hexagon:2611.00240v1

A Parity Theorem for Reverse-Complement-Invariant Minimum Decycling Sets in q-ary de Bruijn Graphs

Oleksiy Babanskyy

Abstract

Let $\mathcal{A}$ be a finite alphabet with a fixed-point-free involutive complement, and let reverse complement act on the directed de Bruijn graph $B(q,k)$ by reversing a word and complementing every letter. We prove that a reverse-complement-invariant ordinary minimum decycling set exists if and only if k is even, for every integer $k \geq 2$. This simultaneously covers the binary alphabet and the DNA alphabet with $A \leftrightarrow T$ and $C \leftrightarrow G$.

For odd k, a primitive rotation cycle joins two reverse-complement-fixed lower states in opposite directions. Disjoint complementary alphabet pairs force at least q deletions beyond the ordinary minimum. For even k, choose injective antisymmetric real weights on the alphabet and form a weighted cyclotomic selector. We prove that arbitrary representatives on its zero-embedding rotation cycles preserve acyclicity, resolve those ties equivariantly, and establish the exact defect identity

$$\mathbf{1}_{\mathrm{RC}(M)} - \mathbf{1}_M = \nabla\chi.$$

A finite sequence of legal full-alphabet source firings removes this nonnegative symmetric defect two coordinates at a time. We also identify a self-dual spectral ideal that gives a deterministic direct selector, rather than a simulated firing path. For integer weights, membership is decidable with polynomial bit complexity. The executable finite panel in the repository checks the binary specialization only; it is convention-level assurance and is not used in the all-order proof.

Contents

1 Introduction **2**

2 Normalization and prior work **3**
2.1 Words, graphs, and reverse complement 3
2.2 Pure cycling registers and the ordinary minimum 4
2.3 Full-alphabet source firings 4
2.4 Scope of the imported results 4

3 The odd-order obstruction **5**

4 The weighted cyclotomic seed and zero-sector ties **5**
4.1 The selector 6
4.2 Equivariant zero ties 7

5 The half-plane coboundary **7**

6 Legal defect descent and even existence **7**

7 A direct spectral construction **9**

8 The q-ary parity theorem **10**

9 Exact finite assurance and reproducibility **11**

10 Boundaries and outlook **12**

Use of generative AI **12**

A Convention dictionary **13**

B Stable PCRs and exact algebraic signs **13**
B.1 Stable rotation cycles 13
B.2 Polynomial-bit exact sign determination 14

C Finite validation boundaries **15**

1 Introduction

Minimum decycling sets in directed de Bruijn graphs are classical combinatorial objects and underlie several finite-window sequence-sampling constructions. Their ordinary minimum is governed by rotation cycles and unavoidable sets [5, 1]. More recent work made the move structure of minimum decycling sets explicit [3]. Reverse-complement symmetry imposes a different constraint: whenever a word is chosen, its reverse complement must also be chosen. This is the natural strand symmetry for DNA and was posed directly in the k-nonical setting [4].

In Section 5.1 of [4], the authors identify the lack of known algorithms for symmetric MDSs; Section 6 also notes the unknown symmetric minimum size. We answer the following existence part of that problem: for which word lengths can literal reverse-complement symmetry attain the ordinary minimum? The answer below classifies every fixed-point-free complemented alphabet. It does not determine the exact odd-order constrained minimum or solve the additional path-length optimization in that work.

Theorem 1.1 (Parity theorem)**.** *Let $\mathcal{A}$ be a finite nonempty alphabet with a fixed-point-free involutive complement, and put $q = |\mathcal{A}|$. For every integer $k \geq 2$, the directed de Bruijn graph $B(q, k)$ has a reverse-complement-invariant ordinary minimum decycling set if and only if k is even.*

In particular, Theorem 1.1 applies to binary words and to DNA words under $A \leftrightarrow T$ and $C \leftrightarrow G$. The case $k = 1$ is deliberately outside the theorem: every loop must be deleted, so the unique ordinary minimum is the whole alphabet and is invariant even though k is odd.

The two directions use different mechanisms. At odd order, a primitive pure-rotation cycle has two reverse-complement-fixed lower endpoints. One deletion from that cycle cannot destroy both endpoint-to-endpoint directions, and reverse complement supplies the opposing residual path. Counting this obstruction over disjoint alphabet pairs gives at least q additional deletions. At even order, we assign distinct nonzero real weights $s(a)$ satisfying $s(\bar{a}) = -s(a)$. A weighted cyclotomic version of the classical selector then has an exact reverse-complement defect equal to the gradient of an open half-plane indicator. Legal full-alphabet source firings remove that defect monotonically.

The proof yields two complementary constructions. The first is a finite descent and exposes the move-theoretic mechanism; every allowed monotone scheduler uses an explicitly

determined number of firings. The second cuts the same self-dual residual poset at a spectral midpoint and gives a direct word-membership rule. When the weight table is integer-valued, exact real-algebraic sign determination makes the latter a polynomial-bit algorithm. Neither statement supplies a bound on the longest residual path or a biological performance guarantee.

Relation to prior terminology. The phrase "symmetric decycling set" has also been used for the ordinary Mykkeltveit construction based on opposite geometric half-planes [6]. We use the unambiguous phrase *reverse-complement-invariant* throughout.

Evidence discipline. The all-order proof occupies Sections 3–8. Section 9 records an independent exact implementation of the binary specialization only as a sign, convention, and mutation check. No finite census is a premise of Theorem 1.1.

Organization. Section 2 fixes the graph, rotation, and gradient conventions. Section 3 proves the odd obstruction. Sections 4–6 construct and symmetrize the weighted seed. Section 7 gives the direct spectral selector, and Section 8 assembles the parity theorem and its DNA corollary. The appendices collect convention maps, technical details, and the finite binary assurance table.

> **Claim boundary.** The mathematical theorem is alphabet-general under a fixed-point-free complement. We do not claim an exact odd-order symmetry premium, an optimal firing path, a quantitative window guarantee, biological performance, or blanket novelty or priority over unexamined sources.

2 Normalization and prior work

2.1 Words, graphs, and reverse complement

Let $\mathcal{A}$ be a finite nonempty alphabet of cardinality q, equipped with a fixed-point-free involution $a \mapsto \bar{a}$. Thus q is even. The binary specialization has $\bar{0} = 1$, and the DNA specialization has $\bar{A} = T$, $\bar{C} = G$. For $w = w_0 \cdots w_{k-1}$ define

$$\mathrm{RC}(w)_i = \overline{w_{k-1-i}}.$$

The directed vertex de Bruijn graph $B(q,k)$ has vertex set $\mathcal{A}^k$ and an arc $u \to v$ exactly when $u_1 \cdots u_{k-1} = v_0 \cdots v_{k-2}$. Loops are retained. Reverse complement is a directed anti-automorphism:

$$u \to v \quad \Longrightarrow \quad \mathrm{RC}(v) \to \mathrm{RC}(u).$$

A set $M \subseteq \mathcal{A}^k$ is a *decycling set* when $B(q,k) \setminus M$ is a directed acyclic graph (DAG), and is an *ordinary minimum decycling set* when its cardinality is minimum among all decycling sets. We abbreviate the latter to MDS. Reverse-complement invariance means $\mathrm{RC}(M) = M$ as a literal set of labelled words.

The line-graph identity $B(q,k) = L(B(q,k-1))$ lets us regard each length-k word as the labelled lower-state arc

$$\mathrm{t}(w) = w_0 \cdots w_{k-2} \longrightarrow \mathrm{h}(w) = w_1 \cdots w_{k-1}.$$

Deleting M as vertices of $B(q,k)$ is equivalent to deleting those labelled arcs from $B(q,k-1)$; acyclicity is equivalent in the two representations.

2.2 Pure cycling registers and the ordinary minimum

Let $R(w_0 \cdots w_{k-1}) = w_1 \cdots w_{k-1} w_0$. A pure cycling register (PCR) is an orbit under R, using distinct rotations only. Its cardinality is the minimal period of its words. PCRs partition $\mathcal{A}^k$, and each is a directed cycle. Their number is

$$N_q(k) = \frac{1}{k} \sum_{d \mid k} \varphi(d) q^{k/d}. \tag{1}$$

Every decycling set meets every PCR. The matching construction for arbitrary finite alphabets is the unavoidable-set theorem of Champarnaud–Hansel–Perrin [1]; the original binary construction is due to Mykkeltveit [5]. Hence the ordinary minimum cardinality is $N_q(k)$, and an MDS selects exactly one distinct word from each PCR. The converse is false: one-per-PCR is only a local cardinality condition and can leave cycles using several PCRs. Every homopolymer a^k is a singleton PCR loop and is a mandatory selection.

2.3 Full-alphabet source firings

For $u \in \mathcal{A}^{k-1}$ write

$$L(u) = \{au : a \in \mathcal{A}\}, \qquad R(u) = \{ua : a \in \mathcal{A}\}.$$

A forward firing, or full-alphabet F-move, is valid at u for a selected set M when $L(u) \subseteq M$, and then

$$F_u(M) = (M \setminus L(u)) \cup R(u).$$

These are literal sets. The two companion sets overlap only at a homopolymer loop, which is removed and reinserted. A valid F-move preserves ordinary-MDS status; this is the standard move of [3, Section 4.1]. Indeed, before the move the residual lower graph has no arc entering u, whereas afterward it has no arc leaving u. Any newly created directed cycle would have to enter u along a restored arc and then leave through a deleted one, which is impossible. The result is therefore decycling. If a newly inserted outgoing companion other than the common loop had already belonged to M, the result would be a smaller decycling set, contradicting ordinary minimality; hence cardinality is preserved.

For an integer function f on lower states, fix the gradient convention

$$(\nabla f)(w) = f(\mathrm{t}(w)) - f(\mathrm{h}(w)). \tag{2}$$

The indicator update of a valid firing is $\nabla \mathbf{1}_u$, including at a homopolymer loop where both sides vanish.

2.4 Scope of the imported results

We import the ordinary cardinality theorem and the standard F-move terminology. We prove move preservation above, as well as the odd obstruction, the weighted arbitrary-zero-tie acyclicity theorem, equivariant zero ties, the exact coboundary, the terminating descent, and the direct midpoint selector. The k-nonical paper supplies a dated formulation of the reverse-complement-constrained problem [4]; it does not supply the parity theorem or the construction below. Our graph-theoretic result does not replace that paper's optimization, implementation, or window-guarantee analysis.

3 The odd-order obstruction

Theorem 3.1 (Odd obstruction)**.** *For every odd $k = 2m+1 \geq 3$, every reverse-complement-invariant decycling set M in $B(q,k)$ satisfies*

$$|M| \geq N_q(k) + q.$$

In particular, ordinary minimum cardinality is impossible.

Proof. Let M be such a set. Fix a complementary pair $a, \bar{a}$ and define the length-$2m$ lower states

$$f_m = (a\bar{a})^m, \qquad g_m = (\bar{a}a)^m.$$

Both are fixed by reverse complement. The length-k word

$$w_m = (a\bar{a})^m\bar{a}$$

contains m copies of a and $m+1$ copies of $\bar{a}$. If it were a proper r-fold power, then $r > 1$ would divide both counts, contradicting $\gcd(m, m+1) = 1$. Its PCR therefore has full length k.

In the lower-state graph this PCR contains the direct arc $g_m \to f_m$ and, through its other $k-1$ arcs, a directed path from f_m to g_m. Denote this PCR by C_a. Every decycling set meets C_a. If M selected exactly one arc of C_a, deleting the direct arc would leave the long path, while deleting any other arc would leave the direct path. Thus one endpoint-to-endpoint direction would survive.

By reverse-complement invariance of M, that surviving path would have an image directed from the other fixed endpoint back to the first. Their concatenation is a nonempty closed directed walk and contains a directed cycle, a contradiction.

Consequently $|M \cap C_a| \geq 2$. The mate PCR $\mathrm{RC}(C_a)$ contains the same number of selected words. These two PCRs are distinct: their counts of a and $\bar{a}$ are interchanged. Repeating the argument for each of the $q/2$ complementary alphabet pairs gives q distinct PCRs, each requiring at least two selected words. PCRs from different alphabet pairs are disjoint because their words use disjoint sets of letters. Every other PCR still requires at least one selected word. Since PCRs partition $\mathcal{A}^k$,

$$|M| \geq 2q + (N_q(k) - q) = N_q(k) + q.$$

□

The proof does not determine the exact odd excess; no formula inferred from small odd orders is used below.

4 The weighted cyclotomic seed and zero-sector ties

Assume for this section that $k \geq 4$ is even. Choose an injective map

$$s : \mathcal{A} \longrightarrow \mathbb{R} \setminus \{0\}, \qquad s(\bar{a}) = -s(a).$$

Such a map exists by assigning distinct positive magnitudes to the complement pairs. For DNA one may use $s(A) = -2, s(C) = -1, s(G) = 1, s(T) = 2$. The existence, descent, and direct-selector arguments use only real arithmetic, signs, and injectivity. Integrality will

be imposed separately in Corollary 7.4, where a finite exact encoding and a bit complexity bound are required. Put $\zeta = e^{2\pi i/k}$ and, for $w = w_0 \cdots w_{k-1}$, define

$$P_s(w) = \sum_{j=0}^{k-1} s(w_j)\zeta^{j+1}, \qquad W_s(w) = \sum_{j=0}^{k-1} s(w_j).$$

For $u = u_0 \cdots u_{k-2}$ define

$$\lambda_s(u) = \sum_{j=0}^{k-2} s(u_j) \sin\left(\frac{2\pi(j+1)}{k}\right), \qquad \chi(u) = \mathbf{1}[\lambda_s(u) > 0].$$

Lemma 4.1 (Weighted covariance)**.** *For every length-k word w,*

$$P_s(Rw) = \zeta^{-1}P_s(w), \quad P_s(\mathrm{RC}\, w) = -\overline{P_s(Rw)}, \quad P_s(R\,\mathrm{RC}\, w) = -\overline{P_s(w)}.$$

Moreover

$$\operatorname{Im} P_s(w) = \lambda_s(\mathrm{t}(w)), \qquad \operatorname{Im} P_s(Rw) = \lambda_s(\mathrm{h}(w)),$$

and $\lambda_s(\mathrm{RC}\, u) = \lambda_s(u)$ on lower words.

Proof. The rotation identity follows by reindexing and $\zeta^k = 1$. Antisymmetry of s gives

$$P_s(\mathrm{RC}\, w) = -\sum_j s(w_j)\zeta^{-j} = -\overline{P_s(Rw)}.$$

Rotating gives the third identity. The omitted final letter of $\mathrm{t}(w)$ multiplies $\zeta^k = 1$ and contributes no imaginary part; apply the same observation to Rw for the head. Finally, reindexing the lower sum, using antisymmetry and $\sin(2\pi(k-r)/k) = -\sin(2\pi r/k)$, proves lower-word invariance. □

4.1 The selector

On each PCR with nonzero embedding select one word by the rule

1. select the point on the negative real axis, if it occurs;
2. otherwise select the unique w satisfying $\operatorname{Im} P_s(w) < 0 < \operatorname{Im} P_s(Rw)$.

On a PCR where P_s vanishes identically, select an arbitrary representative for now. Call the resulting set M_0 and write $m_0 = \mathbf{1}_{M_0}$.

Lemma 4.2 (Full period and uniqueness)**.** *Every nonzero embedded PCR has full period k, and the selector above is unique.*

Proof. If a word in a nonzero PCR had period $d < k$, then $P_s(R^d w) = \zeta^{-d}P_s(w) = P_s(w)$ would force $P_s(w) = 0$. The nonzero rotations therefore form a nondegenerate regular k-gon. Since k is even, a negative-real vertex is unique when present; if no real vertex is present, clockwise rotation gives exactly one strict lower-to-upper crossing. □

Theorem 4.3 (Arbitrary-zero-tie acyclicity)**.** *Every choice of one representative on each zero-embedding PCR produces an ordinary MDS M_0.*

Proof. Set $Q_s = P_s - W_s$. If $y = w_1 \cdots w_{k-1}a$ follows $w = w_0 \cdots w_{k-1}$ and $\delta = s(a) - s(w_0)$, then

$$P_s(y) = \zeta^{-1}P_s(w) + \delta, \qquad W_s(y) = W_s(w) + \delta, \qquad Q_s(y) = -W_s(w) + \zeta^{-1}P_s(w).$$

A residual path cannot pass from $\operatorname{Im} P_s \le 0$ to $\operatorname{Im} P_s > 0$: the strict crossing and departure from the negative ray are selected. Hence a residual cycle lies wholly in the open upper or closed lower half-plane.

Write $\theta = 2\pi/k$ and $P_s(w) = \rho e^{i\phi} \neq 0$. Along an edge whose endpoints remain in the open upper region,

$$\operatorname{Re} Q_s(y) - \operatorname{Re} Q_s(w) = 2\rho \sin(\theta/2) \sin(\phi - \theta/2) > 0. \tag{3}$$

Along a residual edge in the closed lower region the same expression is strictly negative: the negative-ray starting point was selected and the remaining angular interval is $\pi + \theta \le \phi \le 2\pi$. Thus a residual cycle could only consist of edges beginning at $P_s = 0$.

If consecutive labels w and y both have zero embedding, then $P_s(y) = s(a) - s(w_0) = 0$. Injectivity of s gives $a = w_0$, so the edge is a pure rotation. Such a cycle is exactly a zero-embedding PCR, from which one word was selected. This final contradiction proves acyclicity. Since M_0 meets every PCR exactly once, it has cardinality $N_q(k)$. □

4.2 Equivariant zero ties

Reverse complement permutes the zero PCRs. On a paired orbit $\{C, \mathrm{RC}(C)\}$, choose $w \in C$ and $\mathrm{RC}(w)$ in its mate. On a stable PCR, choose one of its two reverse-complement-fixed rotations; the alphabet-general fixed-rotation lemma is proved in Appendix B. From now on M_0 denotes a seed with these equivariant zero ties. The zero-sector part of m_0 is then literally reverse-complement invariant.

5 The half-plane coboundary

Theorem 5.1 (Exact reverse-complement defect)**.** *For every length-k word w,*

$$m_0(\mathrm{RC}\, w) - m_0(w) = \chi(\mathrm{t}(w)) - \chi(\mathrm{h}(w)) = (\nabla\chi)(w). \tag{4}$$

Equivalently, $\mathbf{1}_{\mathrm{RC}(M_0)} - \mathbf{1}_{M_0} = \nabla\chi$.

Proof. Put $A = \operatorname{Im} P_s(w)$ and $B = \operatorname{Im} P_s(Rw)$. Lemma 4.1 sends the ordered imaginary coordinates of $\mathrm{RC}(w)$ to (B, A). If A and B are nonzero, the selector rule gives

$$m_0(\mathrm{RC}\, w) - m_0(w) = \mathbf{1}[B < 0 < A] - \mathbf{1}[A < 0 < B] = \mathbf{1}[A > 0] - \mathbf{1}[B > 0].$$

The ray cases fix the sign convention. If $A = 0$, write $P_s(w) = r \in \mathbb{R}$. Then $B = -r \sin(2\pi/k)$: for $r < 0$ both sides are -1, and for $r > 0$ both are zero. If $B = 0$, write $P_s(Rw) = r \in \mathbb{R}$: for $r > 0$ both sides are $+1$, and for $r < 0$ both are zero. When $k > 2$, simultaneous vanishing of A and B with $P_s(w) \neq 0$ is impossible. In the zero sector, the chosen ties are equivariant, while both lower imaginary coordinates vanish.

Finally, Lemma 4.1 identifies the sign difference with $\chi(\mathrm{t}(w)) - \chi(\mathrm{h}(w))$ under convention (2). □

The open upper half-plane and the negative-real-axis convention are load-bearing. Equation (4) is a literal edgewise identity, not an averaged symmetry statement.

6 Legal defect descent and even existence

Reverse complement carries an ordinary MDS to an ordinary MDS because it reverses directed cycles bijectively.

Lemma 6.1 (Positive-potential firing)**.** *Let A and N be ordinary MDSs and suppose*

$$\mathbf{1}_N - \mathbf{1}_A = \nabla f$$

for an integral potential $f \geq 0$ with minimum zero. If $A \neq N$, some lower state u with $f(u) > 0$ admits a valid full-alphabet F-move from A.

Proof. Assume otherwise. At every positive state u, failure of the source condition supplies an incoming residual arc $v \to u$. It may be chosen nonloop: a residual loop would already contradict that A is decycling. Since this arc is not in A,

$$\mathbf{1}_N(v \to u) = f(v) - f(u) \in \{0, 1\},$$

so $f(v) \geq f(u) > 0$. Iterating the predecessor choice within the finite positive support eventually produces a directed residual cycle, a contradiction. □

Lemma 6.2 (Defect-paired update)**.** *Suppose an ordinary MDS A satisfies*

$$\mathbf{1}_{\mathrm{RC}(A)} - \mathbf{1}_A = \nabla f, \qquad f \geq 0, \qquad f \circ \mathrm{RC} = f, \qquad \min f = 0.$$

If A is not invariant, fire an enabled positive state u supplied by Lemma 6.1. The new MDS $A' = F_u(A)$ satisfies the same conditions with

$$f' = f - \mathbf{1}_u - \mathbf{1}_{\mathrm{RC}(u)}.$$

Moreover $u \neq \mathrm{RC}(u)$ and $\sum f' = \sum f - 2$.

Proof. The physical firing gives $\mathbf{1}_{A'} - \mathbf{1}_A = \nabla \mathbf{1}_u$. Reverse complement reverses lower arcs, hence sends this gradient to $-\nabla \mathbf{1}_{\mathrm{RC}(u)}$. Subtracting the update from its image gives

$$\mathbf{1}_{\mathrm{RC}(A')} - \mathbf{1}_{A'} = \nabla\big(f - \mathbf{1}_u - \mathbf{1}_{\mathrm{RC}(u)}\big).$$

The lower length $k-1$ is odd, so a reverse-complement-fixed lower word would have a central letter fixed by complement; therefore $u \neq \mathrm{RC}(u)$. Symmetry of f makes both decremented coordinates positive. Nonnegativity, symmetry, and minimum zero are preserved, and the total decreases by two. □

Theorem 6.3 (Even-order descent)**.** *For every even $k \geq 4$, the seed M_0 reaches a literal reverse-complement-invariant ordinary MDS by legal full-alphabet F-moves. Every scheduler that repeatedly uses Lemma 6.1 executes exactly*

$$\frac{1}{2}\big|\{u \in \mathcal{A}^{k-1} : \lambda_s(u) > 0\}\big| \tag{5}$$

physical firings.

Proof. Theorem 5.1 initializes the symmetric nonnegative potential $f = \chi$. Its minimum is zero: every homopolymer lower word a^{k-1} has

$$\lambda_s(a^{k-1}) = s(a) \sum_{j=1}^{k-1} \sin(2\pi j/k) = 0.$$

While the current MDS is not invariant, Lemmas 6.1–6.2 supply another legal move and decrease $\sum f$ by two. The process terminates.

At an invariant terminal, $\nabla f = 0$. The lower de Bruijn graph is strongly connected—append the letters of any target word—so f is constant. Its preserved minimum is zero, hence $f = 0$. Conversely the potential cannot vanish before the indicator defect does. Starting from the 0–1 potential χ therefore gives exactly (5) moves. □

The theorem asserts neither uniqueness of the terminal, confluence, nor shortest distance among unrestricted firing paths. A "paired update" is one physical firing whose defect potential changes at two lower states, not two simultaneous firings.

7 A direct spectral construction

The descent proves existence but need not be simulated. This section selects a particular self-dual ideal of positive lower states and thereby gives a direct terminal formula.

Fix any total order on $\mathcal{A}$ and use its lexicographic extension to lower words. Write $\vartheta(u) = \mathrm{RC}(u)$ and define

$$\psi_s(u) = \sum_{j=0}^{k-2} s(u_j)(\zeta^{j+1} - 1), \qquad r_s(u) = \operatorname{Re} \psi_s(u), \qquad U = \{u : \lambda_s(u) > 0\}.$$

The final coefficient of $P_s(ua) - W_s(ua)$ is $s(a)(\zeta^k - 1) = 0$, so

$$\psi_s(u) = P_s(ua) - W_s(ua) \tag{6}$$

for every appended letter a.

Define $a_\star : \mathcal{A}^{k-1} \to \{0, 1\}$ by

$$a_\star(u) = 1 \quad \Longleftrightarrow \quad u \in U \ \text{ and } \ \big(r_s(u) < 0 \ \text{ or } \ [r_s(u) = 0 \ \text{ and } u < \vartheta(u)]\big). \tag{7}$$

Lemma 7.1 (Self-dual spectral ideal)**.** *The set $A_\star = \{u : a_\star(u) = 1\}$ is an order ideal in the residual seed graph induced by U, and*

$$a_\star(u) + a_\star(\vartheta u) = \chi(u) \tag{8}$$

for every lower word u.

Proof. Reindexing as in Lemma 4.1 gives

$$\psi_s(\vartheta u) = -\overline{\psi_s(u)}.$$

Thus $\lambda_s(\vartheta u) = \lambda_s(u)$ and $r_s(\vartheta u) = -r_s(u)$. Since $k - 1$ is odd and the complement has no fixed letter, ϑ has no fixed lower word. Formula (7) therefore chooses exactly one member of every ϑ-pair in U, proving (8).

Now let $w : u \to v$ be an edge with $u, v \in U$. Both consecutive imaginary coordinates are positive, so w is not selected by the seed. From (6) and rotation covariance,

$$\psi_s(v) - \psi_s(u) = (\zeta^{-1} - 1)P_s(w). \tag{9}$$

Writing $P_s(w) = \rho e^{i\phi}$ and $\alpha = 2\pi/k$, the two positive imaginary coordinates imply $\alpha < \phi < \pi$. Hence

$$r_s(v) - r_s(u) = 2\rho \sin(\alpha/2) \sin(\phi - \alpha/2) > 0.$$

If $v \in A_\star$, then $r_s(v) \le 0$, so every predecessor u has $r_s(u) < 0$ and lies in $A_\star$. This is the ideal property. Strictness also shows that no residual edge joins two zero-level states, so the lexicographic tie in (7) cannot spoil ideality. □

Theorem 7.2 (Direct spectral MDS)**.** *For every even $k \ge 4$, the function*

$$m_\star(w) = m_0(w) + a_\star(\mathrm{t}(w)) - a_\star(\mathrm{h}(w)) \tag{10}$$

is the indicator of a literal reverse-complement-invariant ordinary MDS.

Proof. Fire the states of $A_\star$ in any predecessor-first order. Consider an incoming edge $x \to u$ when u is to be fired. If it was selected by M_0, nothing is required. Otherwise it is an edge of the residual seed graph. If $x \notin U$, the defect identity (4) and Booleanity would force that edge to be selected by M_0, a contradiction. Thus $x \in U$; ideality puts x in $A_\star$, and it was fired earlier, selecting its outgoing edge to u. No state of U is a homopolymer, since the complete sine sum is zero there. Thus all incoming companions of u are selected and the full-alphabet firing is legal.

The cumulative indicator update is exactly (10); in particular it is 0–1 valued and remains an ordinary MDS. For $w : u \to v$, equation (8) gives

$$(\nabla a_\star)(\mathrm{RC}\, w) = (\nabla a_\star)(w) - (\nabla \chi)(w).$$

Combining this with (4) yields $m_\star(\mathrm{RC}\, w) = m_\star(w)$. □

The construction is direct in a precise sense: membership of a queried word uses the seed rule and the two endpoint values in (7), not a materialized list of firings or of q^{k-1} lower states. For a deterministic seed, order PCRs by their lexicographically least word. In each paired zero-PCR orbit, select the least word of the smaller PCR and its reverse complement in the mate; on a stable zero PCR select the least of the two fixed rotations.

Corollary 7.3 (Real-weight family)**.** *For every injective map $s : \mathcal{A} \to \mathbb{R} \setminus \{0\}$ satisfying $s(\bar{a}) = -s(a)$, the seed construction produces an ordinary MDS, and both the legal defect descent and direct selector produce reverse-complement-invariant ordinary MDSs at every even $k \geq 4$. No integrality hypothesis on s is needed for these existence statements.*

Proof. Weighted covariance uses only real linearity, antisymmetry, and complex conjugation. The zero-sector argument additionally uses injectivity to deduce equality of letters from equality of weights. The coboundary uses only signs, while its nonnegative integral potential is the indicator χ, not the alphabet weight map. Finally, the direct construction uses the real order of λ_s and r_s. Thus every preceding step remains valid for the stated real weights. □

Corollary 7.4 (Exact membership complexity)**.** *Given the complement table, the integer weights, and an even length $k \geq 4$, a queried word's membership in the MDS of Theorem 7.2 is decidable with preprocessing time, query time, and working memory polynomial in k, q, and the bit length of the weight table. The algorithm uses exact arithmetic and stores no object of size q^{k-1} or q^k.*

Proof. Appendix B reduces every required real and imaginary sign to the sign of an integer polynomial at $2\cos(2\pi/k)$, supplies an explicit algebraic separation bound, and invokes standard polynomial-bit real algebraic arithmetic. In the zero-PCR branch, at most k distinct rotations and their reverse complements are inspected. All remaining operations are linear scans or lexicographic comparisons of words of length at most k. □

8 The q-ary parity theorem

The only even order not covered by the cyclotomic argument is $k = 2$.

Lemma 8.1 (The base case)**.** *For every fixed-point-free complemented alphabet, $B(q, 2)$ has a literal reverse-complement-invariant ordinary MDS.*

Proof. Choose one representative a_i from each complement pair and impose the complement-reversing order

$$a_1 < \cdots < a_{q/2} < \bar{a}_{q/2} < \cdots < \bar{a}_1.$$

Delete the word xy exactly when $y \le x$. Every residual lower arc is strictly increasing, so the residual graph is acyclic. The deletion rule is invariant because $y \le x$ if and only if $\bar{x} \le \bar{y}$. It deletes all q loops and one orientation of every unordered pair, hence

$$q + \frac{q(q-1)}{2} = \frac{q(q+1)}{2} = N_q(2).$$

□

Proof of Theorem 1.1. For even $k \ge 4$, choose any injective antisymmetric nonzero real weight map. Theorem 6.3 gives an invariant ordinary MDS, and Theorem 7.2 gives a deterministic direct choice of one. The case $k = 2$ is Lemma 8.1. Theorem 3.1 excludes ordinary minimum cardinality at every odd $k \ge 3$. □

Corollary 8.2 (DNA)**.** *Let $\mathcal{A} = \{A, C, G, T\}$ with $A \leftrightarrow T$ and $C \leftrightarrow G$. For every $k \ge 2$, $B(4, k)$ has a reverse-complement-invariant ordinary minimum decycling set if and only if k is even. At even k its cardinality is*

$$N_4(k) = \frac{1}{k} \sum_{d|k} \varphi(d) 4^{k/d}.$$

One concrete direct construction uses weights $(-2, -1, 1, 2)$ in the order (A, C, G, T). At every odd $k \ge 3$, an RC-invariant decycling set has cardinality at least $N_4(k) + 4$.

For completeness, define the symmetry-constrained minimum

$$\tau_{\mathrm{RC}}(q, k) = \min\{|M| : M \subseteq \mathcal{A}^k,\ \mathrm{RC}(M) = M,\ B(q, k) \setminus M \text{ is a DAG}\}$$

and the premium $\Delta_q(k) = \tau_{\mathrm{RC}}(q, k) - N_q(k)$. The theorem says

$$\Delta_q(k) = 0 \quad (k \text{ even}), \qquad \Delta_q(k) \ge q \quad (k \ge 3 \text{ odd}).$$

No exact formula for the odd premium is asserted.

9 Exact finite assurance and reproducibility

The proof above is deductive and alphabet-general. The public executable capsule is deliberately restricted to the binary specialization: it reconstructs $k = 2, 4, 6, 8, 10, 12$ using exact cyclotomic coordinates, verifies one-per-PCR selection, residual acyclicity, the edgewise coboundary, every legal firing, and the final literal symmetry. Keeping this scope explicit avoids presenting a finite sample as evidence for an all-order or alphabet-general quantifier.

Two implementations are used. The producer represents exact coordinates in the appropriate quadratic or cyclotomic integer rings and uses Kahn's algorithm for the residual DAG. The independent validator uses frozen exact coordinate tables and a recursive DFS DAG oracle; it does not import the producer. Both reconstruct the complete scientific payload rather than comparing only summary counts.

The panel reports six rows and 612 legal source firings. Seven scientific mutations alter rotation, gradient, half-plane, ray, zero-tie, paired-update, or firing conventions; seven independent envelope mutations alter retained scientific fields. Every mutation must be rejected. Appendix C gives the finite table, and the repository documentation gives the exact replay commands and stable payload identity.

This executable layer neither implements the direct q-ary membership algorithm of Corollary 7.4 nor tests genomic data. Those absences are intentional: the repository accompanies a mathematical article, not a sequence-analysis software release.

10 Boundaries and outlook

The parity theorem settles whether literal reverse-complement symmetry can be achieved at the ordinary necklace bound for every finite alphabet with a fixed-point-free complement. Several nearby questions remain distinct.

1. The exact premium $\Delta_q(k)$ at odd orders is not determined. The proof gives the uniform lower bound q by counting disjoint obstructions, one for each complementary alphabet pair. This is a lower bound, not an equality claim or a packing–covering theorem.
2. The descent does not identify a unique terminal or a shortest path. The direct selector chooses one terminal after fixing weights and a total alphabet order, but different choices may produce different MDSs.
3. Minimum cardinality implies that the residual graph is acyclic, but it does not bound its longest directed path. Quantitative window guarantees require a separate analysis.
4. The fixed-point-free hypothesis is structural. If a complement has fixed letters, lower reverse-complement fixed points and odd-order PCR arguments change. The exceptional order $k = 1$ also behaves differently.
5. Corollary 7.4 is a word-membership theorem, not a claim that explicitly listing an exponentially large MDS takes polynomial total time.
6. No biological conservation, density, runtime, memory, clinical, or application-performance conclusion follows from the graph theorem alone.

The self-dual ideal description may nevertheless be useful independently of the present proof: it converts a global symmetry condition into an exact local membership rule while retaining the ordinary optimum. Understanding which choices of weights and midpoint ties also optimize residual-path statistics is a natural next mathematical problem.

Use of generative AI

Generative-AI assistants contributed substantively to the mathematical development, including candidate proof arguments, symbolic identities, construction and verification code, adversarial checks, and manuscript drafting. The author selected the problem, directed the research process, and is responsible for the article's claims and final text. The two exact implementations are distinct, but the AI-assisted review process does not constitute independent specialist peer review. No end-to-end formal proof-assistant certification is claimed.

A Convention dictionary

Object	Frozen convention
alphabet	finite, nonempty, fixed-point-free involutive complement
de Bruijn orientation	$u \to v$ when the suffix of u equals the prefix of v
rotation	$R(w_0 \cdots w_{k-1}) = w_1 \cdots w_{k-1} w_0$
reverse complement	reverse, complement; directed anti-automorphism
PCR	set of distinct rotations; never an unfolded length-k list
lower-state arc	$w : \mathrm{t}(w) = w_0 \cdots w_{k-2} \to \mathrm{h}(w) = w_1 \cdots w_{k-1}$
weights	existence: injective nonzero reals with $s(\bar{a}) = -s(a)$; exact bit complexity: integers
gradient	$(\nabla f)(w) = f(\mathrm{t}(w)) - f(\mathrm{h}(w))$
half-plane	open upper half-plane; negative-real-axis selector
F-move	all q incoming companions replaced by all q outgoing companions
set equality	literal labelled-word equality unless stated otherwise
MDS	ordinary cardinality minimum plus a literal residual DAG
finite evidence	binary exact panel; never an all-order or q-ary premise

The theorem starts at $k = 2$. At $k = 1$ every vertex is a loop, the unique ordinary minimum is $\mathcal{A}$, and the parity conclusion would be false.

B Stable PCRs and exact algebraic signs

B.1 Stable rotation cycles

Lemma B.1 (Fixed rotations in a stable PCR)**.** *Let C be an RC-stable PCR at even length k, over any fixed-point-free complemented alphabet, and let d be its minimal period. Then d is even and C contains exactly two reverse-complement-fixed words.*

Proof. Choose $w \in C$ and $s \pmod d$ with $\mathrm{RC}(w) = R^s w$. In cyclic position notation,

$$w_r = \overline{w_{s-1-r}}.$$

The positional involution $r \mapsto s - 1 - r$ cannot have a fixed point. If d were odd, 2 would be invertible modulo d and $2r \equiv s - 1 \pmod d$ would have a solution, a contradiction. Thus d is even. For even d, that congruence is solvable exactly when s is odd, so s is even.

A rotation $R^i w$ is reverse-complement fixed precisely when

$$\mathrm{RC}(R^i w) = R^{-i} \mathrm{RC}(w) = R^{s-i} w = R^i w,$$

or $2i \equiv s \pmod d$. Since d and s are even, this congruence has exactly $\gcd(2, d) = 2$ solutions modulo d. □

Lemma B.2. *On words of odd length, reverse complement has no fixed point.*

Proof. The central letter of a fixed odd-length word would equal its complement. □

These facts permit equivariant representatives on stable zero PCRs, but do not by themselves imply global acyclicity; Theorem 4.3 is still essential.

B.2 Polynomial-bit exact sign determination

We justify Corollary 7.4. In this subsection only, specialize the real weight map used by the existence proof to an integer-valued map. Put

$$c = 2\cos(2\pi/k), \qquad d = \varphi(k)/2, \qquad B = \max_{a\in\mathcal{A}} |s(a)|.$$

For $k \geq 4$, c is an algebraic integer of degree d. If Ψ_k is its monic minimal polynomial, it is recovered from the reciprocal cyclotomic identity

$$\Phi_k(z) = z^d \Psi_k(z + z^{-1}). \tag{11}$$

Classical minimal-polynomial formulas provide the same representation [7]. Among the real conjugates of c, the chosen value is the largest one; this property gives a canonical exact isolating interval for the root of Ψ_k at which all subsequent signs are evaluated.

Define integer polynomials

$$C_0 = 2,\ C_1 = x,\ C_{j+1} = xC_j - C_{j-1}, \qquad S_0 = 0,\ S_1 = 1,\ S_{j+1} = xS_j - S_{j-1}.$$

At $x = c$,

$$C_j(c) = 2\cos(2\pi j/k), \qquad S_j(c) = \frac{\sin(2\pi j/k)}{\sin(2\pi/k)}.$$

Since $\sin(2\pi/k) > 0$, every sign or zero used in the seed and midpoint rules reduces to an integer polynomial at c:

$$\begin{aligned} 2\operatorname{Re} P_s(w) &= \sum_j s(w_j) C_{j+1}(c), \\ \operatorname{sign} \operatorname{Im} P_s(w) &= \operatorname{sign} \sum_j s(w_j) S_{j+1}(c), \\ \operatorname{sign} \lambda_s(u) &= \operatorname{sign} \sum_j s(u_j) S_{j+1}(c), \\ 2r_s(u) &= \sum_j s(u_j)\big(C_{j+1}(c) - 2\big). \end{aligned}$$

Reducing a polynomial modulo Ψ_k decides exact zero: the value vanishes at c precisely when the remainder is the zero polynomial.

For a nonzero remainder g, each conjugate c_i of c lies in $[-2, 2]$. The elementary bounds $|C_j(c_i)| \leq 2$ and $|S_j(c_i)| \leq j$ imply the common safe estimate

$$|g(c_i)| \leq M := 2Bk(k+1).$$

Because $g(c)$ is a nonzero algebraic integer, its field norm is a nonzero rational integer. Therefore

$$|g(c)| \geq M^{-(d-1)}. \tag{12}$$

Certified interval evaluation below half this gap determines the sign.

The coefficients of Φ_k, Ψ_k, the recurrences, and their remainders have bit length polynomial in k and $\log(B+1)$. Standard Sturm–Habicht real-root isolation and sign evaluation operate in polynomial bit complexity [2]. Equation (12) requires only $O(k(\log B + \log k))$ refinement bits. Scanning at most k rotations handles a deterministic zero-PCR tie. This proves the claimed polynomial preprocessing, query, and workspace bounds without a real-RAM or floating-point assumption.

C Finite validation boundaries

The executable binary panel has four epistemic roles: it verifies the left-rotation and gradient conventions; exercises negative-ray, strict-crossing, and zero-sector cases; replays every legal firing and terminal invariant; and rejects convention mutations. It neither interpolates in k nor tests the alphabet-general theorem.

k	PCRs	zero PCRs	stable zero	paired zero	$\lvert\operatorname{supp}\chi\rvert$	firings
2	3	0	0	0	0	0
4	6	3	1	1	2	1
6	14	5	1	2	10	5
8	36	6	2	2	52	26
10	108	9	1	4	220	110
12	352	19	5	7	940	470

Table 1: Exact binary convention-assurance panel. PCRs use distinct rotations, and "paired zero" counts PCR pairs.

The total is 612 legal firings. Runtime and envelope metadata are excluded from the stable scientific payload. The paper proof, not this table or its hash, is authority for every all-order statement.

References

[1] Jean-Marc Champarnaud, Georges Hansel, and Dominique Perrin. Unavoidable sets of constant length. *International Journal of Algebra and Computation*, 14(2):241–251, 2004. https://doi.org/10.1142/S0218196704001700.

[2] Ioannis Z. Emiris and Elias P. Tsigaridas. Computations with one and two real algebraic numbers, 2005. https://arxiv.org/abs/cs/0512072.

[3] Guillaume Marçais, Dan DeBlasio, and Carl Kingsford. Sketching methods with small window guarantee using minimum decycling sets. *Journal of Computational Biology*, 31(7):597–615, 2024. https://doi.org/10.1089/cmb.2024.0544.

[4] Guillaume Marçais, C. S. Elder, and Carl Kingsford. k-nonical space: Sketching with reverse complements. *Bioinformatics*, 40(11):btae629, 2024. https://doi.org/10.1093/bioinformatics/btae629.

[5] Johannes Mykkeltveit. A proof of golomb's conjecture for the de bruijn graph. *Journal of Combinatorial Theory, Series B*, 13(1):40–45, 1972. https://doi.org/10.1016/0095-8956(72)90006-8.

[6] David Pellow, Lianrong Pu, Bariş Ekim, Lior Kotlar, Bonnie Berger, Ron Shamir, and Yaron Orenstein. Efficient minimizer orders for large values of k using minimum decycling sets. *Genome Research*, 33(7):1154–1161, 2023. https://doi.org/10.1101/gr.277644.123.

[7] William Watkins and Joel Zeitlin. The minimal polynomial of $\cos(2\pi/n)$. *The American Mathematical Monthly*, 100(5):471–474, 1993. https://doi.org/10.1080/00029890.1993.11990434.