Proving the Herman-Protocol Conjecture*

Maria Bruna[1], Radu Grigore[2], Stefan Kiefer[†3], Joël Ouaknine[‡4], and James Worrell[5]

1 Mathematical Institute, University of Oxford, Oxford, United Kingdom
2 School of Computing, University of Kent, Canterbury, United Kingdom
3 Department of Computer Science, University of Oxford, Oxford, United Kingdom
4 Department of Computer Science, University of Oxford, Oxford, United Kingdom
5 Department of Computer Science, University of Oxford, Oxford, United Kingdom

Abstract

Herman's self-stabilization algorithm, introduced 25 years ago, is a well-studied synchronous randomized protocol for enabling a ring of N processes collectively holding any odd number of tokens to reach a stable state in which a single token remains. Determining the worst-case expected time to stabilization is the central outstanding open problem about this protocol. It is known that there is a constant h such that any initial configuration has expected stabilization time at most hN^2. Ten years ago, McIver and Morgan established a lower bound of $4/27 \approx 0.148$ for h, achieved with three equally-spaced tokens, and conjectured this to be the optimal value of h. A series of papers over the last decade gradually reduced the upper bound on h, with the present record (achieved in 2014) standing at approximately 0.156. In this paper, we prove McIver and Morgan's conjecture and establish that $h = 4/27$ is indeed optimal.

1998 ACM Subject Classification F.1.2 Modes of Computation

Keywords and phrases randomized protocols, self-stabilization, Lyapunov function, expected time

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.104

1 Introduction

The notion of *self-stabilization* was introduced in a seminal paper of Dijkstra [11], and rose to prominence a decade later, following (among others) an invited talk of Lamport during which he pointed out that "self-stabilization [is] a very important concept in fault tolerance" [22]. Both self-stabilization and fault tolerance have since become central themes in distributed computing (see, e.g., [12]), as recently witnessed by the award of the 2015 Edsger W. Dijkstra Prize in Distributed Computing to Michael Ben-Or and Michael Rabin for "starting the field of fault-tolerant randomized distributed algorithms" in the early 1980s.

In this paper, we examine an early self-stabilization algorithm known as Herman's Protocol [19], whose exact mathematical analysis has proven remarkably challenging over the two-and-a-half decades since its inception. This algorithm considers a ring of N processes

* Full version at `http://arxiv.org/abs/1504.01130`.
† Stefan Kiefer is supported by a University Research Fellowship of the Royal Society and by EPSRC grant EP/M003795/1.
‡ Joël Ouaknine is supported by ERC grant AVS-ISS (648701).

43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016).
Editors: Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi; Article No. 104; pp. 104:1–104:12
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

(or nodes), where each process either holds or doesn't hold a token. Starting from any initial configuration of K tokens, where K is required to be odd, Herman's algorithm proceeds as follows: at each time step, every process that holds a token either keeps it or passes it to its clockwise neighbor with probability 1/2. All updates happen synchronously, and if a process finds itself with two tokens (having simultaneously kept one and received one from its counterclockwise neighbor) then both tokens are annihilated. It is straightforward to see that, starting from an odd number of tokens and following this procedure, almost surely only one token eventually remains, at which point the ring is said to have *stabilized.*

Herman's original paper [19] presents the algorithm in a form amenable to implementation. Each process possesses a bit, which the process can read and write. Each process can also read the bit of its counterclockwise neighbor. In this representation, having the same bit as one's counterclockwise neighbor is interpreted as having a token. At each time step, each process compares its bit with the bit of its counterclockwise neighbor; if the bits differ, the process keeps its bit, whereas if the bits are the same, the process flips its bit with probability 1/2 and keeps it with probability 1/2. It is straightforward to verify that the bit-flipping version is an implementation of the token-passing version: in particular, a process flipping its bit corresponds to passing its token to its clockwise neighbor. If the number of processes is odd, by construction this bit representation forces the number of tokens to be odd as well, which justifies the assumption that K, the number of tokens, is always odd. In this paper we make no assumption about the parity of the number of *processes*, as we abstract from the bit implementation, and simply assume that the number of tokens is odd throughout.

Herman's original paper [19] showed that the expected time (number of synchronous steps) to stabilization is $O(N^2 \log N)$. The same paper also mentions an improved upper bound of $O(N^2)$ due to Dolev, Israeli, and Moran, without giving a proof or a further reference. In 2004, Fribourg et al. [16] established an upper bound of $2N^2$, and the following year Nakata [24] gave a tighter upper bound of $0.936N^2$ and exhibited an initial configuration with expected stablization time $\Omega(N^2)$. At the same time and independently, McIver and Morgan showed in [23] that the initial configuration consisting of three equally-spaced tokens has an expected stabilization time of exactly $\frac{4}{27}N^2$, and conjectured that this value is an upper bound on the expected time to stabilization starting from *any* initial configuration with *any* (odd) number of tokens. The conjecture is intriguing since increasing the initial number of tokens might be thought to lengthen the expected time to stabilization, due to the larger number of collisions required to achieve stabilization.

Nevertheless, McIver and Morgan's Herman-Protocol Conjecture is supported by considerable amount of experimental evidence [5], and in the intervening years a series of papers have gradually reduced the upper bound on the constant h such that stabilization from any initial configuration takes expected time at most hN^2: upper bounds of approximately 0.64, 0.521, 0.167, and 0.156 are given respectively in [21, 13, 14, 18], the last one provided last year by Haslegrave, and coming relatively close to McIver and Morgan's lower bound of $4/27 \approx 0.148$.

In this paper, we prove McIver and Morgan's conjecture and establish that $h = 4/27$ is indeed optimal. Writing T_z for the stabilization time starting from an initial configuration z, we seek to prove that $\mathbb{E}T_z \leq \frac{4}{27}N^2$. To this end, one of the key ideas is to work with a Lyapunov function $V(z)$ in lieu of the (more complicated) function $\mathbb{E}T_z$. The domain of the function V is *continuous*: a domain element describes a configuration in terms of the distances between adjacent tokens. Combinatorial arguments exploiting the highly symmetrical structure of $V(z)$ enable us to establish that, for an arbitrary configuration z, we have $\mathbb{E}T_z \leq V(z)$, with equality holding for all three-token configurations. Finally, in what constitutes the most technically challenging part of this paper, we combine induction on the number of tokens with analytical techniques to show that V is bounded by $\frac{4}{27}N^2$. Taken together, we obtain $\mathbb{E}T_z \leq \frac{4}{27}N^2$, entailing the Herman-Protocol Conjecture.

The case of there being an *even* number K of tokens is equally natural from a mathematical point of view, although it does not correspond to a concrete bit-flipping protocol. It was established in [14] that the worst-case configuration in this variant is the equidistant *two*-token configuration, with an expected stabilization time of $\frac{1}{2}N^2$; the analysis underlying that result is considerably simpler than what is required in case the number of tokens is odd, as in the present paper.

Herman's protocol is also related to the notion of *coalescing random walks* [2, 8, 1]. There, one considers multiple independent random walks on $\mathbb{Z}^d$ (or on the vertices of a connected graph). When two walks meet, they coalesce into a new random walk. A protocol for self-stabilizing mutual exclusion based on such random walks was proposed in [20]. The expected coalescence time was studied in [7, 25, 6].

It is interesting to note that Herman's ring is closely related to widely-studied models of random walks and Brownian motion in statistical physics. Observe that by a simple modification of the formalism, one may equivalently view Herman's model as a ring in which tokens randomly move in discrete step in *any* direction, with pairwise collisions leading to annihilation; this precisely corresponds to Fisher's *vicious drunks* model [15] (with periodic boundary conditions). Similar models have been studied in chemical physics [10, 3, 28] and statistical mechanics [17, 26, 27], among others.

The rest of the paper is organized as follows. In Section 2 we review previous results in the literature that are relevant to our proof. In Section 3 we outline the structure of our proof, identifying two key lemmas, Lemma 8 and Lemma 9. Those are proved in [4] and Section 4, respectively.

Another solution of the conjecture, using different techniques, is independently shown in [9].

2 Relevant Previous Results

For the rest of the paper we fix the number N of processes. We assume that the number K of tokens is odd, and both N and K are at least 3.

Processes are numbered from 1 to N, clockwise, according to their position in the ring. A configuration with K tokens is formalized as a function $z : \{1, \ldots, K\} \to \{1, \ldots, N\}$ with $z(1) < \cdots < z(K)$, where the ith token ($i \in \{1, \ldots, K\}$) is held by the processor with the number $z(i)$. We write Z_K for the set of configurations with K tokens, and Z for the set of all possible configurations, that is, $Z = Z_1 \cup Z_3 \cup Z_5 \cup \ldots$

For a fixed initial configuration $z = z_0$ we write $(z_t)_{t\geq 0}$ for the stochastic process of configurations emanating from z. The *stabilization time* T_z is the smallest $t \geq 0$ such that $z_t \in Z_1$, i.e., the time until only one token is left. In this paper we focus on the expectation $\mathbb{E}T_z$. It is shown in [23] that if N is odd and a multiple of 3, then there is a configuration $z \in Z_3$ (with the 3 tokens maximally separated in an equilateral triangle) such that $\mathbb{E}T_z = \frac{4}{27}N^2$.

In this paper we show:

▶ **Theorem 1.** *We have $\mathbb{E}T_z \leq \frac{4}{27}N^2$ for all $z \in Z$.*

Equivalently, the Herman conjecture states that for all odd $K \geq 3$ and all $z \in Z_K$ we have $\mathbb{E}T_z \leq \frac{4}{27}N^2$. Only the case $K = 3$ was previously known [23].

The following proposition has been used in a similar form in various papers on Herman's protocol, for instance in [23, Lemma 5]. It bounds the stabilization time by a Lyapunov function V.

▶ **Proposition 2** (Bound by a Lyapunov function). *Given $z \in Z$, denote by $z' \in Z$ the random successor configuration of z. Let $V : Z \to \mathbb{R}$ be a function with*

$$\mathbb{E}(V(z') \mid z) \;\le\; V(z) - 1 \qquad \text{for all } z \in Z \setminus Z_1, \text{ and} \tag{1}$$

$$0 \;\le\; V(z) \qquad \text{for all } z \in Z_1. \tag{2}$$

Then $\mathbb{E}T_z \le V(z)$ for all $z \in Z$. In particular, $V(z) \ge 0$ for all $z \in Z$.

Although this result is not new, we give a short proof based on a martingale argument. The proof is inspired by [18], and may provide some intuition.

Proof. Let $z \in Z$. Consider the stochastic process $(z_t)_{t \ge 0}$ of configurations emanating from $z = z_0$. Define $W_t := V(z_t) + t$. By (1) the process $(W_t)_{t \ge 0}$ is a supermartingale. The stabilization time $T_z = T_{z_0}$ is a stopping time with finite expectation, and the differences $|W_{t+1} - W_t|$ are bounded as the Markov chain reachable from z has finitely many states. Hence, the optional stopping theorem applies, yielding $\mathbb{E}W_{T_z} \le \mathbb{E}W_0 = V(z)$. By definition of W_t we have $\mathbb{E}W_{T_z} = \mathbb{E}V(z_{T_z}) + \mathbb{E}T_z$. Since $z_{T_z} \in Z_1$, we have $\mathbb{E}T_z \le \mathbb{E}W_{T_z}$ by (2). By combining the previous two inequalities, we obtain $\mathbb{E}T_z \le V(z)$. ◀

Following [14, 18] we associate with a configuration $z \in Z_K$ the *gap vector* $\mathbf{g}(z) = (g_0, \ldots, g_{K-1}) \in \mathbb{N}^K$ by setting $g_0 := N + z(1) - z(K)$, and $g_i := z(i+1) - z(i)$ for $i \in \{1, \ldots, K-1\}$. Then $\mathbf{g}(z)/N$ lives in the so-called standard $(K-1)$-simplex $D^{(K)}$, defined by

$$D^{(K)} := \left\{ \mathbf{x} = (x_0, \ldots, x_{K-1}) \in [0,1]^K \;\mid\; x_0 + \cdots + x_{K-1} = 1 \right\}.$$

Towards a suitable Lyapunov function V we define the cubic polynomial $f_3^{(K)} : D^{(K)} \to [0, \infty)$ by

$$f_3^{(K)}(\mathbf{x}) := \sum_{\substack{0 \le i_0 < i_1 < i_2 < K \\ i_2 - i_1,\, i_1 - i_0 \text{ odd}}} x_{i_0} x_{i_1} x_{i_2}.$$

For instance, we have $f_3^{(5)}(\mathbf{x}) = x_0x_1x_2 + x_0x_1x_4 + x_0x_3x_4 + x_1x_2x_3 + x_2x_3x_4$.

The following lemma was implicitly proved in previous works:

▶ **Lemma 3** (Lyapunov function V_3 [14, Page 240, Proof of Theorem 1] and [18, Theorem 4]). *Let $V_3 : Z \to [0, \infty)$ be defined by $V_3(z) := 4N^2 f_3^{(K)}(\mathbf{g}(z)/N)$ for $z \in Z_K$. Denote by $z' \in Z_1 \cup Z_3 \cup \ldots \cup Z_K$ the random successor configuration of $z \in Z_K$. Then $\mathbb{E}(V_3(z') \mid z) = V_3(z) - \frac{K-1}{2}$ for all $z \in Z_K$. Hence, by Proposition 2, $\mathbb{E}T_z \le 4N^2 f_3^{(K)}(\mathbf{g}(z)/N)$.*

For $K = 3$ Lemma 3 gives $\mathbb{E}T_z \le 4N^2 f_3^{(K)}(\mathbf{g}(z)/N) = \frac{4}{N} g_0g_1g_2$. In fact, for $K = 3$ it was shown before in [23] that $\mathbb{E}T_z$ is identically equal to $\frac{4}{N} g_0g_1g_2$, providing an exact formula for the expected stabilization time of configurations with three tokens. Lemma 3 suggests analyzing f_3:

▶ **Lemma 4** (Maximum of f_3 [14, Proof of Theorem 2], [18, Theorem 3]). *For all $K \ge 3$ odd we have*

$$\max_{\mathbf{x} \in D} f_3^{(K)}(\mathbf{x}) \;=\; f_3^{(K)}\left(\frac{1}{K}, \ldots, \frac{1}{K}\right) \;=\; \frac{1}{24}\left(1 - \frac{1}{K^2}\right).$$

By combining Lemmas 3 and 4 one obtains $\mathbb{E}T_z \le \frac{N^2}{6}(1 - \frac{1}{K^2})$, which is the bound obtained in [14]. A slightly better bound is given in [18].

3 Proof of the Herman Conjecture

The function V_3 from Lemma 3 leaves room for improvement since $\mathbb{E}(V_3(z') \mid z) = V_3(z) - \frac{K-1}{2}$, which is strictly less than $V_3(z) - 1$ for $K > 3$. The idea for obtaining an optimal bound is to decrease the gap between $\frac{K-1}{2}$ and 1, by decreasing the Lyapunov function V. One could think that the scaled function $\frac{2}{K-1}V_3$ is also a Lyapunov function satisfying (1), but this is not true; in particular, note that the number of tokens K might be different for a configuration z and its successor z'. Since scaling does not work, we decrease the Lyapunov function by subtracting a quintic polynomial, as follows. Define a quintic polynomial $f_5^{(K)} : D^{(K)} \to [0, \infty)$, similar to $f_3^{(K)}$:

$$f_5^{(K)}(\mathbf{x}) = \sum_{\substack{0 \le i_0 < i_1 < \cdots < i_4 < K \\ i_4 - i_3, \ldots, i_1 - i_0 \text{ odd}}} x_{i_0} x_{i_1} x_{i_2} x_{i_3} x_{i_4}$$

For instance, $f_5^{(3)}(\mathbf{x}) = 0$, $f_5^{(5)}(\mathbf{x}) = x_0x_1x_2x_3x_4$, and $f_5^{(7)}(\mathbf{x}) = x_0x_1x_2x_3x_4 + x_0x_1x_2x_3x_6 + x_0x_1x_2x_5x_6 + x_0x_1x_4x_5x_6 + x_0x_3x_4x_5x_6 + x_1x_2x_3x_4x_5 + x_2x_3x_4x_5x_6$. We also define a polynomial $f^{(K)} : D^{(K)} \to [0, \infty)$:

$$f^{(K)}(\mathbf{x}) := f_3^{(K)}(\mathbf{x}) - \alpha f_5^{(K)}(\mathbf{x}) \qquad \text{with } \alpha := 24 \tag{3}$$

For example, $f^{(5)}(\mathbf{x}) = x_0x_1x_2 + x_0x_1x_4 + x_0x_3x_4 + x_1x_2x_3 + x_2x_3x_4 - \alpha x_0x_1x_2x_3x_4$. Throughout the paper we use α in the expression of $f^{(K)}$ for notational convenience. From now onwards we may drop the superscript K from the domain $D^{(K)}$ of the functions $f_3^{(K)}$, $f_5^{(K)}$ and $f^{(K)}$ to avoid notational clutter when K is understood.

The following properties of f are fundamental:

▶ **Lemma 5** (Symmetry and continuity properties)**.** *The function f has the following properties.*

(a) *It is symmetric with respect to rotation:*

$$f(x_0, \ldots, x_{K-1}) = f(x_1, \ldots, x_{K-1}, x_0)$$

(b) *It is continuous: For $K \ge 5$ we have*

$$f^{(K)}(x_0, 0, x_2, x_3, \ldots, x_{K-1}) = f^{(K-2)}(x_0 + x_2, x_3, \ldots, x_{K-1}).$$

Analogous properties were shown for f_3 in [14]. Their proof carries over to f_5 and hence to f. The following lemma uses f to define a tighter Lyapunov function.

▶ **Lemma 6** (Lyapunov function V)**.** *Define $V : Z \to [0, \infty)$ by $V(z) := 4N^2 f(\mathbf{g}(z)/N)$. Let $z \in Z$ and denote by z' the random successor configuration of z. Then $\mathbb{E}(V(z') \mid z) \le V(z) - 1$. Hence, by Proposition 2, $\mathbb{E}T_z \le 4N^2 f(\mathbf{g}(z)/N)$.*

We remark that a similar Lyapunov function has been investigated in [14, Equation (15)], but did not lead to a proof of the Herman conjecture. It seems that $V(z)$ needs to be chosen with great care, since even slight variations do not work.

Lemma 6 suggests analyzing f:

▶ **Lemma 7** (Maximum of f)**.** *For all $K \ge 3$ odd we have*

$$\max_{\mathbf{x} \in D} f^{(K)}(\mathbf{x}) = \frac{1}{27}.$$

With this in hand our main result follows:

Proof of Theorem 1. Immediate by combining Lemmas 6 and 7. ◀

It remains to prove Lemmas 6 and 7.

3.1 Proof of Lemma 6

Towards Lemma 6 we show:

▶ **Lemma 8** (Lyapunov function V_5). *Define* $V_5 : Z \to [0, \infty)$ *by* $V_5(z) := 4N^2 f_5(\mathbf{g}(z)/N)$. *Let* $K \geq 5$ *and* $z \in Z$ *and denote by* z' *the random successor configuration of* z. *Then*

$$\mathbb{E}(V_5(z') \mid z) = V_5(z) + \frac{1}{32}\frac{(K-1)(K-3)}{N^2} - \frac{1}{2}(K-3) f_3\left(\frac{\mathbf{g}(z)}{N}\right).$$

The proof in [4] requires an analysis of correlations among the changes in gaps between tokens in each step of the protocol. Using Lemma 8 one can readily prove Lemma 6:

Proof of Lemma 6. For $K = 3$ the statement follows from Lemma 3. For $K \geq 5$ we have:

$$\begin{aligned}
\mathbb{E}(V(z') \mid z) &= \mathbb{E}((V_3(z') - 24V_5(z')) \mid z) && \text{by the definitions} \\
&= \mathbb{E}(V_3(z') \mid z) - 24\mathbb{E}(V_5(z') \mid z) && \text{linearity of expectation} \\
&= V_3(z) - \frac{K-1}{2} - 24V_5(z) - \frac{3}{4}\frac{(K-1)(K-3)}{N^2} \\
&\qquad + 12(K-3) f_3\left(\frac{\mathbf{g}(z)}{N}\right) && \text{Lemmas 3 and 8} \\
&\leq V(z) - \frac{K-1}{2} + 12(K-3) f_3\left(\frac{\mathbf{g}(z)}{N}\right) && \text{since } K \geq 3 \\
&\leq V(z) - \frac{K-1}{2} + \frac{K-3}{2} && \text{Lemma 4} \\
&= V(z) - 1
\end{aligned}$$

◀

3.2 Proof of Lemma 7

Towards Lemma 7 we show:

▶ **Lemma 9** (Local maxima of f). *Let* $K \geq 5$ *and odd. There is no* $\mathbf{v} \in D^{(K)}$ *in the interior of* $D^{(K)}$ *such that* $\mathbf{v}$ *is a local maximum and* $f^{(K)}(\mathbf{v}) > \frac{1}{27}$.

The proof in Section 4 involves a combinatorial analysis of inequalities arising from conditions on the derivatives of $f^{(K)}$. Using Lemma 9 one can readily prove Lemma 7:

Proof of Lemma 7. We proceed by induction on K. For the induction base we have $K = 3$. It is straightforward to check that the maximum of $f^{(3)}(\mathbf{x}) = f_3^{(3)}(\mathbf{x}) = x_0 x_1 x_2$ is $f^{(3)}(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) = \frac{1}{27}$.

For the induction step we have $K \geq 5$. Let $\mathbf{v} \in D^{(K)}$ with $f^{(K)}(\mathbf{v}) = \max_{\mathbf{x} \in D^{(K)}} f^{(K)}(\mathbf{x})$. If $\mathbf{v}$ is in the interior of $D^{(K)}$, then by Lemma 9 we have $f^{(K)}(\mathbf{v}) \leq \frac{1}{27}$. If $\mathbf{v}$ is at the boundary of $D^{(K)}$, then $v_i = 0$ for some i. By Lemma 5(a) we can assume that $v_1 = 0$. Using Lemma 5(b) the statement follows from the induction hypothesis. ◀

4 Proof of Lemma 9

In this section we prove Lemma 9. In Section 4.1 we state several properties that an interior local maximum of $f^{(K)}$ would have to satisfy. In Section 4.2 we prove Lemma 9 for $K = 5$ for a first taste of the general argument. In Section 4.3 we prove Lemma 9 for $K = 7$ to illustrate some fine points that occur only for larger values of K. In Section 4.4 we state some combinatorial facts needed for the general case. Finally, in Section 4.5 we prove Lemma 9.

4.1 Properties of an Interior Local Maximum

The following lemma is obtained by considering first and second derivatives of f evaluated at an interior local maximum.

▶ **Lemma 10.** *Let* $\mathbf{v}$ *be a local maximum of* $f^{(K)}$ *in the interior of* $D^{(K)}$ *and define* $c \in \mathbb{R}$ *by*

$$c \quad = \sum_{\substack{1<i_2<K \\ i_2 \text{ even}}} v_{i_2} \quad - \quad \alpha \sum_{\substack{1<i_2<i_3<i_4<K \\ i_2,\, i_4 \text{ even} \\ i_3 \text{ odd}}} v_{i_2} v_{i_3} v_{i_4}\,. \tag{4}$$

This expression holds for the same value of c *if the indices are rotated by an arbitrary* k*: for all* j *the index* i_j *becomes* $(i_j + k) \bmod K$*. Further, we have*

$$\sum_{\substack{3\le i_3<i_4<K \\ i_3 \text{ odd} \\ i_4 \text{ even}}} v_{i_3} v_{i_4} \le \frac{1}{\alpha}\,. \tag{5}$$

Again, this inequality also holds when indices are rotated.

For example, for $K = 7$ we have $c = v_2 + v_4 + v_6 - \alpha(v_2v_3v_4 + v_2v_3v_6 + v_2v_5v_6 + v_4v_5v_6) = v_1 + v_3 + v_5 - \alpha(v_1v_2v_3 + v_1v_2v_5 + v_1v_4v_5 + v_3v_4v_5)$.

Proof of Lemma 10. The idea of the proof is as follows. We pick a particular direction in $D^{(K)}$, namely $\mathbf{d} = (-1, 0, 1, 0, 0, \dots, 0)$, and consider the function $f(\mathbf{v} + \epsilon\mathbf{d})$ as a univariate function of ϵ. Since $\mathbf{v}$ is a local maximum, the first derivative must be zero and the second derivative must be nonpositive. Exploiting the fact that $v_i > 0$ for all i holds in the interior, we obtain (4) and (5), respectively. See [4] for the detailed proof. ◀

Let $S_j^{(K)}(\mathbf{x})$ denote the scalar product of $\mathbf{x}$ with a copy of itself rotated j times:

$$S_j^{(K)}(\mathbf{x}) := \sum_{i=0}^{K-1} x_i x_{i+j}$$

In all formulas it will be the case that the subscript of S is odd. Also, the superscript will be omitted when unimportant or understood from context.

▶ **Corollary 11.** *Let* $\mathbf{v}$ *be a local maximum of* $f^{(K)}$ *in the interior of* $D^{(K)}$*. Then the following inequality holds:*

$$\sum_{\substack{1\le i<K-2 \\ i \text{ odd}}} \frac{K-i-2}{2} S_i(\mathbf{v}) \quad \le \quad \frac{K}{\alpha}\,.$$

For example, for $K = 11$ we have $4S_1(\mathbf{v}) + 3S_3(\mathbf{v}) + 2S_5(\mathbf{v}) + S_7(\mathbf{v}) \le 11/\alpha$.

▶ **Lemma 12** (Bound for f_5)**.** *Suppose that* $\mathbf{v} \in D^{(K)}$ *satisfies* $f^{(K)}(\mathbf{v}) > \frac{1}{27}$*. Then* $\alpha f_5(\mathbf{v}) < \frac{1}{216}$*.*

Proof. By Lemma 4 we have $f_3(\mathbf{v}) \le \frac{1}{24}$ and hence $\alpha f_5(\mathbf{v}) = f_3(\mathbf{v}) - f(\mathbf{v}) < \frac{1}{24} - \frac{1}{27} = \frac{1}{216}$. ◀

4.2 Proof of Lemma 9 for $K = 5$

Let $K = 5$. Then

$$\begin{aligned} f(\mathbf{x}) &= f_3(\mathbf{x}) - \alpha f_5(\mathbf{x}) \\ &= x_0x_1x_2 + x_0x_1x_4 + x_0x_3x_4 + x_1x_2x_3 + x_2x_3x_4 - \alpha x_0x_1x_2x_3x_4 \end{aligned}$$

Towards a contradiction, suppose that there is a local maximum $\mathbf{v}$ with $f(\mathbf{v}) > \frac{1}{27}$ in the interior of D. By (4), the value

$$c = v_2 + v_4 - \alpha v_2v_3v_4 \tag{6}$$

is invariant under rotations. Indeed, $v_{2+k} + v_{4+k} - \alpha v_{2+k}v_{3+k}v_{4+k} \equiv c$ for all k, but we shall avoid explicitly mentioning rotations, for notational simplicity. Summing (6) over all K rotations we obtain:

$$5c = 2 - \alpha f_3(\mathbf{v}) \tag{7}$$

By (6) we have $v_0v_1c = v_0v_1v_2 + v_0v_1v_4 - \alpha f_5(\mathbf{v})$ and, summing over all K rotations,

$$cS_1(\mathbf{v}) = 2f(\mathbf{v}) - 3\alpha f_5(\mathbf{v}) \tag{8}$$

Moreover,

$$cS_1(\mathbf{v}) \overset{\text{Cor. 11}}{\leq} \frac{5c}{\alpha} \overset{(7)}{=} \frac{2}{\alpha} - f_3(\mathbf{v}) = \frac{2}{\alpha} - f(\mathbf{v}) - \alpha f_5(\mathbf{v}).$$

Combining this with (8) gives:

$$\frac{2}{\alpha} \geq 3f(\mathbf{v}) - 2\alpha f_5(\mathbf{v}) \overset{\text{Lemma 12}}{\geq} \frac{3}{27} - 2 \cdot \frac{1}{216}$$

This implies $\alpha \leq 216/11 \approx 19.6$, which is a contradiction as required (since $\alpha = 24$). ◀

4.3 Proof of Lemma 9 for $K = 7$

Let $K = 7$. Towards a contradiction, we suppose again that there is a local maximum $\mathbf{v}$ with $f(\mathbf{v}) > \frac{1}{27}$ in the interior of D. By (4), all K rotations of the following hold with the same $c \in \mathbb{R}$:

$$c = v_2 + v_4 + v_6 - \alpha(v_2v_3v_4 + v_2v_3v_6 + v_2v_5v_6 + v_4v_5v_6) \tag{9}$$

Summing (9) over K rotations we obtain:

$$7c = 3 - 2\alpha f_3(\mathbf{v}) \tag{10}$$

By (9) we have

$$v_0v_1c = v_0v_1v_2+v_0v_1v_4+v_0v_1v_6-\alpha(v_0v_1v_2v_3v_4+v_0v_1v_2v_3v_6+v_0v_1v_2v_5v_6+v_0v_1v_4v_5v_6) \tag{11}$$

and

$$\begin{aligned} v_0v_3c &= v_0v_3v_4 + v_0v_3v_6 - \alpha v_0v_3v_4v_5v_6 + v_0v_2v_3(1 - \alpha(v_3v_4 + v_3v_6 + v_5v_6)) \\ &\geq v_0v_3v_4 + v_0v_3v_6 - \alpha v_0v_3v_4v_5v_6 \end{aligned} \tag{12}$$

where the last inequality is by (5). Summing (11) and (12) over K rotations we obtain:

$$c\big(2S_1(\mathbf{v}) + S_3(\mathbf{v})\big) \;\geq\; 4f_3(\mathbf{v}) - 9\alpha f_5(\mathbf{v}) = 4f(\mathbf{v}) - 5\alpha f_5(\mathbf{v}) \tag{13}$$

Further we have:

$$c\big(2S_1(\mathbf{v}) + S_3(\mathbf{v})\big) \overset{\text{Cor. 11}}{\leq} \frac{7c}{\alpha} \overset{(10)}{=} \frac{3}{\alpha} - 2f_3(\mathbf{v}) \;=\; \frac{3}{\alpha} - 2f(\mathbf{v}) - 2\alpha f_5(\mathbf{v})$$

Combining this with (13) gives:

$$\frac{3}{\alpha} \;\geq\; 6f(\mathbf{v}) - 3\alpha f_5(\mathbf{v}) \overset{\text{Lemma 12}}{\geq} \frac{6}{27} - 3\cdot\frac{1}{216}$$

This leads to $\alpha \leq 14.4$, which is a contradiction as desired. ◀

4.4 Combinatorial Lemmas

In order to generalize the proofs from Sections 4.2 and 4.3 to any odd K, we state some combinatorial lemmas in this subsection. They are proved in [4].

In order to generalize (7) and (10) we show the following lemma:

▶ **Lemma 13.** *We have:*

$$\sum_{k=0}^{K-1} \sum_{\substack{1<i'_0<i'_1<i'_2<K \\ i'_0,\, i'_2 \text{ even} \\ i'_1 \text{ odd}}} x_{i'_0+k}x_{i'_1+k}x_{i'_2+k} = \frac{K-3}{2} \sum_{\substack{0\leq i_0<i_1<i_2<K \\ i_2-i_1,\, i_1-i_0 \text{ odd}}} x_{i_0}x_{i_1}x_{i_2} = \frac{K-3}{2} f_3^{(K)}(\mathbf{x})\,.$$

For example, if $K = 5$, then we obtain that summing the 5 rotations of $x_2x_3x_4$ gives $f_3^{(5)}(\mathbf{x})$. As another example, if $K = 7$, then we obtain that summing the 7 rotations of $x_2x_3x_4 + x_2x_3x_6 + x_2x_5x_6 + x_4x_5x_6$ gives $2f_3^{(7)}(\mathbf{x})$. These two instances of Lemma 13 help establish (7) and (10).

In order to generalize the inequality in (12) we need the following lemma:

▶ **Lemma 14.** *Let $\mathbf{v}$ be a local maximum of $f^{(K)}$ in the interior of $D^{(K)}$. If i_1 is odd and $0 < i_1 < K$, then the following inequality holds:*

$$v_0v_{i_1}\Bigg(\sum_{\substack{1<i_2<K \\ i_2 \text{ even}}} v_{i_2} - \sum_{\substack{1<i_2<i_3<i_4<K \\ i_2,i_4 \text{ even} \\ i_3 \text{ odd}}} \alpha v_{i_2}v_{i_3}v_{i_4}\Bigg) \geq v_0v_{i_1}\Bigg(\sum_{\substack{i_1<i_2<K \\ i_2 \text{ even}}} v_{i_2} - \sum_{\substack{i_1<i_2<i_3<i_4<K \\ i_2,i_4 \text{ even} \\ i_3 \text{ odd}}} \alpha v_{i_2}v_{i_3}v_{i_4}\Bigg)\,.$$

The inequality says that if we drop those terms that do not occur in $f_3^{(K)}$ or $f_5^{(K)}$, then we obtain a lower bound. The proof groups those terms that are not in either of $f_3^{(K)}$ or $f_5^{(K)}$, and then invokes (5) to show that their sum is nonnegative.

In order to generalize (8) and (13) we need Corollary 16 below, which is a consequence of the following lemma:

▶ **Lemma 15.** *Let l be an odd, positive integer. Then:*

$$\sum_{k=0}^{K-1} \sum_{\substack{1\leq i'_1<K-2 \\ i'_1 \text{ odd}}} \frac{K-i'_1-2}{2} \sum_{\substack{i'_1<i'_2<\cdots<i'_{l-1}<K \\ \forall j,\, i'_j\equiv j \pmod 2}} x_kx_{i'_1+k} \prod_{1<j<l} x_{i'_j+k} =$$

$$= \left(\frac{l-1}{2}K - l\right) \sum_{\substack{0\leq i_0<\cdots<i_{l-1}<K \\ i_j-i_{j-1} \text{ odd for } 0<j<l}} \prod_{j=0}^{l-1} x_{i_j}$$

For example, if $K = 5$ and $l = 3$, then we have that summing 5 rotations of $x_0x_1x_2 + x_0x_1x_4$ gives $2f_3^{(5)}(\mathbf{x})$. As another example, if $K = 9$ and $l = 3$, then summing 9 rotations of $3x_0x_1(x_2 + x_4 + x_6 + x_8) + 2x_0x_3(x_4 + x_6 + x_8) + x_0x_5(x_6 + x_8)$ gives $6f_3^{(9)}(\mathbf{x})$.

▶ **Corollary 16.** *We have:*

$$\sum_{k=0}^{K-1} \sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K - i_1 - 2}{2} \sum_{\substack{i_1 < i_2 < K \\ i_2 \text{ even}}} x_{0+k}x_{i_1+k}x_{i_2+k} = (K-3)f_3^{(K)}(\mathbf{x})$$

and also

$$\sum_{k=0}^{K-1} \sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K - i_1 - 2}{2} \sum_{\substack{i_1 < i_2 < i_3 < i_4 < K \\ i_2, i_4 \text{ even} \\ i_3 \text{ odd}}} x_{0+k}x_{i_1+k}x_{i_2+k}x_{i_3+k}x_{i_4+k} = (2K-5)f_5^{(K)}(\mathbf{x})\,.$$

Proof. Instantiate Lemma 15 with $l = 3$ and, respectively, $l = 5$. ◀

4.5 Proof of Lemma 9

Towards a contradiction, suppose that there is a local maximum $\mathbf{v}$ with $f(\mathbf{v}) > \frac{1}{27}$ in the interior of D, i.e., $v_i > 0$ for all $i \in \{0, \dots, K-1\}$. Summing up the K rotations of (4) and using Lemma 13, we obtain:

$$Kc \;=\; \frac{K-1}{2} - \frac{K-3}{2}\alpha f_3(\mathbf{v}) \tag{14}$$

Multiplying (4) on both sides by $\sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K-i_1-2}{2} v_0 v_{i_1}$ we obtain:

$$\begin{aligned} c \sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K - i_1 - 2}{2} v_0 v_{i_1} &= \sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K - i_1 - 2}{2} v_0 v_{i_1} \left(\sum_{\substack{1 < i_2 < K \\ i_2 \text{ even}}} v_{i_2} - \sum_{\substack{1 < i_2 < i_3 < i_4 < K \\ i_2,\, i_4 \text{ even} \\ i_3 \text{ odd}}} \alpha v_{i_2} v_{i_3} v_{i_4} \right) \\ &\ge \sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K - i_1 - 2}{2} v_0 v_{i_1} \left(\sum_{\substack{i_1 < i_2 < K \\ i_2 \text{ even}}} v_{i_2} - \sum_{\substack{i_1 < i_2 < i_3 < i_4 < K \\ i_2,\, i_4 \text{ even} \\ i_3 \text{ odd}}} \alpha v_{i_2} v_{i_3} v_{i_4} \right) \end{aligned}$$

using Lemma 14. Summing K rotations of this inequality yields:

$$\begin{aligned} c \sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K - i_1 - 2}{2} S_{i_1}(\mathbf{v}) \;&\ge\; (K-3)f_3(\mathbf{v}) - (2K-5)\alpha f_5(\mathbf{v}) \\ &= (K-3)f(\mathbf{v}) - (K-2)\alpha f_5(\mathbf{v}) \end{aligned} \tag{15}$$

using Corollary 16. Further we have:

$$c \sum_{\substack{1 \le i_1 < K-2 \\ i_1 \text{ odd}}} \frac{K - i_1 - 2}{2} S_{i_1}(\mathbf{v}) \overset{\text{Cor. 11}}{\le} \frac{Kc}{\alpha} \overset{(14)}{=} \frac{K-1}{2\alpha} - \frac{K-3}{2} f_3(\mathbf{v})\,.$$

Combining this with (15) gives:

$$\frac{K-1}{2\alpha} \;\ge\; \frac{3K-9}{2} f(\mathbf{v}) - \frac{K-1}{2}\alpha f_5(\mathbf{v}) \overset{\text{Lemma 12}}{\ge} \frac{K-3}{2} \cdot \frac{1}{9} - \frac{K-1}{2} \cdot \frac{1}{216}\,.$$

This implies

$$\alpha \leq \frac{216(K-1)}{23K-71} < 19.7\,.$$

Since $\alpha = 24$, this leads to a contradiction as desired. ◀

5 Conclusions

In this paper we have proved the Herman-Protocol Conjecture formulated by McIver and Morgan in [23] a decade ago, which says that the worst-case initial configuration consists of three maximally-separated tokens, for N multiple of 3. This follows from our result that the worst-case self-stabilization time is at most $\frac{4}{27}N^2$, for any number of processes N and any odd number of tokens K.

The proof uses a Lyapunov function approach. To do so, we first find a suitable Lyapunov function and then show that its maximum is $\frac{4}{27}N^2$. Then we show that this function gives an upper bound for the self-stabilization time for *each* possible configuration in Herman's algorithm.

References

1 D. Aldous and J. A. Fill. Reversible Markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

2 R. Arratia. Limiting point processes for rescalings of coalescing and annihilating random walks on Z^d. *The Annals of Probability*, 9(6):909–936, 1981.

3 D. Balding. Diffusion-reaction in one dimension. *J. Appl. Prob.*, 25:733–743, 1988.

4 M. Bruna, R. Grigore, S. Kiefer, J. Ouaknine, and J. Worrell. Proving the Herman-Protocol Conjecture. Technical report, arxiv.org, 2015. Available at `http://arxiv.org/abs/1504.01130`.

5 PRISM case studies. Randomised self-stabilising algorithms. `http://www.prismmodelchecker.org/casestudies/self-stabilisation.php`.

6 C. Cooper, R. Elsässer, H. Ono, and T. Radzik. Coalescing random walks and voting on graphs. In *Proc. PODC*, pages 47–56. ACM, 2012.

7 D. Coppersmith, P. Tetali, and P. Winkler. Collisions among random walks on a graph. *SIAM Journal on Discrete Mathematics*, 6(3):363–374, 1993.

8 J.T. Cox. Coalescing random walks and voter model consensus times on the torus in Z^d. *The Annals of Probability*, 17(4):1333–1366, 1989.

9 E. Csóka and S. Mészáros. Generalized solution for the Herman protocol conjecture. Technical report, arxiv.org, 2015. Available at `http://arxiv.org/abs/1504.06963`.

10 P.-G. de Gennes. Soluble model for fibrous structures with steric constraints. *J. Chem. Phys.*, 48(5):2257–2259, 1968.

11 E. W. Dijkstra. Self-stabilizing systems in spite of distributed control. *Comm. ACM*, 17(11):643–644, 1974.

12 S. Dolev. *Self-Stabilization*. MIT Press, 2000.

13 Y. Feng and L. Zhang. A Tighter Bound for the Self-Stabilization Time in Herman's Algorithm. *Inf. Process. Lett.*, 113(13):486–488, 2013.

14 Y. Feng and L. Zhang. A nearly optimal upper bound for the self-stabilization time in Herman's algorithm. *Dist. Comp.*, pages 1–12, 2015.

15 M. E. Fisher. Walks, walls, wetting, and melting. *J. Stat. Phys.*, 34(5-6):667–729, 1984.

16 L. Fribourg, S. Messika, and C. Picaronny. Coupling and self-stabilization. *Dist. Comp.*, 18:221–232, 2005.

17 S. Y. Grigoriev and V. B. Priezzhev. Random walk of annhilating particles on the ring. *Theor. Math. Phys.*, 146(3):411–420, 2006.

18 J. Haslegrave. Bounds on Herman's algorithm. *Theoretical Computer Science*, 550:100–06, 2014.

19 T. Herman. Probabilistic self-stabilization. *Inf. Process. Lett.*, 35(2):63–67, 1990.

20 A. Israeli and M. Jalfon. Token management schemes and random walks yield self-stabilizing mutual exclusion. In *Proc. PODC*, pages 119–131. ACM, 1990.

21 S. Kiefer, A. Murawski, J. Ouaknine, J. Worrell, and L. Zhang. On stabilization in Herman's algorithm. In *Proc. ICALP*, volume 6756 of *LNCS*. Springer, 2011.

22 L. Lamport. Solved problems, unsolved problems and non-problems in concurrency. In *Proc. PODC*, pages 1–11. ACM, 1984.

23 A. McIver and C. Morgan. An elementary proof that Herman's ring is $\Theta(N^2)$. *Inf. Process. Lett.*, 94(2):79–84, 2005.

24 T. Nakata. On the expected time for Herman's probabilistic self-stabilizing algorithm. *Theor. Comput. Sci.*, 349(3):475–483, 2005.

25 R.I. Oliveira. On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.*, 364(4):2109–2128, 2012.

26 J. Rambeau and G. Schehr. Distribution of the time at which N vicious walkers reach their maximal height. *Phys. Rev. E*, 83, 2011.

27 G. Schehr, S. N. Majumdar, A. Comtet, and P. J. Forrester. Reunion probability of N vicious walkers: Typical and large fluctuations for large N. *J. Stat. Phys.*, 150:491–530, 2013.

28 M. Warner. Aggregation in dense solutions of rods. *J. Chem. Soc. Faraday. Trans.*, 87(6):861–867, 1991.