change of notation. Two reasons for the Watson–Crick partition are given at [6, Rem. 4.16]: empirical stability of the pairing, and the unique partition under which base complementation coincides with a frame operation [6, Prop. 4.14]. Relabelling U as T carries (3) to the same partition of the same four-element set; the classification is independent of that labelling [5, Prop. 4.11]. Codon tables of [8] are published with T ; U is used here and the two are interchangeable throughout the enumerations.

2.4 Imported objects

The following objects are used as already defined. No clause of their definitions is reopened.

Object	Source
marker; non-collapse	[2, Def. 3.3]
vessel; binaries b_1, b_2 ; $\Sigma_{\mathfrak{b}}$	[2, Defs. 3.]
structural group; ρ, τ and their fixed points	[2, Props. 4.]
frame; frame-setters; torsor structure	[2, Defs. 6.]
decomposition of the frame action	[2, Prop. 6.]
initiator	[2, Def. 9.1]
frame group $\text{Fr} = \langle \text{flip}, \text{rev} \rangle$	[3, Def. 2.1]
structured alphabet; count over its word space	[3, Def. 4.1]
class sizes two and four, profile eight and twelve	[3, Cor. 3.7]
transversality of blocks and classes	[4, Rem. 6.]
capacity; parity; production clauses	[5, Thm. 5.]
seed map; demand; propagation; refinement	[6, Defs. 5.]
residue volumes	[9]
translation tables	[8]

Definition 2.8 (Frame classes). Write X for the set of orbits of Fr acting on W . By [3, Lem. 4.5, Cor. 3.7], under Definition 2.6,

$$|X| = 20, \quad (4)$$

of which eight orbits have size two and twelve have size four. Elements of X are called *frame classes*, or simply classes. The orbit map is denoted

$$T : W \rightarrow X, \quad T(w) = \text{Fr} \cdot w. \quad (5)$$

Definition 2.9 (Blocks). A *block* is a set of four words of W that agree in their first two letters. Write $\mathcal{B}$ for the set of blocks. Then $|\mathcal{B}| = 16$, and the block map

$$B : W \rightarrow \mathcal{B} \quad (6)$$

sends each word to the unique block containing it.

The two partitions of W — classes and blocks — are the data on which everything after Section 4 turns. Their relation is not recorded here. It is imported as [6, Prop. 5.2] and restated when first used.

2.5 The experiment

The constructions of [6] assign residues to words. The present paper reads that assignment as a decision rule on a finite statistical experiment. The experiment is defined now; sufficiency and ancillarity are proved in Section 5.

Definition 2.10 (Finite experiment). A *finite experiment* is a triple $\mathcal{E} = (W, \mathcal{F}, \{P_{\theta, \eta}\})$ in which W is a finite set, $\mathcal{F} = 2^W$, the parameter of interest θ ranges over a finite set Θ , the nuisance parameter η ranges over a (possibly dummy) set H , and each $P_{\theta, \eta}$ is a probability measure on W . A *statistic* is a map from W to a finite set. Throughout, unless a different measure is named, $P_{\theta, \eta}$ is the normalized counting measure on W , which does not depend on (θ, η) . Dependence on the parameters enters through the decision maps of Definition 2.14, not through a family of sampling laws on words.

Remark 2.11 (Why the sampling law is flat). The word space W is the domain of a codon table, not a population of observed transcripts. The paper makes no sampling claim about genomic frequencies. The counting measure is the unique Fr-invariant probability on W , and is used because every enumeration in [3, 6] is an enumeration of that set. Replacing it by another Fr-invariant measure does not change the σ -algebras generated by T , A and B , hence does not change the sufficiency statements of Section 5.

Definition 2.12 (Relative frame). Let F denote a set of representatives of the frame-setters of [2, Defs. 6.1, 6.2], so that each word $w \in W$ determines a unique pair $(T(w), A(w))$ with $A(w) \in F$ once a global frame has been fixed. The map

$$A : W \rightarrow F \quad (7)$$

is the *relative-frame statistic*. Under H_{teth} (i)–(iii) the poly-A tether fixes one frame for all clusters along it [6, Rem. 4.13], and

$$8 \cdot 2 + 12 \cdot 4 = 64 = |W| \quad (8)$$

counts the pairs (class, relative frame) [6, Cor. 4.12]. Without clause (iii) of H_{teth} the residual count is thirty-six and no map onto a sixty-four-codon table exists. That clause is Target 9.7 of [6] and is load-bearing for the identification of W with the domain of a codon table. It is not discharged here.

Definition 2.13 (Parameter of interest and nuisance). Let R be a finite residue set with $|R| \leq 20$, and let $*$ be a formal symbol disjoint from R . The *parameter of interest* is the unknown excision map

$$\Xi : X \rightarrow R \quad (9)$$

of Definition 2.17, not the realized value $\Xi(T(w))$. The *nuisance parameter* η ranges over a set E of translational environments, stacking-registry configurations, and third-base engagement data. Liberation ϖ of Definition 2.17 may depend on η ; excision does not.

Definition 2.14 (Decision maps). An *assignment* is a map

$$c : W \rightarrow R \cup \{*\}. \quad (10)$$

Write $\mathcal{C}_{\text{std}}$ for the standard genetic code of [8], table 1, and $\mathcal{C}_{\text{mt}}$ for the vertebrate mitochondrial code of [8], table 2, both read as assignments $W \rightarrow R \cup \{*\}$ under the identification $T \leftrightarrow U$. The value $*$ is an absence, not a twenty-first residue: it records that liberation does not proceed, not that a distinct product is formed. No assignment is constructed in this section.

Remark 2.15 (Convention on $*$ and published stops). Two readings of $*$ are used, and they are not interchangeable.

- (i) *Production*. $*$ means that liberation fails. It is not an element of R .
- (ii) *Audit and published patterns*. The published stop symbol of [8] is identified with $*$ when a table is read as an assignment $W \rightarrow R \cup \{*\}$, and when the pattern of a published block is named. Under that reading, a block that carries one residue and two published stops has pattern $(2, 2)$ in the pair {residue, $*$ }.

Clause (i) is used in the production arguments of Sections 5 and 10. Clause (ii) is used in the range comparison and the audit of Section 8. An empty audit is then an equality of functions into $R \cup \{*\}$, not a claim that a stop is a residue.

Definition 2.16 (Generated σ -algebras). Write $\sigma(T)$, $\sigma(A)$, $\sigma(B)$ and $\sigma(c)$ for the partitions of W generated by the indicated maps, identified with the corresponding finite σ -algebras. Inclusions and failures of inclusion among these partitions are statements about functions on a sixty-four-element set and are proved by exhibition.

2.6 The register

Definition 2.17 (Excision and liberation). Ξ denotes the *excision map*, assigning to each frame class the

residue formed at that cluster:

$$\Xi : X \rightarrow R. \quad (11)$$

ϖ denotes the *liberation predicate*, deciding whether the residue formed is released:

$$\varpi : W \times E \rightarrow \{0, 1\}. \quad (12)$$

The composite of excision with liberation is the production of the cluster. An assignment c compatible with a pair (Ξ, ϖ) satisfies

$$c(w) = \begin{cases} \Xi(T(w)) & \text{if } \varpi(w, \eta) = 1 \text{ and the residue is placed} \\ * & \text{if } \varpi(w, \eta) = 0, \end{cases} \quad (13)$$

the middle clause “placed at w ” being supplied by the seed, propagation and refinement of [6] and not by Ξ alone. Equation (13) is a compatibility condition, not a construction.

Remark 2.18 (Kernels, not only functions). Nothing below forbids viewing Ξ as a Markov kernel $X \rightarrow \text{Prob}(R)$ or ϖ as a kernel $W \times E \rightarrow \text{Prob}(\{0, 1\})$. The register of [6] takes both as deterministic, and the sufficiency argument of Section 5 uses only that Ξ is constant on classes. Determinism of ϖ is not used in any [A] or [D] statement.

Hypothesis 2.19 (H_{teth}). In four clauses, each load-bearing, imported unchanged from [6, Hyp. 2.9].

- (i) The tether is a constant poly-A strand, one adenine–chrysene unit per coding unit.
- (ii) Its register with the coding strand is non-Watson–Crick, so that the tether is constant irrespective of the codon; and the intervening catalyst of the original cluster is displaced and enters no registry.
- (iii) The tether carries a chemical chain polarity.
- (iv) It is attached at the same ring and in the same orientation at every unit.

Hypothesis 2.20 (H_{seed}). Imported from [6, Hyp. 2.10]. Excision seeds a block at the codon it shares with the seeding class, on the pyrimidine half of a block whose halves carry distinct residues, and the seed propagates within the block.

Hypothesis 2.21 ($H_{\text{split-pos}}$). Imported from [6, Hyp. 2.11]. A block carrying two residues is divided

at the third position, pyrimidine against purine.

Hypothesis 2.22 ($H_{\text{split-chem}}$). Imported from [6, Hyp. 2.12]. The residue on the purine half is determined by the reaction vessel — hydrogen bonding, steric hindrance and electrostatics together — and not by size alone.

Hypothesis 2.23 ($H_{\text{lib-env}}$). Imported from [6, Hyp. 2.13]. Liberation depends on the translational environment as well as on the cluster and the frame.

Remark 2.24 (Why the split is two hypotheses). $H_{\text{split-pos}}$ is proved as a statement about published tables at [6, Thm. 6.2] and will be reused; $H_{\text{split-chem}}$ is not proved anywhere in the corpus, and [6, Cor. 6.7] refutes the first account one would try for it. A single hypothesis covering both would conceal that asymmetry.

Remark 2.25 (Imported hypotheses, and one not carried). H_{mk} , H_{stack} , H_{frm} and H_{ind} are imported unchanged from [2, Hyps. 2.6–2.11], and $H_{\text{prod}}(0)$, $(0')$, (1) , (2) from [5, §2.5]. H_{stop} of [4, Hyp. 2.22] is not carried: two of the three designations it makes are consequences of [6, Cor. 6.4], and the third is treated as an environmental decision under $H_{\text{lib-env}}$. Withdrawing a hypothesis is recorded because the ledgers of this corpus have more often added them.

2.7 Declaration: the standard code is used constructively

Declaration. [5, Rem. 2.4] declares that the standard genetic code enters no hypothesis and no construction of that paper, and lists the hypotheses deliberately absent so that the claim can be checked. This paper is the opposite case, as is [6], and the declaration is owed in the same form.

The demand graph of the seed construction, imported from [6, §5], is built from $\mathcal{C}_{\text{std}}$; the refinement is verified against $\mathcal{C}_{\text{std}}$; and the audit of published tables is a statement about [8]. The sufficiency and ancillarity statements of Section 5, and the identification of W as a word space in Section 4, do not depend on that datum. A reader arriving from [5] must not carry its independence claim forward: it is not inherited, and nothing in this paper re-establishes it.

Remark 2.26 (The identification hypothesis is not used). Thesis B of [7] is conditional on an Identification Hypothesis Φ relating the formal graph of $\mathcal{A}$ to biological reality. Φ is not an ingredient of any statement labelled [A] or [D] in this paper. Agreement of a constructed assignment with a published translation table is a comparison of two functions on a sixty-four-element set. It is not a proof that the residues produced at a chrysene cluster are biological amino acids, and it is not offered as one.

2.8 Naming

Remark 2.27 (Three substitutions and two collisions). Excision and liberation name the two chemical operations. Excision is a term of art in homology theory; neither this paper nor [1] uses homology, so the collision is nominal and is recorded rather than avoided, as Δ against Δ_k was recorded at [5, Rem. 2.23].

ϖ , not ρ , denotes the liberation predicate: ρ is the fixed-point-free element of the structural group at [2, Prop. 4.7] and is reserved for that use. Ξ is free across the corpus.

Block is reserved for the four words agreeing in their first two letters (Definition 2.9) and *class* for a frame orbit (Definition 2.8). The two partitions are transverse and are never interchanged.

Sufficient and *ancillary* are used only in the sense of Definition 2.8 below, as properties of statistics on the experiment of Definition 2.10. They are not synonyms for “adequate” or “secondary,” and they are not applied to chemical reagents.

Definition (Sufficiency and ancillarity; forward reference). The formal definitions are given in Section 5. Informally, and only to fix usage: a statistic S is sufficient for θ when the conditional law of the observation given S does not depend on θ ; a statistic S is ancillary for θ when the marginal law of S does not depend on θ . Completeness, minimal sufficiency, and Basu independence are likewise defined there. No property so named is asserted in the present section.

2.9 Notation

Symbol	Meaning
$\mathcal{A}$	the twelve-axiom set $\{A_1, \dots, A_{12}\}$ of [1]
$\mathfrak{C}$	embedded labeled molecular node of A_1
$\Psi_{\mathfrak{C}}$	local Chrysene structural tensor
$\Omega_{\mathfrak{C}}, \mathfrak{A}_{\mathfrak{C}}$	quasi-local algebra; crossed product $\Omega_{\mathfrak{C}} \rtimes_{\alpha}$
Ξ	excision; a map on the twenty classes
ϖ	liberation; a predicate on words and environments
X	the class set; $ X = 20$
Fr	the frame group $\langle \text{flip}, \text{rev} \rangle$
ρ, τ	structural-group elements of [2, Prop. 4.7]
T, A, B	class, relative-frame, and block statistics
c	an assignment $W \rightarrow R \cup \{*\}$
$\mathcal{C}_{\text{std}}, \mathcal{C}_{\text{mt}}$	standard and vertebrate mitochondrial assignments
$R Y, M K, W S$	purine/pyrimidine, amino/keto, Watson-Crick
pattern (4), (2, 2), ...	the multiset of residue multiplicities in a block
$*$	absence of liberation; not a residue

Remark 2.28 (What this section does not fix). This section does not select a seed map among the eight hundred and forty that achieve coverage [6, Prop. 5.14]; does not discharge $H_{\text{teth}}(\text{iii})$, $H_{\text{split-chem}}$, or $H_{\text{lib-env}}$; does not identify Φ ; and does not choose a Dirac operator or a path metric. Those choices, where they are made at all, are made at the point of use and under an explicit label.

3 Traps and Non-Implications

The statements below are the implications that the constructions do not licence. Each clause names a witness already proved. The section is placed after the conventions and before the experiment so that later arguments can cite it; it was written after Sections 4–10, because the witnesses live there.

A restoration of any clause would be a new claim and would have to be labelled as one. None is restored.

3.1 Non-implications

Proposition 3.1 (Non-implications). *[A] None of the following implications holds in this paper.*

- (i) $|\Sigma| = 4$ does not equip Σ with the field operations of $\mathbb{F}_4$. Witness: Proposition 9.12.
- (ii) Sufficiency of T for excision does not give sufficiency of T for the block. Witness: Corollary 6.5.

- (iii) Sufficiency of T for excision does not give sufficiency of T for the assignment. Witness: Theorem 6.11.
- (iv) T -measurability of an assignment does not forbid variation inside a block. Witness: Remark 6.10.
- (v) Class-extension of a non-constant Ξ is not a block-organised assignment. Witness: Corollary 6.9.
- (vi) Joint T - and B -measurability does not admit an assignment of three or more values; it does admit one value per incidence component. Witness: Theorem 6.8.
- (vii) The raw frame coordinate A is not independent of T . Witness: Proposition 5.14.
- (viii) Basu's theorem does not supply that independence. Witness: Remark 5.15.
- (ix) Liberation need not be T -measurable. Witness: Proposition 5.19.
- (x) Hall's criterion on incidence does not give Hall's criterion on demand. Witness: Proposition 7.9.
- (xi) A covering seed does not assign values on unoccupied blocks. Witness: Proposition 7.20.
- (xii) Propagation does not produce a split. Witness: Remark 7.21.
- (xiii) Conditional expectation of a seeded value is not propagation. Witness: Remark 7.15.
- (xiv) The standard table is not in the range of refinement. Witness: Corollary 8.9.
- (xv) An empty audit is not an identification of residues with biological amino acids, and is not Φ . Witness: Remarks 2.26, 8.22.
- (xvi) Residue volume does not discharge $H_{\text{split-chem}}$. Witness: Remark 8.18.
- (xvii) Capacity does not force an odd fibre and does not exclude the standard table. Witness: Remark 10.4.
- (xviii) Hall does not cancel parity. Witness: Proposition 10.14.
- (xix) Sufficiency of T for excision does not apply parity to the seeded assignment. Witness: Theorem 10.15.
- (xx) Parity does not exclude the mitochondrial table, and does not say that no genetic code can be class-functional. Witness: Proposition 10.9, Remark 10.10.

- (xxi) Parity is not the evenness of $|\text{Fr}|$, and does not restore H_{stop} . Witness: Remarks 10.19, 10.18.
- (xxii) The dual cipher is not an element of Fr and is not reverse complementation. Witness: Propositions 9.3, 4.19.
- (xxiii) The complementary experiment does not determine a second codon table. Witness: Proposition 9.13.
- (xxiv) Complementarity of the two ciphers is not geometric orthogonality and is not a Connes distance. Witness: Proposition 9.14.
- (xxv) Axiom A_8 is not used in the matching, the sufficiency theorems, or the audit. Witness: Proposition 2.4, Remark 2.2.

3.2 Traps

Remark 3.2 (Trap: “sufficient”). Sufficient, in this paper, is a property of a statistic for a named parameter on a named experiment (Definitions 5.1, 5.4). It is not a synonym of “adequate,” “enough,” or “chemically complete.” The sentence “the class is sufficient” is not a sentence of the paper. The sentence “the class is a sufficient reduction of the word for excision” is Theorem 5.7.

Remark 3.3 (Trap: “the frame”). Two objects were named frame and then separated (Remark 5.17). The poly- A tether is a constant reagent under H_{teth} (i)–(ii). The within-class coordinate Φ_s is the position of the word in its orbit. Saying that “the frame is ancillary” without saying which is Remark 5.17 used as a trap.

Remark 3.4 (Trap: “parity”). Parity names the evenness of fibres of class-functional maps (Theorem 10.6). It does not name the even order of Fr , an absence number on a block, or a stop codon. The even size of classes is the input to the theorem, not a second theorem.

Remark 3.5 (Trap: “the dual”). The dual cipher is a second four-site reading of the perimeter (A_8). Strand complementarity is the letterwise map ρ inside one alphabet. Reverse complementation is a map on words that is never a frame operation. Three duals; three statements; Proposition 9.3 and Remark 9.15.

Remark 3.6 (Trap: “the code”). Demand is read from $\mathcal{C}_{\text{std}}$. The output comparison is with $\mathcal{C}_{\text{int}}$. The range of refinement excludes $\mathcal{C}_{\text{std}}$ (Corollary 8.9). Using “the genetic code” for all three collapses a distinction the audit is there to keep (Remark 8.14).

Remark 3.7 (Trap: “equals the table”). Theorem 8.13 is an equality of two functions on a sixty-four-element set, under four hypotheses. It is not a derivation of the published table from $\mathcal{A}$, and it is not a chemical mechanism. The hypotheses remain listed at Remark 8.23.

Remark 3.8 (Trap: transporting the class count). Theorem 4.6 is [3, Lem. 4.5] applied to a structured alphabet. The carrier is that word space (Remark 4.7). An argument that first counted orbits of a marker presentation and then “transported” the number twenty to the nucleobase alphabet is one of the superseded architectures of Remark 4.22.

3.3 What would count as restoration

Remark 3.9 (Restoration). A later paper that claims any clause of Proposition 3.1 as a theorem must exhibit a proof from $\mathcal{A}$ and the hypotheses it names, or must add a hypothesis and say so. In particular:

- identifying Σ with $\mathbb{F}_4$ is a new datum;
- discharging $H_{\text{split-chem}}$ by a rule other than size is a new argument (size is already closed);
- naming three stop codons from $\mathcal{A}$ is a restoration of H_{stop} ;
- running the complementary experiment through Hall and an audit is a new paper;
- asserting Φ is an identification hypothesis, not a corollary of an empty audit.

The governing restraint of [7] applies to each item.

4 The Cluster as a Statistical Experiment

Section 2.5 fixed the sample space $W = \Sigma^3$ and the statistics T , A and B as maps on that set. This section identifies W with the word space of the composite stack, records what the tether fixes, and records two facts about the Watson–Crick partition that the later sufficiency statements will use. No seed map is chosen here, and no assignment is constructed.

The carrier is the word space of a structured alphabet [3, Def. 4.1, Lem. 4.5]. The class count is that lemma applied,

not a result transported from the marker presentation to a new one [6, Rem. 4.6].

4.1 The composite stack

Definition 4.1 (Composite stack). Imported from [6, Def. 4.1]. After the first transformation the coding strand carries three nucleotide–chrysene units and the poly-A strand three adenine–chrysene units. The two are paired in the register of H_{teth} (ii), and the A-tethered chrysene is interleaved between successive nucleotide pairings. The resulting stack has six units and three vessels, the A-tethered chrysenes serving as the catalyst units and the nucleotide-bearing chrysenes as the substrate units.

Proposition 4.2 (It is a stack). [D] Under H_{teth} , the composite of Definition 4.1 alternates in species in the sense of [2, Def. 3.1], the interleaved A-tethered chrysenes supplying the catalyst sites.

Proof. [2, Def. 3.1] requires species alternation along the stack. Clause (ii) of H_{teth} places the A-tethered units in the catalyst positions of the original cluster and displaces the intervening catalyst of that cluster, so that the six-unit composite alternates substrate and catalyst. This is [6, Prop. 4.2]. $\square$

Remark 4.3 (The displaced catalyst). Clause (ii) of H_{teth} is doing work and is not an aside: it is why the composite has six units rather than more, and why no registry entry arises between a tether unit and a coding unit. An architecture placing the tether at additional sites rather than complexing it within one refines the twenty classes into thirty-six; see Remark 4.22.

4.2 The carrier is a word space

Definition 4.4 (Structured alphabet). A four-element set equipped with a distinguished bipartition into two pairs is a *structured alphabet* in the sense of [3, Def. 4.1]. Under Definition 2.6 the set $\Sigma = \{A, C, G, U\}$ with the partition $\{A, U\} \mid \{C, G\}$ is such an alphabet.

Proposition 4.5 (The alphabet is structured). [D] Definition 2.6 makes $(\Sigma, \{A, U\} \mid \{C, G\})$ a structured alphabet. The word space of length three is the set W of Definition 2.6.

Proof. Immediate from Definitions 2.6 and 4.4. $\square$

Theorem 4.6 (Twenty classes). *[A] The word space of a structured alphabet at length N carries $2^{2N-2} + 2^{N-1}$ frame classes. At $N = 3$ there are twenty, of sizes two and four in the profile eight and twelve.*

Proof. [3, Lem. 4.5, Cor. 3.7]. The statement is imported and not reproved. It is regenerated under the present assignment as check V1 of Appendix A. Definition 2.8 is that count. $\square$

Remark 4.7 (No transport argument is needed). The carrier is a word space over a structured alphabet, which is the object [3, Lem. 4.5] is about. Theorem 4.6 is therefore that lemma applied, not a result transported from the marker presentation to a new one. Earlier plans that required such a transport are recorded at Remark 4.22.

Remark 4.8 (U and T). Relabelling U as T carries $\{A, U\} \mid \{C, G\}$ to $\{A, T\} \mid \{C, G\}$, the same partition of the same four-element set. By [5, Prop. 4.11] the classification is independent of the labelling. The formalism is indifferent to the substitution; the chemistry is not, thymine being 5-methyluracil, and no claim is made either way. This repeats Remark 2.7 only so that the present section is self-contained on the point.

4.3 The tether and the frame

Proposition 4.9 (Why the tether can be constant). *[D] Under the non-Watson-Crick register of H_{teth} (ii) the tether strand is independent of the codon. Under a Watson-Crick register it would be determined by the codon and would distinguish no frame.*

Proof. A Watson-Crick pairing would send each letter of the coding word to its complement on the tether, so the tether word would be a function of the codon. The register of H_{teth} (ii) is not of that form: the tether is the constant word AAA, irrespective of the codon. This is [6, Prop. 4.8]. $\square$

Proposition 4.10 (The tether is palindromic). *[A] As a word the tether is AAA. Since τ fixes the aligned letters and A is aligned under Definition 2.6, $\text{rev}(AAA) = AAA$: the tether is fixed by reversal.*

Proof. By [2, Prop. 4.8], τ fixes exactly the aligned letters of the assignment. Under (2) the letter A is aligned, so each factor of AAA is fixed by τ , hence by the letterwise action of rev . $\square$

Corollary 4.11 (What each datum fixes). *[A] The states remaining when the tether fixes:*

<i>Tether fixes</i>	<i>Residual group</i>	<i>States</i>
<i>nothing</i>	Fr	20
<i>the projection reference only</i>	$\langle \text{rev} \rangle$	36
<i>the reading direction only</i>	$\langle \text{flip} \rangle$	32
<i>both</i>	<i>trivial</i>	64

By Proposition 4.10 the tether's pattern fixes the projection reference and cannot fix the reading direction, so its pattern alone leaves thirty-six.

Proof. Orbit-counting under the four subgroups of $\text{Fr} \cong \mathbb{Z}/2 \times \mathbb{Z}/2$. The four rows are the class count of Theorem 4.6, the count under the index-two subgroup generated by reversal, the count under the index-two subgroup generated by flip, and the count of words. This is [6, Cor. 4.10]. $\square$

Remark 4.12 (Chain polarity is load-bearing). Clause (iii) of H_{teth} supplies what the pattern cannot: a poly-A chain carries a chemical polarity the formalism does not represent, and the reading direction is fixed by the chain. With (i), (ii) and (iii) together the frame is determined. Without (iii) the state count is thirty-six and no map onto a sixty-four-codon table exists, so the clause is as load-bearing as the constancy and is registered as such. It remains Target 9.7 of [6] and is not discharged here.

Corollary 4.13 (Sixty-four). *[D] With the frame fixed, the pairs (class, relative frame) number $8 \cdot 2 + 12 \cdot 4 = 64$, which is the domain of a codon table and is the set W of Definition 2.6.*

Proof. Theorem 4.6 gives the orbit profile eight and twelve. A size-two class has two relative frames; a size-four class has four. The product is $|W|$. This is [6, Cor. 4.12], and is the content of (8). $\square$

Remark 4.14 (One tether, one frame, one code). A single poly-A strand serves every cluster along it, so the

section it supplies is global: one frame for all clusters, hence one codon table and one handedness across all codons. A per-cluster reference would give per-cluster frames and nothing would require the resulting tables to agree. The statistic A of Definition 2.12 is therefore a single coordinate on W , not a family of coordinates indexed by clusters.

4.4 Identification of the experiment

The experiment of Definition 2.10 is now identified with the composite stack.

Definition 4.15 (Identified experiment). The sample space of $\mathcal{E}$ is the word space W of the structured alphabet of Proposition 4.5. The observation is the codon word $w \in W$, equivalently the pair $(T(w), A(w))$ of Corollary 4.13. The parameter of interest is the unknown excision map Ξ of Definition 2.13, not the realized value $\Xi(T(w))$. The nuisance parameter η ranges over translational environments, stacking-registry values of A_6 , and third-base engagement data, as required by $H_{\text{lib-env}}$ and $H_{\text{split-chem}}$. The sampling law remains the normalized counting measure on W (Remark 2.11).

Proposition 4.16 (Coordinates). [A] *The pair of statistics (T, A) separates points of W : the map*

$$W \rightarrow X \times F, \quad w \mapsto (T(w), A(w)) \quad (14)$$

is bijective once a global frame has been fixed. The block statistic B is a function of w and is not a function of T alone.

Proof. Bijectivity is Corollary 4.13. That B is not a function of T is the content of transversality [6, Prop. 5.2]: a class meets a block in at most one word, so the class does not determine the block. The comparison of $\sigma(T)$ and $\sigma(B)$ is deferred to Section 6. $\square$

Remark 4.17 (What the identification does not claim). Definition 4.15 identifies two finite sets and two maps between them. It does not identify the residues in the image of Ξ with biological amino acids, and it does not discharge Φ (Remark 2.26). The identification is of the domain of a codon table with the word space of the stack, which is a statement about $|W|$ and the action of Fr.

4.5 Why Watson–Crick

The partition $\{A, U\} \mid \{C, G\}$ is used for two independent reasons. Only the second is proved here.

Proposition 4.18 (Complementation is a frame operation only under Watson–Crick). [A] *Base complementation is fixed-point-free, and ρ is the unique fixed-point-free element of the structural group of [2, Prop. 4.7]. Hence complementation coincides with flip under exactly the eight assignments carrying the registry partition to $\{A, U\} \mid \{C, G\}$, and under none of the other sixteen.*

Proof. A fixed-point-free involution of a four-element set is a pairing of the four letters. The structural group of [2, Prop. 4.7] has a unique fixed-point-free element ρ . The eight assignments that carry the registry partition to Watson–Crick send ρ to base complementation; the remaining sixteen send it to a different pairing. This is [6, Prop. 4.14]. $\square$

Proposition 4.19 (Reverse complementation never is). [A] *The reversing elements of the frame group are rev and flip $\circ$ rev, acting letterwise as τ and $\tau\rho$, each of which fixes exactly two letters by [2, Prop. 4.8]. A fixed-point-free map equals neither. Hence the biological reverse complement is a frame operation under no assignment.*

Proof. [6, Prop. 4.15]. The two reversing elements each have two fixed letters, so neither can equal a pairing of all four letters. $\square$

Remark 4.20 (Two reasons, and one consequence). The Watson–Crick partition is adopted for stability, which is empirical, and because it is the unique partition under which complementation is a symmetry of the classification at all, which is Proposition 4.18 and is not. The corpus has no other argument for a partition on symmetry grounds. Proposition 4.19 has a consequence worth stating. Strand complementarity is available to the formalism and strand reading direction is not, so a DNA:RNA duplex is an external datum — like the initiator and like the tether — and not a symmetry. It can break the frame; it cannot halve the registry.

Remark 4.21 (Ancillarity preview). Proposition 4.9 is the reason the relative-frame statistic can be ancillary for excision. If the tether were Watson–Crick, the

tether word would be a function of w and would carry information about the codon; it would then fail to be ancillary for any parameter that depends on the codon. The non-Watson–Crick register of H_{teth} (ii) is what keeps A available as an ancillary coordinate. The formal statement is Theorem 5.2 below; the present remark only records which datum of this section that theorem will use.

4.6 Superseded architectures

Remark 4.22 (Two models raised and abandoned). Two earlier carriers were considered in the development of [6] and are recorded because each was abandoned for a reason that constrains the chemistry.

An architecture placing the tether at its own sites in the stack creates registry entries between tether and coding unit of the form $m_j + c$. These are invariant under flip and are genuinely new information, and the count refines from twenty to thirty-six. Complexing the tether within a site creates no such entry. A constant is not informationless when the thing beside it varies; it is informationless when it is complexed into the same site.

An architecture treating the composite as a binary marker alphabet raised three further questions — whether the composite alternates in species, whether a composite site carries a marker at all, and how many tether units pair with how many coding units — all of which dissolve once the carrier is recognised as a word space over the nucleobase alphabet, which is what the chemistry supplies. These are recorded as superseded rather than repaired: they were consequences of a model the chemistry does not use.

Neither abandoned architecture is available to the experiment of Definition 4.15. In particular the sample space has sixty-four points, not thirty-six, and the class statistic T takes values in a twenty-element set, not a thirty-six-element set.

Remark 4.23 (What this section has fixed). The sample space is W , of cardinality sixty-four, identified with the word space of the composite stack under H_{teth} . The class statistic T takes twenty values. The relative-frame statistic A is a global coordinate, fixed by the tether’s pattern together with chain polarity. Base complementation is a frame operation under the present assignment and the biological reverse complement is not. No assignment c has been chosen, and $\sigma(T)$ has not yet been compared with $\sigma(B)$ or with $\sigma(c)$.

5 Sufficiency of the Class and Ancillarity of the Frame

Excision is a function of the class. That fact is Definition 2.17, not a theorem. The work of this section is to say precisely in what sense the class is a sufficient reduction of the word for excision, in what sense the relative frame is ancillary for excision, and in what sense those statements fail if they are written carelessly.

Two pitfalls are recorded first, because both would make the later theorems look stronger than they are. Under the counting measure of Remark 2.11 the law of w does not depend on Ξ . Every statistic is then Neyman-sufficient for Ξ and every statistic is ancillary for Ξ , and nothing has been said. Second, the raw relative-frame coordinate on a four-element set is not independent of the class, because eight classes have size two and twelve have size four. Basu’s theorem is therefore not available, and is not claimed.

What remains, and what is used after this section, is a pair of finite-set statements: the excised residue is T -measurable; and conditionally on the class, the word is uniformly distributed on that class, by a law that does not depend on Ξ .

5.1 Reductions on a finite set

Definition 5.1 (Measurable reduction). Let $\varphi : W \rightarrow Z$ be a function on the sample space. A statistic $S : W \rightarrow S(W)$ is a *sufficient reduction of w for φ* when φ is S -measurable: there exists $\psi : S(W) \rightarrow Z$ with $\varphi = \psi \circ S$. Equivalently, φ is constant on the fibres of S .

This is the deterministic content of Fisher–Neyman factorisation on a finite set. No family of laws is required. When a family is present it is named.

Definition 5.2 (Joint experiment for excision). The *joint experiment* for excision has observation

$$Y = (w, r) \in W \times (R \cup \{*\}), \quad (15)$$

parameter the unknown map $\Xi : X \rightarrow R$, and law

$$P_{\Xi}(w, r) = \frac{1}{|W|} \mathbf{1}_{\{r=\Xi(T(w))\}} \quad (16)$$

when liberation is suppressed, and

$$P_{\Xi, \eta}(w, r) = \frac{1}{|W|} \mathbf{1}_{\{r=\Xi(T(w))\}} \varpi(w, \eta) + \frac{1}{|W|} \mathbf{1}_{\{r=*\}} (1 - \varpi(w, \eta)) \quad (17)$$

when it is not. Equation (16) is the case $\varpi \equiv 1$. The word-marginal of either law is the counting measure of Remark 2.11 and does not depend on Ξ .

Remark 5.3 (Relation to Section 2.5). Definition 2.10 placed parameter-dependence in the decision maps rather than in a sampling law on words. Definition 5.2 does not alter that: the word-marginal remains flat. What it adds is the residue coordinate, without which a family $\{P_\Xi\}$ on W alone cannot see Ξ and every sufficiency claim about Ξ is vacuous. The joint observation is the chemically correct one: a cluster presents a word and a residue (or an absence).

Definition 5.4 (Neyman sufficiency, finite form). A statistic S of the joint observation Y is *Neyman-sufficient* for Ξ when the conditional law $P_\Xi(Y \in \cdot \mid S)$ does not depend on Ξ . Equivalently, the likelihood factorises as

$$P_\Xi(Y) = g(S(Y), \Xi) h(Y) \quad (18)$$

for functions g, h with h free of Ξ .

Definition 5.5 (Ancillarity, finite form). A statistic S of Y is *ancillary* for Ξ when the marginal law of S does not depend on Ξ . It is *conditionally ancillary* for Ξ given another statistic U when the conditional law of S given U does not depend on Ξ .

Definition 5.6 (Completeness, finite form). A statistic S is *complete* for a family $\{P_\theta\}$ when $\mathbb{E}_\theta[f(S)] = 0$ for every θ implies $f = 0$ almost surely under every P_θ .

5.2 Sufficiency of the class

Theorem 5.7 (Deterministic reduction). *[A] The excised residue, viewed as a function of the word, is T -measurable. The class statistic T is a sufficient reduction of w for excision in the sense of Definition 5.1.*

Proof. By Definition 2.17, Ξ is a map on X . The residue formed at w in the absence of liberation failure is $\Xi(T(w))$. That composite is T -measurable by construction. $\square$

Theorem 5.8 (Neyman sufficiency in the joint experiment). *[A] In the joint experiment of Definition 5.2 with $\varpi \equiv 1$, the class statistic T — more precisely the pair (T, r) — is Neyman-sufficient for the unknown map Ξ .*

Proof. The likelihood (16) is $P_\Xi(w, r) = |W|^{-1} \mathbf{1}_{\{r = \Xi(T(w))\}}$. This is of the form $g(T(w), r, \Xi) h(w, r)$ with $h = |W|^{-1}$ free of Ξ . Fisher–Neyman gives the claim. $\square$

Proposition 5.9 (Minimal sufficiency). *[A] In the joint experiment with $\varpi \equiv 1$, two observations (w_1, r_1) and (w_2, r_2) have proportional likelihood functions of Ξ if and only if $T(w_1) = T(w_2)$ and $r_1 = r_2$. Hence (T, r) is minimal sufficient for Ξ . Among statistics that are functions of the word alone, T is a minimal sufficient reduction for Ξ .*

Proof. The likelihood, as a function of Ξ , is the indicator of the constraint $r = \Xi(T(w))$, scaled by $|W|^{-1}$. Two such indicators coincide for every Ξ if and only if they impose the same constraint, which is the pair of equalities claimed. If $T(w_1) \neq T(w_2)$, an excision map sending the two classes to distinct residues separates the two likelihoods. The last sentence follows because any further reduction of T collapses two classes and the same separator applies. $\square$

Remark 5.10 (What is not claimed). Theorem 5.7 says that T captures all information in the word about *excision*. It does not say that T captures all chemical information, all translational information, or all information about the assignment c . Liberation may depend on η by $H_{\text{lib-env}}$. The assignment c depends on seed, propagation and refinement, which use the block coordinate B . Sufficiency of T for Ξ is compatible with failure of sufficiency of T for c . That failure is Section 6.

5.3 The within-class coordinate

The relative-frame statistic of Definition 2.12 takes values in a set F of frame-setters. Eight classes realise only two of those values. A statistic with that codomain therefore carries the orbit type, which is a function of T , and cannot be independent of T . The ancillary object is the position of the word *inside* its class.

Definition 5.11 (Orbit type and section). Write $\tau : W \rightarrow \{2, 4\}$ for the orbit type, $\tau(w) = |T^{-1}(T(w))|$. A *section* of T is a map $s : X \rightarrow W$ with $T \circ s = \text{id}_X$. Given a section s , every word w may be written

$$w = \gamma_s(w) \cdot s(T(w)) \quad (19)$$

for an element $\gamma_s(w)$ of Fr , unique modulo the stabiliser of $s(T(w))$. The class of $\gamma_s(w)$ in $\text{Fr}/\text{Stab}(s(T(w)))$ is

the *within-class coordinate* determined by s , denoted $\Phi_s(w)$.

Proposition 5.12 (Within-class uniformity). *[A] Fix a section s . Conditionally on $T = x$, the word w is uniformly distributed on the orbit x , and $\Phi_s(w)$ is uniformly distributed on $\text{Fr}/\text{Stab}(s(x))$. Neither conditional law depends on Ξ .*

Proof. The word-marginal is uniform on W . An orbit of size $n \in \{2, 4\}$ therefore carries conditional mass $1/n$ at each of its points. The identification (19) translates that uniformity into uniformity of Φ_s on the quotient of Fr by the stabiliser, which has order n by the orbit–stabiliser theorem. The counting measure does not depend on Ξ , and the fibres of T are the orbits of Fr , which are likewise free of Ξ . $\square$

Theorem 5.13 (Conditional ancillarity). *[A] For any section s , the within-class coordinate Φ_s is conditionally ancillary for Ξ given T . Equivalently: the conditional law of the observation given the class does not depend on the excision map.*

Proof. Proposition 5.12. $\square$

Proposition 5.14 (A leaks the orbit type). *[A] Let $A : W \rightarrow F$ be the relative-frame statistic of Definition 2.12, with $|F| = 4$. Then A is not independent of T under the counting measure. In particular A is not ancillary for the orbit type τ , and the pair (T, A) does not factor as a product measure on $X \times F$.*

Proof. A class of size two realises only two values of A . Hence $\{\tau = 2\} \subset A^{-1}(F_0)$ for a two-element subset $F_0 \subset F$, and $P(\tau = 2 \mid A \in F \setminus F_0) = 0$, while $P(\tau = 2) = 16/64 = 1/4$. Dependence follows. The eight size-two classes contribute sixteen words and the twelve size-four classes contribute forty-eight, so the push-forward of the counting measure on $X \times F$ is not a product. $\square$

Remark 5.15 (Basu is not available). Basu’s theorem [11] would give independence of a complete sufficient statistic and an ancillary statistic. It does not apply, for two independent reasons. First, the family of laws of T under $\{P_\Xi\}$ is a singleton: the word-marginal, hence the law of T , does not depend on Ξ . A singleton family on a twenty-element set is not complete, since any mean-zero function of T integrates to zero. Second,

the raw statistic A is not ancillary for Ξ in any sense that would cancel its dependence on τ , and τ is a function of T .

The content that Basu would have supplied is already in Theorems 5.7 and 5.13. Independence of class and within-class coordinate holds *conditionally on the class*, which is the only independence the geometry of the action provides.

5.4 The tether as a physical ancillary

Proposition 5.16 (The tether does not vary with the class). *[D] Under $H_{\text{teth}}(\text{i})$ –(ii) the tether word is AAA at every cluster. It is therefore a constant statistic on W , and is ancillary for every function of w , in particular for $\Xi \circ T$.*

Proof. Proposition 4.9. A constant map is independent of every random element defined on the same space, and its law is a Dirac mass free of Ξ . $\square$

Remark 5.17 (Two objects named “frame”). The poly-A tether fixes one global frame for all clusters (Remark 4.14). That global frame is a constant of the experiment, Proposition 5.16. The within-class coordinate Φ_s of Definition 5.11 is not a constant; it locates the word inside its orbit once the global frame has been fixed. Ancillarity of the tether is a statement about a physical reagent. Conditional ancillarity of Φ_s is a statement about the action of Fr on W . The two are not interchangeable. The informal sentence “the frame is ancillary for excision” is the conjunction of the two, and is not used below without saying which.

Remark 5.18 (Why the register must be non-Watson–Crick). If the tether paired Watson–Crick with the codon, the tether word would be a function of w and would fail to be a constant. It would then carry information about $T(w)$, hence about excision. Proposition 5.16 uses $H_{\text{teth}}(\text{ii})$ in an essential way. This is the precise form of the preview at Remark 4.21.

5.5 Liberation is not a class statistic

Proposition 5.19 (Liberation need not be T -measurable). *[A] There exist predicates $\varpi : W \times E \rightarrow \{0, 1\}$ for which $w \mapsto \varpi(w, \eta)$ fails to be constant on classes. Under any such predicate, $\text{absence} \ast$ is not a function of T .*

Proof. The fibres of T have size two or four. Any predicate that takes both values on a single fibre fails to be T -measurable. Such predicates exist; $H_{\text{lib-env}}$ permits them. The second sentence is the first applied to (13). $\square$

Remark 5.20 (The stop designations). [6, Cor. 6.4] produces two absences as a consequence of parity inside a covered block, which is a function of T and B together. The third published absence is an environmental decision under $H_{\text{lib-env}}$ and is not claimed to be T -measurable. Sufficiency of T for excision is compatible with both.

Remark 5.21 (What this section has fixed). Excision is a function of the class. In the joint experiment the pair (T, r) is minimal Neyman-sufficient for the unknown map Ξ . Conditionally on the class, the word is uniform on that class by a law free of Ξ . The tether is a constant, hence ancillary for excision, under the non-Watson-Crick register. The raw four-valued frame coordinate A is not independent of T . Basu's theorem is not invoked. Liberation is not assumed T -measurable. Nothing has been said about sufficiency of T for the block, or for the assignment c .

6 Transversality as a Failure of Sufficiency

Theorem 5.7 says that excision is a function of the class. A codon table is a function of the word, organised in published form by blocks. If the class determined the block, the two would be the same datum and the rest of the construction would be empty. They do not. A class meets a block in at most one word; neither partition refines the other; and the only assignments measurable with respect to both are the constants. Transporting a residue from a class onto a codon table therefore requires a choice the class does not make. That choice is the seed of Section 7.

6.1 Two partitions

The maps $T : W \rightarrow X$ and $B : W \rightarrow \mathcal{B}$ of Definitions 2.8 and 2.9 are the two partitions of the sixty-four-element set. Write $\sigma(T)$ and $\sigma(B)$ for the algebras they generate, as in Definition 2.16.

Definition 6.1 (Incidence). A class $x \in X$ and a block $b \in \mathcal{B}$ are *incident* when $x \cap b \neq \emptyset$. The *incidence graph* Γ is the bipartite graph with parts X and $\mathcal{B}$ and an

edge for each incident pair. Write $E(\Gamma)$ for the edge set.

Theorem 6.2 (Transversality). [A] *A class meets a block in at most one word. Equivalently, every edge of Γ corresponds to a unique word, and the map*

$$W \rightarrow E(\Gamma), \quad w \mapsto (T(w), B(w)) \quad (20)$$

is a bijection.

Proof. Two distinct words of the same block agree in their first two letters and differ only in the third. It is enough to show that no non-identity element of $\text{Fr} = \langle \text{flip}, \text{rev} \rangle$ carries a word to a distinct word of the same block.

Write ρ for the letterwise action of flip and τ for the letterwise involution through which rev acts, as in [2, Props. 4.7, 4.8]. Thus $\text{flip}(XYZ) = \rho(X)\rho(Y)\rho(Z)$ and $\text{rev}(XYZ) = \tau(Z)\tau(Y)\tau(X)$. The remaining non-identity element is $\text{flip} \circ \text{rev}$, acting by $(\text{flip} \circ \text{rev})(XYZ) = \sigma(Z)\sigma(Y)\sigma(X)$ where $\sigma = \rho\tau$. By [2, Prop. 4.8], ρ is fixed-point-free and each of τ, σ fixes exactly two letters. If $\text{flip}(XYZ)$ lies in the block of XYZ , then $\rho(X) = X$ and $\rho(Y) = Y$. This is impossible.

If $\text{rev}(XYZ)$ lies in the block of XYZ , then $\tau(Z) = X$ and $\tau(Y) = Y$. The first equality gives $Z = \tau(X)$, whence $\text{rev}(XYZ) = XY\tau(X) = XYZ$. The image is not distinct.

If $(\text{flip} \circ \text{rev})(XYZ)$ lies in the block of XYZ , then $\sigma(Z) = X$ and $\sigma(Y) = Y$. The first equality gives $Z = \sigma(X)$, whence the image is again XYZ .

Hence no class contains two words of the same block. Each word determines an edge and each edge contains at least one word, so (20) is a bijection. This is [6, Prop. 5.2]. $\square$

Corollary 6.3 (Degrees). [A] *The incidence graph has sixty-four edges. The degree of a class is the size of that class, hence two or four. The degree of every block is four.*

Proof. Theorem 6.2 identifies edges with words. A class of size n meets n blocks. A block contains four words, hence meets four classes. $\square$

Proposition 6.4 (Neither partition refines the other). [A] *There exist words w, w' with $T(w) = T(w')$ and $B(w) \neq B(w')$, and words u, u' with $B(u) = B(u')$ and $T(u) \neq T(u')$. Equivalently,*

$$\sigma(T) \not\subseteq \sigma(B), \quad \sigma(B) \not\subseteq \sigma(T). \quad (21)$$

Proof. A class of size four meets four blocks (Corollary 6.3), which supplies the first pair. A block meets four classes, which supplies the second. $\square$

Corollary 6.5 (*T* is not a sufficient reduction for the block). [A] *The block statistic B is not T-measurable. In the sense of Definition 5.1, T is not a sufficient reduction of the word for the block.*

Proof. Proposition 6.4. $\square$

Proposition 6.6 (The pair separates points). [A] *The joint statistic (T, B) separates points of W. Equivalently, $\sigma(T) \vee \sigma(B)$ is the discrete algebra on W, and (T, B) is a sufficient reduction of the word for every function of the word.*

Proof. Theorem 6.2: the fibre of (T, B) over an incident pair is a singleton, and over a non-incident pair is empty. $\square$

6.2 Measurability of assignments

An assignment $c : W \rightarrow R \cup \{*\}$ is *T-measurable* when it is constant on classes, and *B-measurable* when it is constant on blocks. The class-extension of an excision map is the assignment

$$c_{\Xi} := \Xi \circ T, \quad (22)$$

which is *T-measurable* by construction. Published codon tables are organised by blocks and, except for the splits treated in Section 8, are *B-measurable* or fail *B-measurability* on only four blocks.

Proposition 6.7 (Two components). [A] *The incidence graph Γ has exactly two connected components. They are*

$$W_0 = \{w \in W : \text{the middle letter of } w \text{ lies in } \{A, U\}\}, \quad (23)$$

$$W_1 = \{w \in W : \text{the middle letter of } w \text{ lies in } \{C, G\}\}. \quad (24)$$

Each component contains ten classes and eight blocks. Both generators of Fr preserve the Watson–Crick class of the middle letter, so no class and no block meets both components.

Proof. flip acts letterwise by ρ , which swaps *A* with *U* and *C* with *G*, and therefore preserves $\{A, U\}$ and $\{C, G\}$. rev acts letterwise by τ and then reverses; τ fixes *A, U* and swaps *C, G*, so likewise preserves those pairs. The middle letter of flip(*w*) and of rev(*w*) therefore lies in the

same Watson–Crick pair as the middle letter of *w*. Every orbit, and every block XY^{**} , lies in one of W_0, W_1 . It remains to see that each side is connected.

On W_0 the eight blocks are XA and XU for $X \in \Sigma$. Letterwise flip pairs AA with UU , AU with UA , CA with GU , and CU with GA . Reversal of the word XYA , $Y \in \{A, U\}$, lands in the block AY^{**} . Thus every block XA with $X \neq A$ meets a class that also meets AA , and every block XU with $X \neq A$ meets a class that also meets AU . The flip-edge AA – UU together with the reversal-edge UU – AU joins the two families. So the incidence graph of W_0 is connected.

On W_1 the eight blocks are XC and XG . Letterwise flip pairs AC with UG , AG with UC , CC with GG , and CG with GC . Reversal of XYC , $Y \in \{C, G\}$, lands in the block $G\tau(Y)^{**}$. From AC one reaches GG ; from AG one reaches GC ; the flip-edges then absorb the remaining four blocks. So the incidence graph of W_1 is connected. Hence there are exactly two components. The census of Appendix A confirms the count and the partition. $\square$

Theorem 6.8 (Joint measurability forces constancy on components). [A] *An assignment that is both T-measurable and B-measurable is constant on each component of Γ . Equivalently,*

$$\sigma(T) \cap \sigma(B) = \{\emptyset, W_0, W_1, W\}. \quad (25)$$

In particular such an assignment takes at most two values.

Proof. Let c be both *T-measurable* and *B-measurable*. Then each level set $c^{-1}(r)$ is a union of classes and a union of blocks, hence a union of incidence components. Proposition 6.7 supplies exactly two components. $\square$

Corollary 6.9 (Class-extension is a block code only on components). [A] *The class-extension c_{Ξ} is B-measurable if and only if Ξ is constant on the classes of W_0 and constant on the classes of W_1 . In particular a residue assignment that takes three or more values does not, by itself, produce a block-organised assignment.*

Proof. c_{Ξ} is *T-measurable*. Theorem 6.8 forces it, once *B-measurable*, to be constant on each component. $\square$

Remark 6.10 (The claim that would have been too strong). It is not true that a non-constant-on-blocks assignment cannot be *T-measurable*. A *T-measurable* assignment is constant on each class, and a class of size four places the same residue on four different blocks. Variation inside a single block is still allowed,

and is typical: the four words of a block lie in four classes (Corollary 6.3). The obstruction is the other way. Block-organised tables are (approximately) B -measurable; class-extension is T -measurable; the two conditions together admit at most one value per incidence component. That is Theorem 6.8, and is the sense in which T fails to be sufficient for the assignment. A two-valued assignment, one value on W_0 and one on W_1 , is jointly measurable and is not a codon table.

Theorem 6.11 (Sufficiency for excision does not give sufficiency for the assignment). *[A] There exist excision maps Ξ and assignments c compatible with Ξ in the sense of (13) such that c is not T -measurable. Compatibility with excision does not force the assignment to be a function of the class.*

Proof. Take Ξ non-constant and take c equal to c_Ξ on one incident word of each class and equal to a different value on some other word of a size-four class. Compatibility at the seeded word is preserved; T -measurability is not. Concrete maps of this form are constructed in [6, §5] and will be imported in Section 7. Existence alone is the present claim. $\square$

6.3 Information between class and block

The two partitions are dependent. The dependence is finite and can be written down.

Proposition 6.12 (Entropies). *[A] Under the counting measure, with logarithms in base two,*

$$H(B) = 4, \quad (26)$$

$$H(T) = 4.25, \quad (27)$$

$$H(T, B) = 6, \quad (28)$$

$$I(T; B) = 2.25. \quad (29)$$

Proof. There are sixteen blocks of four words, so B is uniform on sixteen values and $H(B) = \log_2 16 = 4$. The class type is eight classes of probability $2/64 = 1/32$ and twelve of probability $4/64 = 1/16$, so

$$H(T) = -8 \cdot \frac{1}{32} \log_2 \frac{1}{32} - 12 \cdot \frac{1}{16} \log_2 \frac{1}{16} = 8 \cdot \frac{5}{32} + 12 \cdot \frac{4}{16} = \frac{5}{4} + 3 = \frac{17}{4} = 4.25 \quad (30)$$

The joint (T, B) takes sixty-four values of probability $1/64$ by Theorem 6.2, so $H(T, B) = \log_2 64 = 6$. Mutual information is the inclusion–exclusion of those three numbers. $\square$

Remark 6.13 (What the 2.25 bits are). $I(T; B)$ is the information each partition carries about the other, not the information either carries about excision. Excision is a function of T , so $I(\Xi(T); B) = I(T; B)$ only after Ξ is known to be injective, which it is not: twenty classes take values in a residue set of cardinality at most twenty, and published tables identify several classes. The numerical value is recorded because it is cheap and checkable. It is not an ingredient of Theorems 6.8 or 6.11. A machine regeneration of (26)–(29) is deferred to Appendix A.

6.4 Why a seed is required

Remark 6.14 (The missing coordinate). A word is a pair $(T(w), B(w))$ by Theorem 6.2, or a pair $(T(w), A(w))$ by Proposition 4.16. Excision uses only the first coordinate. An assignment organised by blocks uses the second. The gap between Theorem 5.7 and a codon table is exactly the failure of B to be T -measurable. A seed is a choice, for each class that produces a residue, of one incident block on which that residue is placed. It is a section of incidence, not a function of T .

Remark 6.15 (What this section has fixed). Classes and blocks are transverse. The incidence graph is a sixty-four-edge bipartite graph of degrees two-or-four against four, with two components split by the Watson–Crick class of the middle letter. T is not a sufficient reduction for B . The pair (T, B) separates points. Jointly measurable assignments are constant on each component and take at most two values. Class-extension of a non-constant excision map is not a block code. Mutual information $I(T; B)$ equals 2.25 bits under the counting measure and is not used below. No seed has been chosen.

7 Seeds as Sections and Propagation as Conditional Expectation

Section 6 left a missing coordinate. A word is an incident pair $(T(w), B(w))$; excision uses only the first; a block-organised assignment uses the second. A seed chooses, for certain classes, one incident block on which the excised residue is placed. Propagation extends that value along the block. Neither operation is a function of T .

The demand graph against which the seeds are chosen is built from the standard table. Declaration 2.7 applies to every statement in this section that mentions demand, coverage, or the count of maps. Those statements are

labelled $[D]$ under H_{seed} together with that constructive datum. They are not theorems of $\mathcal{A}$.

7.1 Seeds as sections of incidence

Definition 7.1 (Seed). Let $S \subset X$ be a set of classes. A *seed* on S is a section of incidence over S : a map

$$s : S \rightarrow W \quad (31)$$

such that $T(s(x)) = x$ for every $x \in S$. Equivalently, s selects, for each chosen class, one edge of Γ incident to that class, and hence one block $B(s(x))$. Write $\text{im } s$ for the set of seeded words and $B(s(S))$ for the set of seeded blocks.

Remark 7.2 (A section of T , restricted). A seed is a section of T over S , not over all of X . Classes outside S are not assigned a word. Nothing in Definition 7.1 requires $S = X$, and the covering theorems below take $|S| = 17$.

Definition 7.3 (Seeded assignment). Given a seed s on S and an excision map Ξ , the *seeded assignment* is the partial map

$$c_s : \text{im } s \rightarrow R, \quad c_s(s(x)) = \Xi(x). \quad (32)$$

It is defined only on seeded words. Compatibility with excision at those words is (13) restricted to $\text{im } s$.

7.2 Demand

Definition 7.4 (Demand of a class). Let $\mathcal{C}_{\text{std}}$ be the standard assignment of Definition 2.14. For each class $x \in X$ write

$$D(x) := \{b \in \mathcal{B} : \text{there exists } w \in x \cap b \text{ with } \mathcal{C}_{\text{std}}(w) = \Xi(x)\} \quad (33)$$

The set $D(x)$ is the *demand* of x . It is a subset of the neighbourhood of x in Γ , possibly empty (if $\Xi(x)$ does not occur on any block that x meets in the standard table), and at most of size $|x|$. The *demand graph* is the spanning subgraph of Γ whose edges are the pairs (x, b) with $b \in D(x)$.

Remark 7.5 (Constructive datum). Equation (33) uses $\mathcal{C}_{\text{std}}$. A different published table produces a different demand graph and a different covering count. The vertebrate mitochondrial table is not used to build de-

mand; it appears as an output in Section 8. This is the constructive use declared at Declaration 2.7, restated so that it can be checked at the point of use.

Definition 7.6 (Matching and coverage). A seed s on S is a *matching of demand* when $B(s(x)) \in D(x)$ for every $x \in S$. The seed *covers* a block b when $b \in B(s(S))$. It *covers* a residue r when $r = \Xi(x)$ for some $x \in S$. It *covers the table* when every block on which $\mathcal{C}_{\text{std}}$ takes a residue value is covered and every residue in the image of $\mathcal{C}_{\text{std}}$ is covered.

The last clause is Hall's problem on the demand graph: select a set of vertices on the class side, and one incident demand-edge at each, so that the selected edges meet every block that the standard table uses and every residue the standard table displays.

7.3 Hall and seventeen seeds

Theorem 7.7 (Seventeen seeds cover). *[D] Under H_{seed} , and with demand constructed from $\mathcal{C}_{\text{std}}$ as in Definition 7.4, there exists a matching of demand on a set S of seventeen classes that covers the table in the sense of Definition 7.6.*

Proof. This is [6, Thm. 5.7]. The argument there is Hall's marriage theorem [10] applied to the demand graph of that paper, which is the graph of Definition 7.4 under the present assignment. Existence of a matching that meets every used block is the content of Hall's criterion on that finite bipartite graph, as enumerated in [6, Thm. 5.7]. The count seventeen is the minimal cardinality of a class-set for which such a matching exists in that graph; it is not a consequence of $|X| = 20$ alone. $\square$

Remark 7.8 (Why not twenty). Three classes are not needed as seeds. Propagation along a block places a residue at words whose classes were not selected; those classes are thereby represented in the assignment without being elements of S . The arithmetic is not $20 - 3 = 17$ forced by parity: it is the output of Hall's criterion on this particular demand graph. A different table, or a different excision map, can change the number.

Proposition 7.9 (Hall is on demand, not on incidence). *[A] The full incidence graph Γ admits many matchings of any seventeen classes: every class has*

degree two or four (Corollary 6.3), so Hall's criterion holds on Γ for every collection of at most sixteen classes and fails to be the constraint. The constraint is the subgraph of Definition 7.4.

Proof. A bipartite graph with right degrees four and left degrees at least two satisfies Hall's criterion for left-sets of size at most sixteen because any $k \leq 16$ left vertices have at least $\lceil k \cdot 2/4 \rceil$ neighbours in the crudest bound, and in fact many more; the present claim does not need the sharp count. Demand deletes edges. After deletion, Hall's criterion is no longer automatic, which is why Theorem 7.7 is labelled [D] and cites an enumeration. $\square$

7.4 Propagation

Definition 7.10 (Occupied blocks). Given a seed s on S , a block b is *occupied* when $b \in B(s(S))$. Write $\mathcal{B}_s$ for the set of occupied blocks. The seed is *non-conflicting* when Ξ is constant on $s^{-1}(B^{-1}(b))$ for every occupied block b : no two seeds in the same block carry distinct residues.

Definition 7.11 (Propagation). Let s be a non-conflicting seed. The *propagation* of c_s is the partial assignment

$$\hat{c}_s : B^{-1}(\mathcal{B}_s) \rightarrow R \quad (34)$$

defined by

$$\hat{c}_s(w) = \Xi(x) \quad \text{whenever } B(w) = B(s(x)). \quad (35)$$

On an occupied block this is well-defined by the non-conflict hypothesis. On an unoccupied block it is undefined.

Theorem 7.12 (Unique block-measurable extension). [A] Let s be a non-conflicting seed. Among maps $B^{-1}(\mathcal{B}_s) \rightarrow R$, the propagation $\hat{c}_s$ is the unique B -measurable extension of the seeded assignment c_s .

Proof. On each occupied block b , the restriction of c_s takes a single value r_b by non-conflict. Any B -measurable extension must be constant on b and agree with c_s at the seeded word, hence must equal r_b on all of b . Distinct occupied blocks are disjoint, so these constants assemble uniquely. Unoccupied blocks are outside the domain. $\square$

Definition 7.13 (Block kernel). The *block kernel* is

the Markov kernel K on W given by

$$K(w, w') := \begin{cases} 1/4 & \text{if } B(w') = B(w), \\ 0 & \text{otherwise.} \end{cases} \quad (36)$$

It is the conditional law of a uniform word given $\sigma(B)$.

Proposition 7.14 (Propagation as support of the kernel). [A] Let s be a non-conflicting seed and let $\delta_{s(x)}$ be the Dirac mass at a seeded word. The measure $\delta_{s(x)}K$ is uniform on the block $B(s(x))$, and the support of that measure is the set of words to which $\Xi(x)$ is propagated.

Proof. Immediate from (36) and (35). $\square$

Remark 7.15 (Conditional expectation is the wrong operator for values). If one regards c_s as a partially defined real-valued function and completes it by zero off $\text{im } s$, the conditional expectation $\mathbb{E}[c_s \mid \sigma(B)]$ averages the seeded value against three zeros on a block of size four, and returns $r_b/4$. That is not an assignment. Propagation copies the value, it does not average it. The Markov reading that survives is Proposition 7.14: the block kernel supplies the *support* on which the value is copied. Conditional expectation is recorded so that it is not used.

Remark 7.16 (Kan extension). Let $i : \mathcal{B}_s \hookrightarrow \mathcal{B}$ be the inclusion of occupied blocks, and let $f : \mathcal{B}_s \rightarrow R$ be the map sending an occupied block to the unique residue of the seeds it contains. In the category of sets, the left Kan extension of f along i is the function $\text{Lan}_i f : \mathcal{B} \rightarrow R \sqcup \{*\}$ equal to f on $\mathcal{B}_s$ and to $*$ off it. Pulling back along B recovers $\hat{c}_s$ on occupied blocks and leaves unoccupied blocks unset, which is Definition 7.11. The observation is that Theorem 7.12 has this name in **Set**. It is not used as a method of calculation.

7.5 The eight hundred and forty maps

Proposition 7.17 (Eight hundred and forty). [D] Under H_{seed} , with demand constructed from $\mathcal{C}_{\text{std}}$, the number of matchings of demand on some seventeen-class set that cover the table is eight hundred and forty.

Proof. [6, Prop. 5.14], by enumeration of the demand graph. $\square$

Remark 7.18 (No choice is made). This paper does not select a seed among the eight hundred and forty. Everything after Theorem 7.12 that refers to “a seed” means an arbitrary matching of demand that covers the table. Where two such seeds produce different partial assignments, the difference lies in which class seeds which block, not in the propagated values on occupied unsplit blocks; those values are the residues of $\mathcal{C}_{\text{std}}$ on those blocks, by construction of demand.

Remark 7.19 (H_{seed} is not discharged). Hypothesis 2.20 asserts three things: that excision seeds the block at a codon the class shares with that block; that on a block whose halves carry distinct residues the seed lies on the pyrimidine half; and that the seed propagates within the block. The first and third clauses are the definitions of this section. The second clause is a restriction on which edges of the demand graph are eligible when the block splits, and is used in Section 8. None of the three is derived from $\mathcal{A}$.

7.6 What propagation does not do

Proposition 7.20 (Unoccupied blocks remain unset). *[A] If $b \notin \mathcal{B}_s$, the propagation $\hat{c}_s$ does not assign a value on b . No clause of H_{seed} fills an unoccupied block.*

Proof. Definition 7.11. $\square$

Remark 7.21 (Splits are not propagation). A block on which the published table takes two values is not a block on which $\hat{c}_s$ can agree with the table: propagation is constant on occupied blocks. The two-valued blocks are the content of $H_{\text{split-pos}}$ and $H_{\text{split-chem}}$, and of Section 8. Treating a split as a failure of propagation would misread Theorem 7.12. The theorem says that propagation, if it occurs on a block, is constant there. Whether it occurs on the whole block or on a half is a different datum.

Remark 7.22 (What this section has fixed). A seed is a section of incidence over a subset of classes. Demand is read from $\mathcal{C}_{\text{std}}$. Under H_{seed} , seventeen seeds cover the table, and eight hundred and forty such matchings exist. Propagation is the unique B -measurable extension of a non-conflicting seed to the occupied blocks. It is the support of the block kernel, and it is a left Kan extension along the inclusion of occupied blocks; it is not an L^2 conditional expectation. No seed is chosen.

Split blocks are untouched. The mitochondrial table has not been used.

8 Refinement, Normal Form as Range, and the Audit

Propagation is constant on occupied blocks. Published tables are not. The standard table realises patterns (4), (2, 2), (3, 1) and (2, 1, 1). The vertebrate mitochondrial table realises only (4) and (2, 2). The two-valued blocks are not failures of Theorem 7.12. They are a further operation, at the third position, under $H_{\text{split-pos}}$ and $H_{\text{split-chem}}$. This section records the operation, proves that its only possible patterns are (4) and (2, 2), and imports the identification of that range with the vertebrate mitochondrial table. The identification uses four hypotheses and the constructive demand of Section 7. It does not discharge Φ .

8.1 The third-position partition

Definition 8.1 (Purine and pyrimidine). Write

$$Y := \{C, U\}, \quad R := \{A, G\} \quad (37)$$

for the pyrimidine and purine pairs in Σ . For each block $b = XY**$ write

$$b_Y := \{XYC, XYU\}, \quad b_R := \{XYA, XYG\} \quad (38)$$

for its *pyrimidine half* and *purine half*. The two halves partition the block.

Remark 8.2 (Why this bipartition). $H_{\text{split-pos}}$ names the third-position cut and names it as pyrimidine against purine. The other two bipartitions of Σ , $M|K = \{A, C\} \mid \{G, U\}$ and $W|S = \{A, U\} \mid \{C, G\}$, are available and unused. No argument from $\mathcal{A}$ selects $R|Y$ among the three. The selection is the hypothesis, and is checked against published tables at Theorem 8.11 below, not derived.

8.2 Refinement

Definition 8.3 (Refinement of an occupied block). Let s be a non-conflicting covering seed and $\hat{c}_s$ its propagation. An occupied block b is *unsplit* when $\hat{c}_s$ is the assignment on all of b . It is *split* when the

assignment on b is the pair of constants

$$c(w) = \begin{cases} \widehat{c}_s(w) & \text{if } w \in b_Y, \\ r_b^\# & \text{if } w \in b_R, \end{cases} \quad (39)$$

for a residue $r_b^\#$ supplied by the vessel under $H_{\text{split-chem}}$. The seed of a split block is required to lie in b_Y , which is the second clause of H_{seed} .

Remark 8.4 (Two residues, one seed). A split block carries two residues and only one seed. The seeded residue occupies the pyrimidine half by the clause just cited. The purine residue is not obtained by a second seed: a second seed in the same block would violate non-conflict unless the two residues agreed, which would unsplit the block. $H_{\text{split-chem}}$ is the source of the second residue, and is a different hypothesis from H_{seed} .

Definition 8.5 (Refined assignment). Given a covering seed s , a set Σ_{split} of occupied blocks marked for splitting, and a choice of purine residues $(r_b^\#)_{b \in \Sigma_{\text{split}}}$, the *refined assignment* $c^\#$ is $\widehat{c}_s$ on unsplit occupied blocks, the split rule (39) on split occupied blocks, and $*$ on unoccupied blocks.

Unoccupied blocks remain unset by Proposition 7.20. In the covering theorems of Section 7 there are none among the blocks the standard table uses. Absence on a used block, when it occurs, is liberation, Section 8.6.

8.3 Patterns and normal form

Definition 8.6 (Block pattern). Let c be an assignment and b a block. Write $m(c, b)$ for the multiset of values of c on b , ignoring $*$ or not according as the clause at hand includes liberation. The *pattern* of b under c is the partition of 4 given by the multiplicities in that multiset. The patterns that occur below are

$$(4), \quad (2, 2), \quad (3, 1), \quad (2, 1, 1), \quad (1, 1, 1, 1). \quad (40)$$

Definition 8.7 (Normal form). An assignment c is in *normal form* when every block on which c takes residue values has pattern (4) or (2, 2), and every block of pattern (2, 2) is partitioned as $b_Y \mid b_R$.

Theorem 8.8 (Range of refinement). *[A] Every refined assignment of Definition 8.5 is in normal form. No refined assignment realises pattern (3, 1), (2, 1, 1) or (1, 1, 1, 1) on an occupied block.*

Proof. On an unsplit occupied block, $\widehat{c}_s$ is constant by Theorem 7.12, which is pattern (4). On a split occupied block, (39) places one residue on the two pyrimidine words and one residue on the two purine words. If the two residues agree, the block is pattern (4) and was split only formally. If they differ, the block is pattern (2, 2) partitioned as $b_Y \mid b_R$. A pattern (3, 1) would require a part of size three, which neither a constant nor a bipartition into pairs supplies. A pattern (2, 1, 1) or (1, 1, 1, 1) would require the purine half, the pyrimidine half, or both, to take two values; each half is a pair on which (39) is constant. $\square$

Corollary 8.9 (The standard table is not in the range). *[A] No refined assignment equals $\mathcal{C}_{\text{std}}$.*

Proof. $\mathcal{C}_{\text{std}}$ realises patterns (3, 1) and (2, 1, 1) (Theorem 8.11(ii)). Theorem 8.8 forbids both. $\square$

Remark 8.10 (The range is forced by the cut, not by the table). Theorem 8.8 does not use $\mathcal{C}_{\text{std}}$ or $\mathcal{C}_{\text{mt}}$. It uses only that refinement is a bipartition of each split block into two pairs, each pair constant. Any other third-position bipartition — $M \mid K$ or $W \mid S$ at position three — would give the same list of patterns and a different notion of half. The mitochondrial table enters as a comparison, not as a hypothesis of the range statement.

8.4 Published splits

Theorem 8.11 (Published two-valued blocks split as $Y \mid R$). *[A] In tables 1 and 2 of [8] the following hold.*

- (i) *Every block of pattern (2, 2) is partitioned as $b_Y \mid b_R$. None is partitioned as amino against keto or as Watson–Crick against its complement, at the third position.*
- (ii) *The standard table has one block of pattern (3, 1), AUN (Ile/Met), and one of pattern (2, 1, 1), UGN (Cys/stop/Trp). It is not in normal form.*
- (iii) *The vertebrate mitochondrial table has no block of pattern (3, 1) or (2, 1, 1). It is in normal form.*

Proof. Enumeration of the sixteen blocks in each table. The standard blocks of pattern (2, 2) are UUN

(Phe/Leu), UAN (Tyr/stop), CAN (His/Gln), AAN (Asn/Lys), GAN (Asp/Glu) and AGN (Ser/Arg); each is Y against R . AUN is Ile on AUY and AUA, Met on AUG. UGN is Cys on UGY, stop on UGA, Trp on UGG. The vertebrate mitochondrial table moves AUA to Met, UGA to Trp, and AGR to stop, which converts AUN and UGN to pattern (2, 2) with the same cut and leaves AGN as Ser against stop. No other pattern appears. This is [6, Thm. 6.2]. The pattern census is regenerated as check V6 of Appendix A. $\square$

Remark 8.12 ($H_{\text{split-pos}}$ as a statement about tables). Theorem 8.11 is $H_{\text{split-pos}}$ evaluated on two published functions. It is not a derivation of $H_{\text{split-pos}}$ from $\mathcal{A}$. It is the reason the hypothesis is cheap to carry: the cut it names is the cut the tables use. The hypothesis remains a hypothesis of the construction.

8.5 The mitochondrial table as range

Theorem 8.13 (Output is the vertebrate mitochondrial assignment). *[D] There exist a covering seed s among the eight hundred and forty, a split set Σ_{split} , and purine residues ($r_b^\#$) such that the refined assignment $c^\#$, after the liberation of Section 8.6, equals $\mathcal{C}_{\text{mt}}$. The four matching hypotheses are H_{seed} , $H_{\text{split-pos}}$, $H_{\text{split-chem}}$ and $H_{\text{lib-env}}$. Chain polarity $H_{\text{teth}}(\text{iii})$ is used only to identify the domain with a sixty-four-codon table and is not a fifth matching hypothesis.*

Proof. [6, Thm. 5.8], under those four matching hypotheses, with demand constructed from $\mathcal{C}_{\text{std}}$. The seed, the split set and the purine residues are exhibited there and are not re-chosen here. Equality is an equality of two functions on a sixty-four-element set, under the convention of Remark 2.15(ii). $\square$

Remark 8.14 (Demand from one table, comparison with another). Demand was read from $\mathcal{C}_{\text{std}}$. The output comparison is with $\mathcal{C}_{\text{mt}}$. Those are two uses of two published functions, and they are not interchangeable. The standard table supplies the edges of the demand graph. The mitochondrial table is the function the refined assignment is claimed to equal. The differences between the two tables — AUA as Met rather than Ile, UGA as Trp rather than stop, AGR as stop rather than Arg — are accounted for by the split set, the purine residues, and liberation, not by a change of demand. Explicitly: demand uses $\mathcal{C}_{\text{std}}$ only to choose, for each residue, which blocks display it. After propagation and the $Y \mid R$ cut, the con-

struction occupies AU as Ile against Met and UG as Cys against Trp, and leaves the purine halves of UA and AG unoccupied (Proposition 8.16). That is the vertebrate mitochondrial assignment. The standard table is that assignment with two further divisions, at AUA and UGA [6, abstract]. Those two divisions lie outside the range of refinement (Corollary 8.9) and are not performed. The mechanical step is therefore: demand from $\mathcal{C}_{\text{std}}$, output the unrefined-at-AUA-and-UGA form of that table, which is $\mathcal{C}_{\text{mt}}$. The existence of a seed that realises this is [6, Thm. 5.8] and is not re-derived.

Remark 8.15 (Equality is not identification). Theorem 8.13 compares two maps $W \rightarrow R \cup \{*\}$. It does not identify the image of Ξ with biological amino acids, and it does not discharge Φ (Remark 2.26). The status is [D] because four hypotheses are in the antecedent.

8.6 Stops and liberation

Proposition 8.16 (Stops are absences). *[D] Under H_{seed} , the blocks UA and AG are in normal form with non-coding purine halves. The words UAA, UAG, AGA and AGG therefore carry $*$. These are the four stop designations of the vertebrate mitochondrial table.*

Proof. [6, Cor. 6.4]. The argument is that a codon carries a residue only where one is seeded or propagated; those four words are the unseeded purine halves of two split blocks. It is not [5, Prop. 5.9]. $\square$

Remark 8.17 (The third absence). The standard table's third stop, UGA, is not a stop in the mitochondrial output; AGR are stops in that output and absences by Proposition 8.16. What remains environmental under $H_{\text{lib-env}}$ is the standard reading of UGA, and any further setting in which the same codon releases in one translational environment and not in another [6, Target 9.3]. Those decisions are not claimed to be T -measurable (Proposition 5.19). H_{stop} of [4] remains withdrawn, as recorded at Remark 2.25.

8.7 What the vessel does not decide

Remark 8.18 ($H_{\text{split-chem}}$ is open). $H_{\text{split-chem}}$ asserts that the purine residue on a split block is determined by hydrogen bonding, steric hindrance and electrostatics together, and not by size alone. [6, Cor. 6.7] exhibits a split block on which residue volume [9] does

not select the published purine residue. Size alone is therefore not a substitute for the hypothesis. No other discharged rule is offered. The purine residues used in Theorem 8.13 are read from the table they are later compared with, which is circular as a derivation and is not presented as one. The hypothesis remains open.

8.8 Audit

Definition 8.19 (Audit). An *audit* of a refined assignment $c^\sharp$ against a published table $\mathcal{C}$ is the set of words at which the two disagree,

$$\Delta(c^\sharp, \mathcal{C}) := \{w \in W : c^\sharp(w) \neq \mathcal{C}(w)\}. \quad (41)$$

The audit is empty when the two functions are equal.

Proposition 8.20 (Audit against the mitochondrial table). [D] For the refined assignment of Theorem 8.13, $\Delta(c^\sharp, \mathcal{C}_{\text{mt}}) = \emptyset$.

Proof. Restatement of Theorem 8.13. The disagreement set of the two published tables is regenerated as check V7 of Appendix A. Equality with $\mathcal{C}_{\text{mt}}$ remains [6, Thm. 5.8] and is not re-enumerated here. $\square$

Proposition 8.21 (Audit against the standard table). [A] For the same assignment, $\Delta(c^\sharp, \mathcal{C}_{\text{std}})$ consists of the words at which the standard and vertebrate mitochondrial tables disagree. Those words are AUA, UGA, AGA and AGG, under the identification $T \leftrightarrow U$.

Proof. Tables 1 and 2 of [8] differ exactly at those four codons (and in the naming of the corresponding stops). The first audit being empty, the second is that difference set. $\square$

Remark 8.22 (What an empty audit is not). An empty audit is an equality of functions into $R \cup \{*\}$, under Remark 2.15(ii). It is not evidence that the residues produced at a chrysene cluster are biological amino acids, and it is not a verification of any chemical mechanism. Those readings are [E] and are not made.

8.9 Hypotheses still live

Remark 8.23 (Register after the audit). The following remain undischarged, and are used in Theorem 8.13:

- H_{teth} (iii), chain polarity, Target 9.7 of [6];

- H_{seed} , including the pyrimidine-half clause;
- $H_{\text{split-pos}}$, as a rule of the construction (the tables agree with it);
- $H_{\text{split-chem}}$, open, size-alone refuted;
- $H_{\text{lib-env}}$, for the third absence and for any further environmental decision.

No new hypothesis has been added. H_{stop} remains withdrawn. Φ remains unused.

Remark 8.24 (What this section has fixed). Refinement is a $Y \mid R$ cut at the third position on marked blocks. Every refined assignment is in normal form: patterns (4) or (2, 2) with that cut. Under four hypotheses and the constructive demand of Section 7, there is a refined assignment equal to the vertebrate mitochondrial table. Two absences come from parity; a third from the environment. $H_{\text{split-chem}}$ is not a theorem. The audit against $\mathcal{C}_{\text{mt}}$ is empty; the audit against $\mathcal{C}_{\text{std}}$ is the four-codon difference of the two published tables.

9 The Dual Cipher as a Complementary Experiment

Axiom A_8 supplies an oriented perimeter and a dual numbering of its sites. Nothing in Sections 5–8 used that axiom (Proposition 2.4, Remark 2.2). The present section records what the dual cipher is, defines the complementary experiment it names, and states what complementarity does not give. No second codon table is constructed. No matching theorem is repeated on a dual demand graph.

9.1 What A_8 supplies

Definition 9.1 (Dual cipher). Imported from A_8 of [1]. The oriented perimeter of the embedded node $\mathfrak{C}$ carries two complementary numberings of a twelve-site cycle,

$$\{3, 6, 9, 12\} \quad \text{and} \quad \{2, 5, 8, 11\}. \quad (42)$$

Each numbering selects a four-site reading. The two readings are the *cipher* and the *dual cipher*. Write Σ for the four-letter alphabet of the cipher under Definition 2.6, and Σ^* for the four-letter alphabet of the dual cipher, as a set. No identification of Σ^* with Σ is made yet.

Remark 9.2 (Two attachment sets). Axiom A_1 distinguishes two attachment sets. The cipher of Definition 2.6 reads one of them, through the binaries b_1, b_2 of [2, Def. 3.12]. The dual cipher is the reading of the other. That is the content of A_8 used below. The tensor $\Psi_{\mathfrak{C}}$, the crossed product $\mathfrak{A}_{\mathfrak{C}}$, and the spectral triple of A_9 – A_{11} are not used.

Proposition 9.3 (The dual is not a frame operation). *[A] Passage from the cipher to the dual cipher is not an element of Fr, and is not the biological reverse complement.*

Proof. Fr acts on words in a single alphabet Σ , letterwise through $\{\rho, \tau\}$. The dual cipher is a change of which four-site reading is taken as the alphabet, not an action on Σ^3 . The biological reverse complement is not a frame operation under any assignment (Proposition 4.19). The two maps are therefore distinct from each other and from every element of Fr. $\square$

9.2 The complementary experiment

Definition 9.4 (Complementary experiment). Let $\mathcal{E}$ be the identified experiment of Definition 4.15. The *complementary experiment* $\mathcal{E}^*$ has

- sample space $W^* := (\Sigma^*)^3$,
- class statistic $T^* : W^* \rightarrow X^*$ the orbit map of the copy of Fr acting on W^* ,
- block statistic B^* by first two letters of Σ^* ,
- parameter of interest a dual excision $\Xi^* : X^* \rightarrow R^*$,
- sampling law the counting measure on W^* .

The two experiments share the underlying node $\mathfrak{C}$ and the tether of H_{teth} . They do not share an alphabet unless an identification $\Sigma^* \cong \Sigma$ is supplied.

Proposition 9.5 (The dual count is twenty). *[A] If Σ^* is given the structure of a structured alphabet in the sense of Definition 4.4, then $|X^*| = 20$, of profile eight and twelve.*

Proof. Theorem 4.6 is a statement about structured alphabets, not about a particular chemical reading. It applies to $(\Sigma^*, \text{a bipartition})$ as soon as a bipartition is chosen. No such bipartition is chosen here beyond the existence claim. $\square$

Definition 9.6 (Pairing of experiments). A *pairing* of $\mathcal{E}$ with $\mathcal{E}^*$ is an identification $\iota : \Sigma^* \rightarrow \Sigma$ of structured

alphabets, inducing a bijection $W^* \rightarrow W$ and carrying (T^*, B^*) to a pair of statistics (T_{ι}, B_{ι}) on W . Under a pairing the two experiments are two readings of the same sixty-four-element set.

Remark 9.7 (No pairing is chosen). Definition 2.6 chose one assignment of binaries to letters. A pairing would choose a second, or would identify the two. Neither is done. The eight Watson–Crick assignments of Proposition 4.18 are pairings of the cipher with itself under complementation, not pairings of the cipher with the dual cipher. The distinction is Proposition 9.3.

9.3 Complementarity as ancillarity between experiments

The intended reading is that a statistic sufficient for the parameter of $\mathcal{E}^*$ is ancillary for the parameter of $\mathcal{E}$, and conversely. That reading is a target, not a theorem of the present constructions.

Target 9.8 (Cross-ancillarity). Let ι be a pairing. The statistic T_{ι} is conditionally ancillary for Ξ given T , and T is conditionally ancillary for $\Xi^* \circ \iota^{-1}$ given T_{ι} , in the sense of Definition 5.5.

Remark 9.9 (Why this is not Theorem 5.13). Theorem 5.13 is ancillarity of the within-class coordinate Φ_s for Ξ , inside a single experiment. Target 9.8 is ancillarity of one class statistic for the other experiment's parameter. The two statements have the same shape and different σ -algebras. Proving the target requires a pairing and a relation between Ξ and Ξ^* that Sections 5–8 do not supply.

Target 9.10 (Commutant reading). Under the representation of the sequence algebra $\mathfrak{A}_{\mathfrak{C}}$ on the Hilbert space of the spectral triple of [1], the observables of $\mathcal{E}^*$ generate a subalgebra of the commutant of the observables of $\mathcal{E}$.

Remark 9.11 (Status of the commutant). Target 9.10 uses A_7 , A_9 and A_{10} , which this paper has not used. It is recorded because A_8 is stated in [1] as a dual of readings, and the commutant is the operator-algebraic name of that duality. It is not an ingredient of Theorems 7.7, 8.8 or 8.13, and those theorems are not applied to it.

9.4 What the dual does not give

Proposition 9.12 (No field structure). *[A] The cardinality $|\Sigma| = 4$ does not equip Σ with the structure of the field $\mathbb{F}_4$. The same holds for Σ^* . No statement in this paper uses a multiplication on either alphabet.*

Proof. [1, Prop. 3.1]: four elements are the underlying set of exactly one field, up to isomorphism, but the field operations are an additional datum and are not supplied by A_1 – A_{12} . The present constructions use Σ as a structured alphabet, which is a set with a bipartition, not a field. $\square$

Proposition 9.13 (No second table). *[A] The complementary experiment does not determine an assignment $W^* \rightarrow R \cup \{*\}$. In particular Theorem 8.13 is not repeated for $\mathcal{E}^*$.*

Proof. Demand, Hall, and the audit use $\mathcal{C}_{\text{std}}$ and $\mathcal{C}_{\text{mt}}$, which are functions on $W = \Sigma^3$ under Definition 2.6. No published table is a function on W^* , and no demand graph has been defined on Γ^* . Existence of a dual refined assignment is not asserted. $\square$

Proposition 9.14 (No geometric orthogonality). *[A] Complementarity of the two ciphers is not an orthogonality of subspaces of a Hilbert space attached to $\mathfrak{C}$, and is not a statement about the Connes distance of A_{11} . Those readings require the spectral triple, which is unused.*

Proof. Definition 9.1 is a pair of numberings of a cycle. Orthogonality and the Connes distance are structures of A_9 – A_{11} . Remark 2.2 recorded that those axioms are unused in the matching, sufficiency, and audit arguments. They remain unused. $\square$

Remark 9.15 (Strand complementarity is a different dual). Remark 4.20 already recorded that a DNA:RNA duplex is an external datum and not a symmetry of the classification. Strand complementarity pairs letters inside Σ by ρ . The dual cipher pairs two readings of the perimeter. Using one as a name for the other would identify A_8 with a frame operation, which Proposition 9.3 forbids.

9.5 Why the section is here

Remark 9.16 (Ancillarity has two directions). Sections 5–8 treat ancillarity inside one experiment: the tether and the within-class coordinate are ancillary for excision; the block is not. The dual cipher is the corpus’s name for a second experiment on the same node, whose sufficient statistic would be ancillary for the first experiment’s parameter if Target 9.8 holds. Recording the target keeps that direction of the sufficiency language from being forgotten. Proving it is not required for the range theorem or the audit.

Remark 9.17 (What this section has fixed). The dual cipher is a second four-site reading of the perimeter, not an element of Fr and not reverse complementation. The complementary experiment is defined and not run. Cross-ancillarity and the commutant reading are targets. No field structure, no second codon table, and no geometric orthogonality are claimed. Nothing in Sections 5–8 is altered.

10 Parity as a Special Case and the Second Obstruction

[5, Prop. 5.9] is not a statement about absence numbers on blocks. It is a statement about fibres of maps that are constant on classes. This section imports that statement, records what it excludes and what it does not, and records that the refined assignment of Section 8 is not a class-functional map, so the obstruction does not apply to it. The unnamed obstruction that still prevents the mitochondrial table from being class-functional is Target 9.4 of [6] and is not identified here.

10.1 Capacity, imported

The production clauses of [5, §2.5] are used by name below and are restated so that they can be checked.

Hypothesis 10.1 (H_{prod}). Imported from [5, §2.5].

- (i) $H_{\text{prod}}(0)$: the product depends on the cluster’s marker assignment alone.
- (ii) $H_{\text{prod}}(0')$: the pathway does not involve the initiator.
- (iii) $H_{\text{prod}}(1)$: every class produces an amino acid.
- (iv) $H_{\text{prod}}(2)$: distinct classes produce distinct amino acids.

Theorem 10.2 (Capacity). *[D] Assume $H_{\text{prod}}(0)$ and $H_{\text{prod}}(0')$. A cluster of $2N$ sites produces at most $2^{2N-2} + 2^{N-1}$ distinct values; at $N = 3$, at most twenty. If in addition $H_{\text{prod}}(1)$ and $H_{\text{prod}}(2)$ hold, the products are amino acids and the number of distinct ones is exactly twenty.*

Proof. [5, Thm. 5.4]. The ceiling is the class count of Theorem 4.6 applied to a function on X . Saturation is injectivity and totality of that function. $\square$

Remark 10.3 (Ceiling and saturation are not of the same kind). The first clause is a bound. The second asserts that the system uses its capacity. [5, Rem. 5.5] records the distinction. This paper uses the ceiling. It does not use saturation, and $H_{\text{prod}}(1)$ is the reason three classes can remain unassigned after a covering seed of seventeen [6, Prop. 5.20].

Remark 10.4 (Capacity is not the obstruction). Pigeonhole from twenty residues to sixty-four words forces repetition. It does not force an odd fibre, and it does not force absence. Capacity alone does not exclude the standard table.

10.2 Parity

Definition 10.5 (Class-functional). A map $c : W \rightarrow R \cup \{*\}$ is *class-functional* when it is T -measurable: constant on each frame class. Equivalently, c factors through T .

Theorem 10.6 (Parity). *[A] Every frame class has size two or four. Hence every fibre of a class-functional map, pulled back to W , has even cardinality.*

Proof. [5, Prop. 5.9], from [3, Prop. 3.6(i)]: flip has no fixed point, so no class has size one. A fibre of a class-functional map is a union of classes, hence even. Restated as [6, Prop. 8.1]. $\square$

Corollary 10.7 (The standard table is not class-functional). *[A] The standard assignment sends methionine to AUG alone and tryptophan to UGG alone: two fibres of cardinality one. By Theorem 10.6 it is not class-functional. No operator satisfying $H_{\text{prod}}(0)$ and $H_{\text{prod}}(0')$ produces it.*

Proof. [5, Cor. 5.10], imported as [6, Cor. 8.2] (there cited as [5, Cor. 5.9]; the statement in [5] is Corollary 5.10). $\square$

Corollary 10.8 (A start codon is not class-functional). *[A] A uniquely specified start codon is a singleton fibre. No singleton fibre arises from a class-functional map. The start codon of the standard table therefore cannot arise in a class-functional phase.*

Proof. [6, Cor. 8.3]. The cost is Theorem 10.6 and nothing else. $\square$

Proposition 10.9 (Parity is silent on the mitochondrial table). *[A] The vertebrate mitochondrial assignment has fibre sizes twelve of two, seven of four and two of six. It has no singleton fibre and no odd fibre. Theorem 10.6 does not exclude it. It is nonetheless not class-functional: nineteen of the twenty classes carry more than one residue under it.*

Proof. [6, Prop. 8.5], by enumeration. $\square$

Remark 10.10 (What parity does not say). Parity does not say that no genetic code can be class-functional. It says that of the standard table. The mitochondrial table is a counterexample to the general reading [6, Rem. 8.6]. Both tables fail to be class-functional; only one of those failures is parity.

10.3 The refined assignment is not class-functional

Proposition 10.11 (Seeding frees the fibres). *[A] A refined assignment $c^\sharp$ of Definition 8.5 is not class-functional, except in the degenerate case where Ξ is constant and no block splits. Consequently Theorem 10.6 does not apply to $c^\sharp$.*

Proof. A size-four class meets four blocks (Corollary 6.3). Propagation and refinement place residues blockwise, so two words of the same class that land in different blocks need not carry the same value. That is Theorem 6.11. The fibres of $c^\sharp$ are therefore not unions of classes, and the even-cardinality constraint does not bind them. This is the observation of [6, §8]: the fibres are freed, so the obstruction does not apply. $\square$

Remark 10.12 (Why the seed crosses the gap). [5, Cor. 5.10] forbids a class-functional production of the standard table. The construction of Sections 7 and 8

is not class-functional. That is how it can output a table with singleton residues at the codon level without contradicting Theorem 10.6. The bound of twenty survives because every residue placed is still seeded by a class [6].

10.4 The two obstructions

Definition 10.13 (The two obstructions). The *first obstruction* is transversality (Theorem 6.2): T is not a sufficient reduction for B , so a seed is required. The *parity obstruction* is Theorem 10.6: a class-functional map cannot have an odd fibre, and therefore cannot equal the standard table. The *unnamed second obstruction* is whatever prevents the mitochondrial table from being class-functional, parity being silent on it (Proposition 10.9). That obstruction is Target 9.4 of [6] and is not identified here.

Proposition 10.14 (Hall does not cancel parity). *[A] Existence of a covering matching of demand (Theorem 7.7) is a statement about edges of a bipartite graph. Parity is a statement about fibres of class-functional maps. Neither implies the other. Hall is how the first obstruction is answered. It is not an answer to parity, because after Hall the map is no longer class-functional (Proposition 10.11).*

Proof. The two predicates quantify over different sets. $\square$

Theorem 10.15 (Sufficiency does not reach parity). *[A] Sufficiency of T for excision (Theorem 5.7) is the statement that Ξ is class-functional. Parity then applies to $\Xi \circ T$, and forbids that class-extension from equalling the standard table. Parity does not apply to the assignment c produced after a seed. Sufficiency of T for Ξ is therefore compatible with an assignment that the standard table realises and that $\Xi \circ T$ cannot.*

Proof. Corollary 10.7 applied to $c_{\Xi} = \Xi \circ T$; Proposition 10.11 applied to $c^{\sharp}$. $\square$

10.5 Termination, not absence-number

Corollary 10.16 (Termination is not an assembly quantity). *[A] The words UAG and GCU lie in one class of size two. A class-functional map cannot assign termination to one and a residue to the other. Termination, when it occurs inside a class that also carries*

a residue, is supplied by the reading and not by the cluster.

Proof. [5, Cor. 5.12]. The class is class 4 of [4, Tab. 1]. $\square$

Remark 10.17 (The absences on UA and AG). Proposition 8.16 records four absences on the purine halves of UA and AG as a fact about the refined assignment after liberation [6, Cor. 6.4]. That fact is not Theorem 10.6. Even absence numbers on blocks were a misreading of [5, Prop. 5.9] and are not used.

Remark 10.18 (H_{prod} is not H_{stop}). The production clauses constrain how a cluster produces a residue. They do not name UAA, UAG, or UGA. H_{stop} named three codons and was withdrawn (Remark 2.25). Parity allocates termination to the reading level (Corollary 10.16); it does not restore the withdrawn hypothesis.

Remark 10.19 (Parity is not the evenness of $|\text{Fr}|$). The frame group has even order, and every class has even size. The second of those is the input to Theorem 10.6. The first is not. Collapsing the two evennesses would make Corollary 4.11 look like a consequence of Theorem 10.6. It is not: residual state counts are orbit counts under subgroups of Fr , and do not mention fibres of an assignment.

Remark 10.20 (What this section has fixed). Capacity is a ceiling of twenty under $H_{\text{prod}}(0)$ and $(0')$. Parity is even fibre size of class-functional maps. It excludes the standard table and is silent on the mitochondrial table. The refined assignment is not class-functional, so parity does not apply to it. The unnamed obstruction that still splits nineteen mitochondrial classes is Target 9.4 of [6]. Even absence numbers on blocks are not a theorem of this paper.

11 Targets, Predecessors, and Conclusion

11.1 Targets left open

The following are not theorems of this paper. They are collected so that a later argument can cite a single list.

- (1) Chain polarity, H_{teth} (iii). Load-bearing for the identification of W with a sixty-four-codon domain (Re-

mark 4.12). Remains Target 9.7 of [6].

- (2) H_{seed} , including the pyrimidine-half clause on split blocks (Remark 7.19).
- (3) $H_{\text{split-pos}}$, as a rule of the construction. The published tables agree with it (Theorem 8.11); the agreement is not a derivation from $\mathcal{A}$.
- (4) $H_{\text{split-chem}}$. Size alone is refuted (Remark 8.18). No replacement rule is offered.
- (5) $H_{\text{lib-env}}$, for the environmental pair and for any liberation that is not T -measurable (Proposition 5.19, Remark 8.17).
- (6) Cross-ancillarity of the complementary experiment (Target 9.8).
- (7) The commutant reading of A_8 (Target 9.10).
- (8) The Identification Hypothesis Φ of [7]. Unused (Remark 2.26).
- (9) The unnamed second obstruction: what prevents the mitochondrial table from being class-functional [6, Target 9.4].

No target on this list is discharged by the empty audit of Proposition 8.20.

11.2 What was proved

The paper's own theorems, as against imported ones, are the reduction statements.

- T is a sufficient reduction of the word for excision, and (T, r) is minimal Neyman-sufficient for Ξ in the joint experiment (Theorems 5.7, 5.8, Proposition 5.9).
- The within-class coordinate is conditionally ancillary for Ξ ; the raw four-valued frame coordinate is not independent of T ; Basu is not available (Theorem 5.13, Proposition 5.14, Remark 5.15).
- Classes and blocks are transverse; (T, B) separates points; jointly measurable assignments are constant (Theorems 6.2, 6.8, Proposition 6.6).
- A seed is a section of incidence; propagation is the unique B -measurable extension of a non-conflicting seed (Definition 7.1, Theorem 7.12).
- Every refined assignment is in normal form; the standard table is not in that range (Theorem 8.8, Corollary 8.9).
- Parity excludes the standard table and is silent on the mitochondrial table; the refined assignment is not class-functional, so the obstruction does not apply to it (Theorem 10.6, Proposition 10.9, Proposition 10.11).

- The unnamed obstruction that still prevents the mitochondrial table from being class-functional is Target 9.4 of [6] and is not identified.

Imported under named hypotheses: the twenty-class count, the residual state counts, complementation as a frame operation, seventeen seeds, eight hundred and forty matchings, the mitochondrial equality, stops as absences on the purine halves of UA and AG, capacity.

11.3 Predecessors

The paper takes its carrier from [1, 2, 3], its partitions from [4], its capacity and parity from [5], and its seed, demand, refinement and audit from [6]. It adds no axiom. It uses the standard table constructively, which [5] did not, and says so at Declaration 2.7.

The review architecture against which the paper is intended to be read is that of [7] and of the orthogonal pipeline $M \cap C \cap H$: mathematical validity, concordance with the corpus, and fidelity to named hypotheses. Concordance failures to watch are the clauses of Proposition 3.1.

11.4 Conclusion

Excision is a function of the class. A codon table is a function of the word, organised by blocks. The two partitions of the sixty-four-element set are transverse, so the first function does not give the second. A seed is the missing section of incidence; propagation is the unique block-measurable extension; refinement is a third-position cut whose only patterns are (4) and (2, 2). Parity then excludes any class-functional copy of the standard table and does not apply to the refined assignment. Under four hypotheses and a demand graph read from the standard table, one refined assignment equals the vertebrate mitochondrial table. The equality is an equality of functions. The hypotheses remain hypotheses. The complementary experiment remains unrun.

A Tables and Computed Information Quantities

The objects below are functions of a sixty-four-element set and of two published tables. They are generated by `scripts/enumerate_w.py` from Definitions 2.6, 2.8 and 2.9 and from tables 1 and 2 of [8]. They are not transcribed by hand. Re-running the script is the verification.

The script takes `flip` to be letterwise Watson–Crick complementation and `rev` to be letterwise τ followed by reversal, with τ fixing A, U and swapping C, G . That is the action used in the proof of Theorem 6.2.

Proposition A.1 (Generated census). *[A] Under Definition 2.6, Fr acting by `flip` = letterwise complementation and `rev` = letterwise τ then reverse, there are 20 orbits: 8 of size two and 12 of size four. The incidence graph has two connected components, split by the Watson–Crick class of the middle letter.*

Proof. Direct enumeration of the sixty-four words. Regenerated by `scripts/enumerate_w.py`. □

Size	Representative	Orbit
2	AAA	AAA, UUU
2	ACU	ACU, UGA
2	AGU	AGU, UCA
2	AUA	AUA, UAU
2	CAG	CAG, GUC
2	CCC	CCC, GGG
2	CGC	CGC, GCG
2	CUG	CUG, GAC
4	AAC	AAC, CUU, GAA, UUG
4	AAG	AAG, CAA, GUU, UUC
4	AAU	AAU, AUU, UAA, UUA
4	ACA	ACA, AGA, UCU, UGU
4	ACC	ACC, CCU, GGA, UGG
4	ACG	ACG, CGA, GCU, UGC
4	AGC	AGC, CGU, GCA, UCG
4	AGG	AGG, CCA, GGU, UCC
4	AUC	AUC, CAU, GUA, UAG
4	AUG	AUG, CUA, GAU, UAC
4	CAC	CAC, CUC, GAG, GUG
4	CCG	CCG, CGG, GCC, GGC

Proposition A.2 (Entropies, regenerated). *[A] $H(T) = 4.2500$, $H(B) = 4.0000$, $H(T, B) = 6.0000$, $I(T; B) = 2.2500$ bits.*

Proof. Orbit profile 8 of probability $2/64$ and 12 of probability $4/64$; sixteen uniform blocks; sixty-four incident pairs of mass $1/64$. □

Proposition A.3 (Fibre profiles). *[A] The standard table has odd fibres $\ast(3)$, $I(3)$, $M(1)$, $W(1)$. The mitochondrial table has odd fibres none. The two tables disagree at AGA, AGG, AUA, UGA.*

Table	Pattern	Number of blocks	Blocks
standard	(2, 1, 1)	1	UG
standard	(2, 2)	6	AA, AG, CA, GA, UA, UU
standard	(3, 1)	1	AU
standard	(4)	8	AC, CC, CG, CU, GC, GG, GU, UC
mitochondrial	(2, 2)	8	AA, AG, AU, CA, GA, UA, UG, UU
mitochondrial	(4)	8	AC, CC, CG, CU, GC, GG, GU, UC

Remark A.4 (What the census corrects). An earlier draft of Proposition 6.7 asserted that Γ is connected. The census finds two components. The proof that both generators of Fr preserve the Watson–Crick class of the middle letter is in the revised Proposition 6.7; the census confirms that each side is internally connected. Theorem 6.8 was revised with the proposition. The item is recorded in Appendix B.

B Concordance, Ledger, and Verification

The pipeline is that of the orthogonal review $F_{\text{adm}} = M \cap C \cap H$: mathematical validity, concordance with the corpus, fidelity to named hypotheses. Items below are the defects raised in the drafting review and their disposition. Status: open, revised, or verified.

B.1 Defect ledger

#	Status	Item
1	revised	Parameter of interest in §2 was the realized value $\Xi(T(w))$. Reset to the unknown map Ξ .
2	revised	Introduction and abstract overstated minimal sufficiency of T . Aligned with Proposition 5.9: (T, r) is minimal; T is minimal among functions of the word.
3	revised	Four matching hypotheses were not distinguished from chain polarity. Distinguished at Theorem 8.13.
4	revised	* versus published * undeclared. Remark 2.15.
5	revised	Theorem 10.6 had been written as even absence numbers on blocks. Reset to [5, Prop. 5.9]: even fibres of class-functional maps.
6	revised	Stops on UA and AG had been attributed to that proposition. They are [6, Cor. 6.4].
7	revised	[5] was cited as Cor. 5.9 for the standard-code exclusion. The statement in [5] is Cor. 5.10; [6] calls it Cor. 5.9. Drift recorded at Corollary 10.7.
8	revised	Proposition 6.7 asserted that Γ is connected. Census of Appendix A finds two components. Theorem 6.8 revised.
9	verified	$H(T) = 4.25$, $I(T; B) = 2.25$ bits under counting measure, base two. Regenerated in Appendix A.
10	verified	Standard odd fibres: Met 1, Trp 1, Ile 3, * 3. Mitochondrial: none. Disagreement set {AUA, UGA, AGA, AGG}.
11	verified	Standard patterns: eight of (4), six of (2, 2), one of (3, 1), one of (2, 1, 1). Mitochondrial: eight of (4), eight of (2, 2).
12	revised	Demand-from- $\mathcal{C}_{\text{std}}$, output- $\mathcal{C}_{\text{mt}}$: Remark 8.14 now records the mechanical step. Existence remains [6, Thm. 5.8].
13	revised	Invented V-tags pointing at [6] removed. Remaining V-tags refer to Appendix A of this paper.
14	revised	Generating argument for each component of Γ written in the proof of Proposition 6.7. Census remains confirmatory.

B.2 Verification record

V1. Orbit count under the present assignment: twenty, profile eight and twelve. Script, Appendix A.

V2. Incidence degrees: class degree equals class size; every block has degree four. Script.

V3. Two incidence components, split by the Watson–Crick class of the middle letter. Script and Proposition 6.7.

V4. Entropies of Proposition 6.12. Script.

V5. Fibre profiles of $\mathcal{C}_{\text{std}}$ and $\mathcal{C}_{\text{mt}}$. Script.

V6. Block patterns of both tables. Script.

V7. Disagreement set of the two tables. Script.

These seven checks verify statements of *this* paper. They do not re-verify Hall, the eight hundred and forty maps, or the mitochondrial equality; those remain imported from [6]. Draft tags that pointed at [6] without a matching number have been removed.

B.3 Hypotheses still live

Unchanged from Remark 8.23, together with Target 9.4 of [6] (the unnamed obstruction) and Target 9.8, Target 9.10 of this paper. Φ unused. H_{stop} withdrawn.

C Established Literature

Borrowed machinery used by name and kept distinct from the organisation of this paper: Hall's theorem [10]; Basu's theorem [11], cited only to record that it does not apply; Fisher–Neyman factorisation on a finite set [12]; residue volumes [9]; translation tables [8]. Categorical and spectral sources [13, 14] are ambient and unused in the matching.

References

- [1] C. D. Schaper. The Chrysene Formalism Revisited: A Twelve-Axiom Foundation for Molecular Characterization, Tensorial Noncommutative Geometry, and Operator Response. *Annals of the Chrysene Formalism*, 1(2):476–547, 2026.
- [2] C. D. Schaper. The Encoded Stack. *Annals of the Chrysene Formalism*, 1(2):646–675, 2026.
- [3] C. D. Schaper. The Twenty Registry Classes of a Chrysene Stack. *Annals of the Chrysene Formalism*, 1(2):676–704, 2026.
- [4] C. D. Schaper. Coarsenings of the Genetic Code by the Twenty Registry Classes of a Chrysene Stack. *Annals of the Chrysene Formalism*, 1(2):730–767, 2026.
- [5] C. D. Schaper. Why Three and Why Twenty: The Capacity of a Chrysene Cluster, the Base-Four Intermediary, and the Production Operator Δ . *Annals of the Chrysene Formalism*, 1(2):768–793, 2026.
- [6] C. D. Schaper. Seeded Assignments on the Chrysene Registry: Transversality, Third-Position Refinement, and the Mitochondrial Code as Normal Form. *Annals of the Chrysene Formalism*, 1(2):858–882, 2026.
- [7] C. D. Schaper. The Chrysene Formalism as Candidate Mathematics and Candidate Biogenesis. *Annals of the Chrysene Formalism*, 1(2):476–488, 2026.
- [8] A. Elzanowski and J. Ostell. The Genetic Codes. National Center for Biotechnology Information, Bethesda, MD. Last update 23 September 2024.
- [9] A. A. Zamyatnin. Protein volume in solution. *Progress in Biophysics and Molecular Biology*, 24:107–123, 1972.
- [10] P. Hall. On representatives of subsets. *Journal of the London Mathematical Society*, s1-10(1):26–30, 1935.
- [11] D. Basu. On statistics independent of a complete sufficient statistic. *Sankhyā*, 15(4):377–380, 1955.
- [12] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222:309–368, 1922.

- [13] S. Mac Lane. *Categories for the Working Mathematician*. GTM 5. Springer, 2nd ed., 1998.
- [14] A. Connes. *Noncommutative Geometry*. Academic Press, 1994.