# Approximability of Dodgson's Rule

JOHN C. M$^{\text{C}}$CABE-DANSTED
Department of Mathematics, Auckland University
gmatht@gmail.com

DR GEOFFREY PRITCHARD
Department of Statistics, Auckland University
geoff@stat.auckland.ac.nz

DR ARKADII SLINKO
Department of Mathematics, Auckland University
a.slinko@auckland.ac.nz

25 August 2006

**ABSTRACT.** It is known that Dodgson's rule is computationally very demanding. Tideman (1987) suggested an approximation to it but did not investigate how often his approximation selects the Dodgson winner. We show that under the Impartial Culture assumption the probability that the Tideman winner is the Dodgson winner tend to 1. However we show that the convergence of this probability to 1 is slow. We suggest another approximation — we call it Dodgson Quick — for which this convergence is exponentially fast. Also we show that Simpson and Dodgson rules are asymptotically different. We formulate, and heavily use in construction of examples, the generalization of M$^{\text{c}}$Garvey's theorem (1953) for weighted majority relations.

**JEL CLASSIFICATION:** D7

**KEYWORDS:** Dodgson's rule, Simpson's rule, M$^{\text{c}}$Garvey's Theorem, weighted majority relation, Impartial Culture assumption.

The official version of this paper is published in Social Choice and Welfare.
See http://dx.doi.org/10.1007/s00355-007-0282-8.

## 1 Introduction

Condorcet proposed that a winner of an election is not legitimate unless a majority of the population prefer that alternative to all other alternatives. However such a winner does not always exist. A number of voting rules have been proposed which select the Condorcet winner if it exists, and otherwise selects an alternative that is in some sense closest to being a Condorcet Winner. A prime example of such as rule was the one proposed by Dodgson (1876).

Unfortunately, Bartholdi et al. (1989) proved that finding the Dodgson winner is an NP-hard problem. Hemaspaandra et al. (1997) refined this result by proving that it is $\Theta_2^p$-complete and hence is not NP-complete unless the polynomial hierarchy collapses. As Dodgson's rule is hard to compute, it is important to have simple and fast approximations to it. We investigate the asymptotic behaviour of simple approximations to the Dodgson rule as the number of agents gets large. Tideman (1987) suggested an approximation but did not investigate its convergence to Dodgson. We prove that under the assumption that all votes are independent and each type of vote is equally likely (the Impartial Culture (IC) assumption), the probability that the Tideman (1987) approximation picks the Dodgson winner asymptotically converge to 1, but not exponentially fast.

We propose a new social choice rule, which we call Dodgson Quick. The Dodgson Quick approximation does exhibit exponential convergence to Dodgson, and we can quickly verify that it has chosen the Dodgson winner. This, together with it simplicity and other nice properties, makes our new approximation useful in computing the Dodgson winner. Despite its simplicity, our approximation picked the correct winner in all of 1,000,000 elections with 85 agents and 5 alternations, each generated randomly according to the Impartial Culture assumption. Our approximation can also be used to develop an algorithm to determine the Dodgson winner with $\mathcal{O}(\ln n)$ expected running time for a fixed number of alternatives and $n$ agents.

A result independently obtained by Homan and Hemaspaandra (2005) has a lot in common with our result formulated in the previous paragraph, but there are important distinctions as well. They developed a "greedy" algorithm that, given a profile, finds the Dodgson winner with certain probability. Under the Impartial Culture assumption this probability also approaches 1 as we increase the number of agents. However they do not suggest any rule with a simple mathematical definition and, unlike their algorithm, the Dodgson Quick rule requires only the information in the weighted majority relation. This makes the Dodgson Quick rule easier to study and compare to other simple rules such as the Tideman rule.

Our experimental results (M$^c$Cabe-Dansted and Slinko, 2006) showed that Simpson's and Dodgson's rules are very close. However, in the present paper we discover that under the Impartial Culture assumption, the frequency that the Simpson rule picks the Dodgson winner does not converge to one.

M$^c$Garvey (1953) proved that for any tournament, we can find a profile such that the tournament is the majority relation on that profile. However, when we pass from a profile to the majority relation on that profile, we lose all the information about the margins of defeat. By retaining this information we obtained the so-called weighted majority relation on the profile. Weighted majority relation can be represented as an ordinary tournament with advantages attached to its edges, i.e. by a weighted tournament. For the purposes of constructing the examples in this paper we need a generalisation of McGarvey's theorem to weighted tournaments. Fortunately it can be done, and for any weighted tournament, it is possible to find a society with that weighted tournament as its weighted majority relation, if and only if all the weights are even or all the weights are odd.

The first paper that mentions this result was probably Debord's PhD thesis (1987) as quoted by Vidu (1999). However this source is inaccessible to the authors. We view this result as important and give an independent proof of it in this paper.

## 2 Preliminaries

Let $A$ and $\mathcal{N}$ be two finite sets of cardinality $m$ and $n$ respectively. The elements of $A$ will be called alternatives, the elements of $\mathcal{N}$ agents. We assume that the agents have preferences over the set of alternatives represented by (strict) linear orders. By $\mathcal{L}(A)$ we denote the set of all linear orders on $A$. The elements of the Cartesian product

$$\mathcal{L}(A)^n = \mathcal{L}(A) \times \cdots \times \mathcal{L}(A) \quad (n \text{ times}) \tag{1}$$

are called **profiles**. The profiles represent the collection of preferences of an $n$-element society of agents $\mathcal{N}$. A family of mappings $F = \{F_n\}$, $n \in \mathbb{N}$,

$$F_n \colon \mathcal{L}(A)^n \to A, \tag{2}$$

is called a **social choice function** (SCF).

Let $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ be a profile. If a linear order $P_i \in \mathcal{L}(A)$ represents the preferences of the $i^{\text{th}}$ agent, then by $aP_ib$, where $a, b \in A$, we denote that this agent prefers $a$ to $b$. We define $n_{xy}$ to be the number of linear orders in $\mathcal{P}$ that rank $x$ above $y$, i.e. $n_{xy} = \#\{i \mid xP_iy\}$. The approximations we consider depend upon the information contained in the matrix $N_{\mathcal{P}}$, where $(N_{\mathcal{P}})_{ab} = n_{ab}$. A function $W^{\mathcal{P}} \colon A \times A \to \mathbb{Z}$ given by $W^{\mathcal{P}}(a,b) = n_{ab} - n_{ba}$ for all $a, b \in A$, will be called the **weighted majority relation** on $\mathcal{P}$. It is obviously skew symmetric, i.e. $W^{\mathcal{P}}(a,b) = -W^{\mathcal{P}}(b,a)$ for all $a, b \in A$.

Many of the rules to determine the winner use the numbers

$$\operatorname{adv}(a,b) = \max(0, n_{ab} - n_{ba}) = (n_{ab} - n_{ba})^+, \tag{3}$$

which will be called **advantages**. Note that $\operatorname{adv}(a,b) = \max(0, W(a,b)) = W(a,b)^+$ where $W$ is the weighted majority relation on $\mathcal{P}$.

A **Condorcet winner** is an alternative $a$ for which $\operatorname{adv}(b,a) = 0$ for all other alternatives $b$. A Condorcet winner does not always exist. The rules we consider below attempt to pick an alternative that is in some sense closest to being a Condorcet winner. These rules will always pick the Condorcet winner when it exists; such rules are called Condorcet consistant rules.

The social choice rules we consider are based on calculating the vector of **scores** and the alternative with the lowest score wins. Let the lowest score be $s$. It is possible that more than one alternative has a score of $s$. In this case we may have a set of winners with cardinality greater than one. Strictly speaking, to be a social choice function, a rule has to output a single winner. Rules are commonly modified to achieve this by splitting ties. One of the most popular methods of splitting ties is to use for this purpose the preference of the first agent. However we will usually study the set of tied winners rather than the single winner output from a tie-breaking procedure, as this will give us more information about the rules.

The **Dodgson score** (Dodgson 1876, see e.g. Black 1958; Tideman 1987), which we denote as $Sc_{\mathbf{d}}(a)$, of an alternative $a$ is the minimum number of neighbouring alternatives that must be swapped to make $a$ a Condorcet winner. We call the alternative(s) with the lowest Dodgson score the **Dodgson winner**(s).

The **Simpson score** (Simpson 1969, see e.g. Laslier 1997) $Sc_{\mathbf{s}}(a)$ of an alternative $a$ is

$$Sc_{\mathbf{s}}(a) = \max_{b \neq a} \mathrm{adv}(b, a). \tag{4}$$

We call the alternative(s) with the lowest Simpson score the **Simpson winner**(s). That is, the alternative with the smallest maximum defeat is the Simpson winner. This is why the rule is often known as the Maximin or Minimax rule.

The **Tideman score** (Tideman, 1987) $Sc_{\mathbf{t}}(a)$ of an alternative $a$ is

$$Sc_{\mathbf{t}}(a) = \sum_{b \neq a} \mathrm{adv}(b, a). \tag{5}$$

We call the alternative(s) with the lowest Tideman score the **Tideman winner**(s). Tideman (1987) suggested the rule based on this score as an approximation to Dodgson.

The **Dodgson Quick (DQ) score** $Sc_{\mathbf{q}}(a)$ of an alternative $a$, which we introduce in this paper, is

$$Sc_{\mathbf{q}}(a) = \sum_{b \neq a} F(b, a), \tag{6}$$

where

$$F(b, a) = \left\lceil \frac{\mathrm{adv}(b, a)}{2} \right\rceil. \tag{7}$$

We call the alternative(s) with the lowest Dodgson Quick score the **Dodgson Quick winner**(s) or **DQ-winner**.

The **Impartial Culture** assumption (IC) stipulates that all possible profiles $\mathcal{P} \in \mathcal{L}(A)^n$ are equally likely to represent the collection of preferences of an $n$-element society of agents $\mathcal{N}$, i.e. all agents are independent and they choose their linear orders from the uniform distribution on $\mathcal{L}(A)$. This assumption is of course does not accurately reflect the voting behaviour of most voting societies. Worse, we have found that the choice of probability model for the population can affect the similarities between approximations to the Dodgson rule (M$^{c}$Cabe-Dansted and Slinko, 2006). However the IC is the simplest assumption available. As noted by Berg (1985), many voting theorists have chosen to focus their research upon the IC. Thus an in depth study of the approximability of Dodgson's rule under the Impartial Culture is a natural first step.

The IC leads to the following $m!$-dimensional multinomial distribution. Let us enumerate all $m!$ linear orders in some way. Let $\mathcal{P} \in \mathcal{L}(A)^n$ be a random profile. Let then $X$ be a vector where each $X_i$, for $i = 1, 2, \ldots, m!$, represents the number of occurrences of the $i^{\text{th}}$ linear order in the profile $\mathcal{P}$. Then, under the IC, the vector $X$ is $(n, k, \mathbf{p})$-multinomially distributed with $k = m!$ and $\mathbf{p} = \mathbf{1}_k/k = (\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k})$.

## 3 A M^c^Garvey Theorem for Weighted Tournaments

The McGarvey Theorem (1953) is states that every tournament can be represented as a majority relation for a certain society of voters. We will prove a generalization of the McGarvey Theorem to weighted tournaments and weighted majority relations.

Like most other authors, Laslier (1997) defines weighted tournaments as matrices and tournaments as (complete and asymmetric) binary relations. However Laslier notes that there are many different equivalent definitions of tournaments, of which Laslier gives four examples. In this paper we define both tournaments and weighted tournaments as functions, for consistency.

**Definition 3.1** *A* ***weighted tournament*** *on a set $A$ is any function $W\colon A \times A \to \mathbb{Z}$ satisfying $W(a,b) = -W(b,a)$ for all $a, b \in A$.*

We call $W(a,b)$ the **weight** of an ordered pair of distinct elements $(a,b)$. One can view weighted tournaments as complete directed graphs whose edges are assigned integers characterising the intensity and the sign of the relation between the two vertices that this particular edge connects. The only condition is that if an edge from $a$ to $b$ is assigned integer $z$, then the edge from $b$ to $a$ is assigned the integer $-z$.

Weighted majority relation $W^{\mathcal{P}}$ on a profile $\mathcal{P}$ defined earlier in this paper is a prime example of a weighted tournament. We say that a profile $\mathcal{P}$ **generates** a weighted tournament $W$ if $W = W^{\mathcal{P}}$. We note that $\text{adv}(a,b) = W^{\mathcal{P}}(a,b)^+$, where $x^+ = \max(0,x)$. Similarly $W^{\mathcal{P}}(a,b) = \text{adv}(a,b) - \text{adv}(b,a)$. Another important example is the classical tournament on $A$ which can be identified with a weighted tournament $W$ whose all weights $W(a,b)$ are 1 or $-1$.

**Definition 3.2** *For a weighted tournament $W$, for which all weights $W(a,b)$ are non-zero, we define the* ***reduction*** *of $W$ to be the tournament $W_S$, where:*

$$W_S(a,b) \;=\; \begin{cases} 1 & \text{if} \quad W(a,b) > 0, \\ -1 & \text{if} \quad W(a,b) < 0, \\ 0 & \text{if} \quad a = b. \end{cases} \tag{8}$$

For any weighted tournament $W$ for which there exist $a \neq b$ such that $W(a,b) = 0$, this reduction is not defined. We note that if $n$ is odd and $W^{\mathcal{P}}$ is the weighted majority relation on the profile $\mathcal{P}$, then the reduction $W_S^{\mathcal{P}}$ is the classical majority relation on the profile $\mathcal{P}$.

We are now ready to formulate the main result of this section.

**Theorem 3.3** *Let $W$ be a weighted tournament. Then there exists a profile that generates a weighted tournament $W$ if and only if all weights in $W$ have the same parity (either all odd or all even).*

The proof will be split into three lemmata.

**Lemma 3.4** *Let $W^{\mathcal{P}}$ be a weighted majority relation on a profile $\mathcal{P}$ with $n$ agents, then all weights in $W^{\mathcal{P}}$ have the same parity as $n$. That is, for each pair of distinct alternatives $a$ and $b$, the weight $W^{\mathcal{P}}(a,b)$ is even if and only if $n$ is even.*

Proof. We know that for all alternatives $a$ and $b$ we have $W^{\mathcal{P}}(a,b) = n_{ab} - n_{ba}$ and $n = n_{ba} + n_{ab}$. Hence $W^{\mathcal{P}}(a,b) + n = 2n_{ab}$ and so $W^{\mathcal{P}}(a,b)$ and $n$ have the same parity. □

**Lemma 3.5** *For a weighted tournament $W$ with all weights being even, we may construct a profile $\mathcal{P}$ that generates $W$.*

Proof. We may construct such a profile as follows. We start with an empty profile $\mathcal{P}$. For each ordered pair of alternatives $(a,b)$, for which the weight $W(a,b)$ is positive, we let $k = W(a,b)/2$. We take the linear order $\mathbf{v}$, on the set of alternatives $A$, such that $a\mathbf{v}b$ and $b\mathbf{v}x$ for all $x \neq a, b$ and the linear order $\mathbf{w}$ such that $a\mathbf{w}b$ and $x\mathbf{w}a$ for all $x \neq a, b$. We add $k$ instances of $\mathbf{v}$ and $k$ instances of $\mathbf{w}$ to the profile $\mathcal{P}$. After we have done this for every pair $(a,b)$, it is easy to check that the resulting profile generates weighted majority relation $W^{\mathcal{P}}$ such that $W^{\mathcal{P}}(a,b) = W(a,b)$ for all $a, b \in A$.

□

For the last lemma we will need the following definition.

**Definition 3.6** *Let $W_1$ and $W_2$ be two weighted tournaments. We define their sum $W_1 + W_2$ and their difference $W_1 - W_2$ as functions, which for all alternatives $a$ and $b$ satisfy*

$$(W_1 + W_2)(a,b) = W_1(a,b) + W_2(a,b), \qquad (W_1 - W_2)(a,b) = W_1(a,b) - W_2(a,b). \tag{9}$$

It is easy to check that the sum and the difference of two weighted tournaments is again a weighted tournament.

**Lemma 3.7** *For a weighted tournament $W$ with all weights being odd, we may construct a profile which generates this weighted tournament.*

Proof. Let $W_1$ be the weighted majority relation of a profile consisting of a single arbitrarily chosen linear order $\mathbf{v}$. Let $W_2 = W - W_1$. Note that as $W_1$ is generated from a profile with an odd number (i.e. one) of linear orders, all the weights in $W_1$ must be odd. Thus all weights in $W_2$ are the difference between two odd numbers. Hence all weights in $W_2$ are even and we can construct a profile for which $W_2$ is the majority relation, as shown by Lemma 3.5. Since $W = W_1 + W_2$, joining the profile that generates $W_1$, which is the linear order $\mathbf{v}$, and the profile that generates $W_2$ we obtain a profile that generates $W$.

$\square$

Our generalisation to the M$^c$Garvey theorem easily follows from these lemmata.

## 4 Dodgson Quick, A New Approximation

In this section we work under the Impartial Culture assumption.

**Definition 4.1** *Let $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ be a profile. We say that the $i^{th}$ agent ranks $b$* ***directly above*** *$a$ if and only if $aP_ib$ and there does not exist $c$ different from $a, b$ such that $aP_ic$ and $cP_ib$. We define $D(b, a)$ as the number of agents who rank $b$ directly above $a$.*

**Lemma 4.2** *The probability that $D(x, a) > F(x, a)$ for all $x$ converges exponentially fast to 1 as the number of agents $n$ tends to infinity.*

Proof. As $n_{ba}$ and $D(b, a)$ are binomially distributed with means of $n/2$ and $n/m$, respectively, from Chomsky's (Dembo and Zeitouni, 1993) large deviation theorem, we know that for a fixed number of alternatives $m$ there exist $\beta_1 > 0$ and $\beta_2 > 0$ such that

$$P\left(\frac{D(b,a)}{n} < \frac{1}{2m}\right) \leq e^{-\beta_1 n}, \qquad P\left(\frac{n_{ba}}{n} - \frac{1}{2} > \frac{1}{4m}\right) \leq e^{-\beta_2 n}.$$

We can rearrange the second equation to involve $F(b, a)$,

$$P\left(\frac{n_{ba}}{n} - \frac{1}{2} > \frac{1}{4m}\right) = P\left(\frac{2n_{ba}}{n} - 1 > \frac{1}{2m}\right) \tag{10}$$

$$= P\left(\frac{n_{ba} - n_{ab}}{n} > \frac{1}{2m}\right) \tag{11}$$

$$= P\left(\frac{\text{adv}(b,a)}{n} > \frac{1}{2m}\right). \tag{12}$$

Since $\text{adv}(b, a) \geq F(b, a)$,

$$P\left(\frac{n_{ba}}{n} - \frac{1}{2} > \frac{1}{4m}\right) \geq P\left(\frac{F(b,a)}{n} > \frac{1}{2m}\right). \tag{13}$$

From this and the law of probability $P(A \vee B) \leq P(A) + P(B)$ it follows that

$$P\left(\frac{F(b,a)}{n} > \frac{1}{2m}\right) \leq e^{-\beta_2 n}, \qquad P\left(\frac{D(b,a)}{n} < \frac{1}{2m}\right) \leq e^{-\beta_1 n},$$

and so for $\beta = \min(\beta_1, \beta_2)$ we obtain

$$P\left(\frac{F(b,a)}{n} > \frac{1}{2m} \text{ or } \frac{D(b,a)}{n} < \frac{1}{2m}\right) \leq e^{-\beta_1 n} + e^{-\beta_2 n} \leq 2e^{-\beta n}. \tag{14}$$

Hence

$$P\left(\exists_x \frac{F(x,a)}{n} > \frac{1}{2m} \text{ or } \frac{D(x,a)}{n} < \frac{1}{2m}\right) \quad \leq \quad 2me^{-\beta n}. \tag{15}$$

Using $P(\bar{E}) = 1 - P(E)$, we find that

$$P\left(\forall_x \frac{F(x,a)}{n} < \frac{1}{2m} < \frac{D(x,a)}{n}\right) \quad \geq \quad 1 - 2me^{-\beta n}. \tag{16}$$

□

**Lemma 4.3** *The DQ-score $Sc_{\mathbf{q}}(a)$ is a lower bound for the Dodgson Score $Sc_{\mathbf{d}}(a)$ of $a$.*

Proof. Let $\mathcal{P}$ be a profile and $a \in A$. Suppose we are allowed to change linear orders in $\mathcal{P}$, by repeated swapping neighbouring alternatives. Then to make $a$ a Condorcet winner we must reduce $\text{adv}(x, a)$ to 0 for all $x$ and we know that $\text{adv}(x, a) = 0$ if and only if $F(x, a) = 0$. Swapping $a$ over an alternative $b$ ranked directly above $a$ will reduce $n_{ba} - n_{ab}$ by two, but this will not affect $n_{ca} - n_{ac}$ where $a \neq c$. Thus swapping $a$ over $b$ will reduce $F(b, a)$ by one, but will not affect $F(c, a)$ where $b \neq c$. Therefore, making $a$ a Condorcet winner will require at least $\Sigma_b F(b, a)$ swaps. This is the DQ-Score $Sc_{\mathbf{q}}(a)$ of $a$. □

**Lemma 4.4** *If $D(x, a) \geq F(x, a)$ for every alternative $x$, then the DQ-Score $Sc_{\mathbf{q}}(a)$ of $a$ is equal to the Dodgson Score $Sc_{\mathbf{d}}(a)$ and the DQ-Winner is equal to the Dodgson Winner.*

Proof. If $D(b, a) \geq F(b, a)$, we can find at least $F(b, a)$ linear orders in the profile where $b$ is ranked directly above $a$. Thus we can swap $a$ directly over $b$, $F(b, a)$ times, reducing $F(b, a)$ to 0. Hence we can reduce $F(x, a)$ to 0 for all $x$, making $a$ a Condorcet winner, using $\Sigma_x F(x, a)$ swaps of neighbouring alternatives. In this case, $Sc_{\mathbf{q}}(a) = \Sigma_b F(b, a)$ is an upper bound for the Dodgson Score $Sc_{\mathbf{d}}(a)$ of $a$. From Lemma 4.3 above, $Sc_{\mathbf{q}}(a)$ is also a lower bound for $Sc_{\mathbf{d}}(a)$. Hence $Sc_{\mathbf{q}}(a) = Sc_{\mathbf{d}}(a)$. □

**Theorem 4.5** *The probability that the DQ-Score $Sc_{\mathbf{q}}(a)$ of an arbitrary alternative $a$ is equal to the Dodgson Score $Sc_{\mathbf{d}}(a)$, converges to 1 exponentially fast.*

Proof. From Lemma 4.4, if $D(x, a) \geq F(x, a)$ for all alternatives $x$ then $Sc_{\mathbf{q}}(a)$= $Sc_{\mathbf{d}}(a)$. From Lemma 4.2, the probability of this event converges exponentially fast to 1 as $n \to \infty$. □

**Corollary 4.6** *The probability that the DQ-Winner is the Dodgson Winner converges to 1 exponentially fast as we increase the number of agents.*

**Corollary 4.7** *Suppose that the number of alternatives $m$ is fixed. Then there exists an algorithm that computes the Dodgson score of an alternative $a$ taking as input the frequency of each linear order in the profile $\mathcal{P}$ with expected running time logarithmic with respect to the number of agents (i.e. is $\mathcal{O}(\ln n)$).*

Proof. The are at most $m!$ distinct linear orders in the profile. Hence for a fixed number of alternatives the number of distinct linear orders is bounded. Hence we may find the DQ-score and check whether $D(x, a) \geq F(x, a)$ for all alternatives $x$ using a fixed number of additions. The largest possible number of additions needed is proportional to the number of agents $n$. Additions can be performed in time linear with respect to the number of bits and logarithmic with respect to the size of the number. So we have only used an amount of time that is logarithmic with respect to the number of agents.

If $D(x, a) \geq F(x, a)$ for all alternatives $x$, we know that the DQ-score is the Dodgson score and we do not need to go further. From Lemma 4.2 we know that the probability that we need go further declines exponentially fast, and, if this happens, we can still find the Dodgson score in time polynomial with respect to the number of agents (Bartholdi et al., 1989). □

**Corollary 4.8** *There exists an algorithm that computes the Dodgson winner taking as input the frequency of each linear order in the profile $\mathcal{P}$ with expected running time that is logarithmic with respect the number of agents.*

# 5 Tideman's Rule

In this section we focus our attention on the Tideman rule which was defined in Section 2. We continue to assume the IC.

**Lemma 5.1** *Given an even number of agents, the Tideman winner and the DQ-winner will be the same.*

Proof. Since the $n$ is even, we know from Lemma 3.4 that all weights in the majority relation $W$ are even. Since the $\text{adv}(a, b) \equiv W(a, b)^+$ it is clear that all advantages will also be even. Since $\text{adv}(a, b)$ will always be even, $\lceil \text{adv}(a, b)/2 \rceil$ will be exactly half $\text{adv}(a, b)$ and so the DQ-score will be exactly half the Tideman score. Hence the DQ-winner and the Tideman winner will be the same. □

**Corollary 5.2** *Let $\mathcal{P}$ be a profile for which the Tideman winner is not the DQ-winner. Then all non-zero advantages are odd.*

Proof. As we must have an odd number of agents, from Lemma 3.4 all weights in the majority relation $W^{\mathcal{P}}$ must be odd. Since the $\text{adv}(a, b) = W^{\mathcal{P}}(a, b)^+$ the advantage $\text{adv}(a, b)$ must be zero or equal to the weight $W^{\mathcal{P}}(a, b)$. □

**Lemma 5.3** *There is no profile with three alternatives such that the Tideman winner is not the DQ-winner.*

Proof. The Tideman and Dodgson Quick rules both pick the Condorcet winner when it exists, so if a Condorcet winner exists the Tideman winner and DQ-winner will be the same. It is well known that the absence of a Condorcet winner on three alternatives means that we can rename these alternatives $a, b$ and $c$ so that $\mathrm{adv}(a, b) > 0$, $\mathrm{adv}(b, c) > 0$, and $\mathrm{adv}(c, a) > 0$. These advantages must be odd from the previous corollary. Hence for some $i, j, k \in \mathbb{Z}$ such that $\mathrm{adv}(a, b) = 2i-1$, $\mathrm{adv}(b, c) = 2j-1$, and $\mathrm{adv}(c, a) = 2k-1$. The DQ-Scores and Tideman scores of $a, b, c$ are $i, j, k$ and $2i-1,\ 2j-1,\ 2k-1$ respectively. From here the result is clear, since if $i > j > k$ then $2i-1 > 2j-1 > 2k-1$. □

**Lemma 5.4** *For a profile with four alternatives there does not exist a pair of alternatives such that $a$ is a DQ-winner but not a Tideman winner, and $b$ is a Tideman winner but not a DQ-winner.*

Proof. By way of contradiction assume that such alternatives $a, b$ exist. Thus there is no Condorcet winner, and so for each alternative $c$ there are one to three alternatives $d$ such that $\mathrm{adv}(c, d) > 0$. Also, since the set of Tideman winners and DQ-winners differ, $n$ must be odd and hence all non-zero advantages must be odd. The relationship between the Tideman score $Sc_{\mathbf{t}}(c)$ and the DQ-score $Sc_{\mathbf{q}}(c)$ is as follows:

$$Sc_{\mathbf{t}}(c) = \sum_{d\in A} \mathrm{adv}(c, d) = 2\sum_{d\in A} \left\lceil \frac{\mathrm{adv}(c, d)}{2} \right\rceil - \#\{c : \mathrm{adv}(c, d) \notin 2\mathbb{Z}\} \tag{17}$$

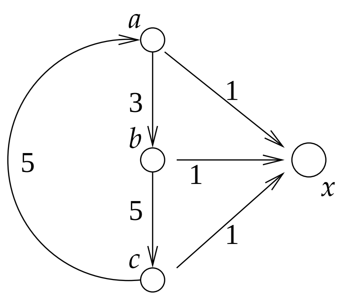
$$= 2Sc_{\mathbf{q}}(c) - (1 \text{ or } 2 \text{ or } 3). \tag{18}$$

Thus, $2Sc_{\mathbf{q}}(c) - 3 \leq Sc_{\mathbf{t}}(c) \leq 2Sc_{\mathbf{q}}(c) - 1$, and, in particular,

$$Sc_{\mathbf{t}}(a) \leq 2Sc_{\mathbf{q}}(a) - 1, \qquad 2Sc_{\mathbf{q}}(b) - 3 \leq Sc_{\mathbf{t}}(b). \tag{19}$$

Given that $a$ is DQ-winner and $b$ is not, we know that $Sc_{\mathbf{q}}(a) \leq Sc_{\mathbf{q}}(b) - 1$. Thus by substitution, $Sc_{\mathbf{t}}(a) \leq 2(Sc_{\mathbf{q}}(b) - 1) - 1 = 2Sc_{\mathbf{q}}(b) - 3 \leq Sc_{\mathbf{t}}(b)$. This shows that if $b$ is a Tideman winner, so is $a$. By contradiction the result must be correct. □

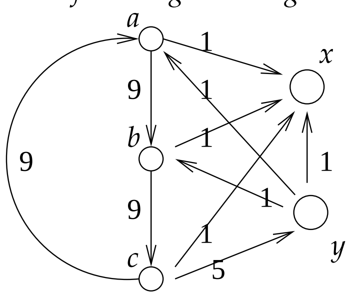
**Example 5.5** *There do exist profiles with four alternatives where the set of tied Tideman winners differs from the set of tied DQ-winners. By Theorem 3.3, we know we may construct a profile whose weighted majority relation has the following advantages:*




| Scores | $a$ | $b$ | $c$ | $x$ |
|---|---|---|---|---|
| Tideman | 5 | 3 | 5 | 3 |
| DQ | 3 | 2 | 3 | 3 |

Here $x, b$ are tied Tideman winners, but $b$ is the sole DQ-winner.

**Example 5.6** *There do exist profiles with five alternatives where there is a unique Tideman winner that differs from the unique DQ-winner. By Theorem 3.3, we know we may construct a profile whose weighted majority relation has the following advantages:*




| *Scores* | $a$ | $b$ | $c$ | $x$ | $y$ |
|---|---|---|---|---|---|
| *Tideman* | *10* | *10* | *9* | *4* | *5* |
| *DQ* | *6* | *6* | *5* | *4* | *3* |

*Here $x$ is the sole Tideman winner, but $y$ is the sole DQ-winner.*

**Theorem 5.7** *For any $m \geq 5$ there exists a profile with $m$ alternatives and an odd number of agents, where the Tideman winner is not the DQ-winner.*

In Example 5.6 we gave an example of a profile with $m = 5$ alternatives for which the Tideman winner is not the Dodgson Quick winner. To extend this example for larger numbers of alternatives, we may add additional alternatives who lose to all of $a$, $b$, $c$, $x$, $y$. From Theorem 3.3 and Lemma 3.4, there exists a profile with an odd number of agents that generates that weighted majority relation.

**Theorem 5.8** *If the number of agents is even, the probability that all of the advantages are 0 does not converge to 0 faster than $\mathcal{O}(n^{-\frac{m!}{4}})$.*

Proof. Let $\mathcal{P}$ be a random profile, $V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m!}\}$ be an ordered set containing all $m!$ possible linear orders on $m$ alternatives, and $X$ be a random vector, with elements $X_i$ representing the number of occurrences of $\mathbf{v}_i$ in $\mathcal{P}$. Under the Impartial Culture assumption, $X$ is distributed according to a multinomial distribution with $n$ trials and $m!$ possible outcomes. Let us group the $m!$ outcomes into $m!/2$ pairs $S_i = \{\mathbf{v}_i, \bar{\mathbf{v}}_i\}$. Denote the number of occurrences of $\mathbf{v}$ as $n(\mathbf{v})$. Let the random variable $Y_i^1$ be $n(\mathbf{v}_i)$ and $Y_i^2$ be $n(\bar{\mathbf{v}}_i)$. Let $Y_i = Y_i^1 + Y_i^2$.

It is easy to show that, given $Y_i = y_i$ for all $i$, each $Y_i^1$ is independently binomially distributed with $p = 1/2$ and $y_i$ trials. It is also easy to show that for an arbitrary integer $n > 0$, a $(2n, 0.5)$-binomial random variable $X$ has a probability of at least $\frac{1}{\sqrt{2n}}$ of equaling $n$; thus if $y_i$ is even then the probability that $Y_i^1 = Y_i^2$ is at least $\frac{1}{2\sqrt{y_i}}$. Combining these results we get

$$P(\forall_i Y_i^1 = Y_i^2 \mid \forall_i Y_i = y_i \in 2\mathbb{Z}) \quad \geq \quad \prod_i \frac{1}{2\sqrt{y_i}} \geq \prod_i \frac{1}{2\sqrt{n}} = 2^{-\frac{m!}{2}} n^{-\frac{m!}{4}}. \tag{20}$$

It is easy to show that for any $k$-dimensional multinomially distributed random vector, the probability that all $k$ elements are even is at least $2^{-k+1}$; hence the probability that all $X_i$ are even is at least $2^{-k+1}$ where $k = m!/2$. Hence

$$P(\forall_i X_{i,1} = X_{i,2}) \quad \geq \quad \left(2^{-\frac{m!}{2}+1}\right)\left(2^{-\frac{m!}{2}} n^{-\frac{m!}{4}}\right) = 2^{1-m!} n^{-\frac{m!}{4}}. \tag{21}$$

If for all $i$, $X_{i,1} = X_{i,2}$ then for all $i$, $n(\mathbf{v}_i) = n(\bar{\mathbf{v}}_i)$, i.e. the number of each type of vote is the same as its complement. Thus

$$n_{ba} = \sum_{\mathbf{v} \in \{\mathbf{v}: b\mathbf{v}a\}} n(v) = \sum_{\bar{\mathbf{v}} \in \{\bar{\mathbf{v}}: a\bar{\mathbf{v}}b\}} n(\bar{v}) = \sum_{\mathbf{v} \in \{\mathbf{v}: a\mathbf{v}b\}} n(v) = n_{ab}, \tag{22}$$

so $\mathrm{adv}(b, a) = 0$ for all alternatives $b$ and $a$. □

**Lemma 5.9** *The probability that the Tideman winner is not the DQ-winner does not converge to 0 faster than $\mathcal{O}(n^{-\frac{m!}{4}})$ as the number of agents $n$ tends to infinity.*

Let $\mathcal{P}$ be a random profile from $\mathcal{L}(A)^n$ for some odd number $n$. Let $|C|$ be the size of the profile from Theorem 5.7. Let us place the first $|C|$ agents from profile $\mathcal{P}$ into sub-profiles $C$ and the remainder of the agents into sub-profile $D$. There is a small but constant probability that $C$ forms the example from Theorem 5.7, resulting in the Tideman winner of $C$ differing from its DQ-winner. As $n$, $|C|$ are odd, $|D|$ is even. Thus from Theorem 5.8 the probability that the advantages in $D$ are zero does not converge to 0 faster than $\mathcal{O}(n^{-\frac{m!}{4}})$. If all the advantages in $D$ are zero then adding $D$ to $C$ will not affect the Tideman or DQ-winners. Hence the probability that the Tideman winner is not the DQ-winner does not converge to 0 faster than $\mathcal{O}(n^{-\frac{m!}{4}})$.

**Theorem 5.10** *The probability that the Tideman winner is not the Dodgson winner does not converge to 0 faster than $\mathcal{O}(n^{-\frac{m!}{4}})$ as the number of agents $n$ tends to infinity.*

Proof. From Corollary 4.6 the DQ-winner converges to the Dodgson winner exponentially fast. However, the Tideman winner does not converge faster than $\mathcal{O}(n^{-\frac{m!}{4}})$ to the DQ-winner, and hence also does not converge faster than $\mathcal{O}(n^{-\frac{m!}{4}})$ to the Dodgson winner. □

Our next goal is to prove that under the IC the probability that the Tideman winner and Dodgson winner coincide converges asymptotically to 1.

**Definition 5.11** *We define the adjacency matrix $M$, of a linear order $\mathbf{v}$, as follows:*

$$M_{ij} = \begin{cases} 1 & \textit{if} \quad i\mathbf{v}j \\ -1 & \textit{if} \quad j\mathbf{v}i \\ 0 & \textit{if} \quad i = j \end{cases} . \tag{23}$$

**Lemma 5.12** *Suppose that $\mathbf{v}$ is a random linear order chosen from the uniform distribution on $\mathcal{L}(A)$. Then its adjacency matrix $M$ is an $m^2$-dimensional random variable satisfying the following equations for* *all* $i, j, r, s \in A$.

$$\begin{aligned} E[M] &= 0 \qquad (24)\\ cov(M_{ij}, M_{rs}) &= E[M_{ij}M_{rs}] \qquad (25)\\ &= \begin{cases} 1 & \text{if} \quad i = r \neq j = s, \\ 1/3 & \text{if} \quad i = r, \text{ but } i, j, s \text{ distinct } \vee \ j = s, \text{ others distinct}, \\ -1/3 & \text{if} \quad i = s, \text{ others distinct } \vee \ j = r, \text{ others distinct}, \\ 0 & \text{if} \quad i, j, r, s \text{ distinct } \vee \ i = j = r = s, \\ -1 & \text{if} \quad i = s \neq j = r. \end{cases} \qquad (26) \end{aligned}$$

Proof. Clearly, $E[M_{ij}] = \frac{(1)+(-1)}{2} = 0$. As $\text{cov}(X, Y) = E[XY] - E[X]E[Y]$ (see e.g. Walpole and Myers 1993; p97), $\text{cov}(M_{ij}, M_{rs}) = E[M_{ij}M_{rs}] - (0)(0) = E[M_{ij}M_{rs}]$. Note that for all $i \neq j$ we know that $M_{ii}M_{ii} = 0$, $M_{ij}M_{ij} = 1$, and $M_{ij}M_{ji} = -1$. If $i = r$ and $i, j, s$ are all distinct then the sign of $M_{ij}M_{is}$ for each permutation of $i, j$ and $s$ is as shown below.

$$\begin{array}{rcccccc} & i & i & j & j & s & s \\ & j & s & i & s & i & j \\ & s & j & s & i & j & i \\ M_{ij} & + & + & - & - & + & - \\ M_{is} & + & + & + & - & - & - \\ M_{ij}M_{is} & + & + & - & + & - & + \end{array} \qquad (27)$$

Thus, $E[M_{ij}M_{rs}] = \frac{+1+1-1+1-1+1}{6} = \frac{1}{3}$.

If $i, j, r, s$ are all distinct then there are six linear orders $\mathbf{v}$ where $i\mathbf{v}j$ and $r\mathbf{v}s$, six linear orders $\mathbf{v}$ where $i\mathbf{v}j$ and $s\mathbf{v}r$, six linear orders $\mathbf{v}$ where $j\mathbf{v}i$ and $r\mathbf{v}s$, and six linear orders $\mathbf{v}$ where $j\mathbf{v}i$ and $s\mathbf{v}r$. Hence,

$$E[M_{ij}M_{rs}] = \frac{6(1)(1)+6(1)(-1)+6(-1)(1)+6(-1)(-1)}{24} = 0 \quad . \qquad (28)$$

We may prove the other cases for $\text{cov}(M_{ij}, M_{rs})$ in much the same way. □

We note that as $\text{var}(X) = \text{cov}(X, X)$ we also have, $\text{var}(M_{ij}) = 1$ if $i \neq j$ and $\text{var}(M_{ij}) = 0$ $i = j$.

**Example 1** *For example, for* $m = 4$ *the covariances with* $M_{12}$ *are shown in the matrix*

$$\mathfrak{L} = \begin{bmatrix} 0 & 1 & 1/3 & 1/3 \\ -1 & 0 & -1/3 & -1/3 \\ -1/3 & 1/3 & 0 & 0 \\ -1/3 & 1/3 & 0 & 0 \end{bmatrix}, \qquad (29)$$

*where* $\mathfrak{L}_{ij} = cov(M_{ij}, M_{12})$.

Define $Y$ to be a collection of random normal variables indexed by $i, j$ for $1 \leq i < j \leq m$ each with mean of 0, and covariance matrix $\Omega$, where

$$\Omega_{ij,rs} \quad = \quad \text{cov}(Y_{ij}, Y_{rs}) = \text{cov}(M_{ij}, M_{rs}), \tag{30}$$

We may use the fact that $i < j$, $r < s$ implies $i \neq j$, $r \neq s$, $(s = i \Rightarrow r \neq j)$ and $(r = j \Rightarrow s \neq i)$ to simplify the definition of $\Omega$ as shown below:

$$\Omega_{ij,rs} \quad = \quad \begin{cases} 1 & \text{if} \quad (r,s) = (i,j), \\ 1/3 & \text{if} \quad r = i, s \neq j \text{ or } s = j, r \neq i, \\ -1/3 & \text{if} \quad s = i \text{ or } r = j, \\ 0 & \text{if} \quad i, j, r, s \text{ are all distinct,} \end{cases} \tag{31}$$

i.e. if $i, j, r, s$ are all distinct then

**Lemma 5.13** *Let $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ be a profile chosen from the uniform distribution on $\mathcal{L}(A)^n$. Let $M_i$ be the adjacency matrix of $P_i$. Then, as $n$ approaches infinity, $\sum_{i=1}^{n} M_i/\sqrt{n}$ converges in distribution to*

$$\begin{bmatrix} 0 & Y_{12} & Y_{13} & \cdots & Y_{1m} \\ -Y_{12} & 0 & Y_{23} & \cdots & Y_{2m} \\ -Y_{13} & -Y_{23} & 0 & \cdots & Y_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -Y_{1m} & -Y_{2m} & -Y_{3m} & \cdots & 0 \end{bmatrix}, \tag{32}$$

*where $Y$ is a collection of random normal variables indexed by $i, j$ for $1 \leq i < j \leq m$ each with mean of 0, and covariance matrix $\Omega$, where*

$$\Omega_{ij,rs} \quad = \quad cov(Y_{ij}, Y_{rs}) = cov(M_{ij}, M_{rs}). \tag{33}$$

Proof. As $M_1, M_2, \ldots, M_n$ are independent identically-distributed (i.i.d.) random variables, we know from the multivariate central limit theorem (see e.g. Anderson, 1984; p81) that $\sum_{i=1}^{n} M_i/\sqrt{n}$ converges in distribution to the multivariate normal distribution with the same mean and covariance as the random matrix $M$ from Lemma 5.12. As $M^T = -M$ and $M_{ii} = 0$, we have the result. □

**Lemma 5.14** *$\Omega$ is non-singular.*

Proof. Consider $\Omega^2$ with elements

$$(\Omega^2)_{ij,kl} \quad = \quad \sum_{1 \leq r < s \leq m} \Gamma_{ij,kl}(r,s), \tag{34}$$

where $\Gamma_{ij,kl}(r,s) = \Omega_{ij,rs}\Omega_{rs,kl}$.

If $i, j, r, s$ distinct, then

$$
\begin{aligned}
\Gamma_{ij,ij}(i,j) = \Omega_{ij,ij}\Omega_{ij,ij} &= (1)(1) = 1, && (35)\\
\Gamma_{ij,ij}(r,j) = \Omega_{ij,rj}\Omega_{rj,ij} &= (1/3)(1/3) = 1/9, && (36)\\
\Gamma_{ij,ij}(i,s) = \Omega_{ij,is}\Omega_{is,ij} &= (1/3)(1/3) = 1/9, && (37)\\
\Gamma_{ij,ij}(r,i) = \Omega_{ij,ri}\Omega_{ri,ij} &= (-1/3)(-1/3) = 1/9, && (38)\\
\Gamma_{ij,ij}(j,s) = \Omega_{ij,js}\Omega_{js,ij} &= (-1/3)(-1/3) = 1/9, && (39)\\
\Gamma_{ij,ij}(r,s) = \Omega_{ij,rs}\Omega_{ij,rs} &= 0. && (40)
\end{aligned}
$$

Let us consider the case $(i, j) = (k, l)$. If $(i, j) = (k, l)$ then

$$
\Gamma_{ij,ij}(r,s) = \Omega_{ij,rs}\Omega_{rs,ij} = \begin{cases} (1)^2 & \text{if} \quad (r,s) = (i,j), \\ (1/3)^2 & \text{if} \quad r = i, s \neq j \text{ or } s = j, r \neq i, \\ (-1/3)^2 & \text{if} \quad s = i, (r \neq j) \text{ or } r = j, (s \neq i), \\ 0 & \text{if} \quad i, j, r, s \text{ are all distinct.} \end{cases} \qquad (41)
$$

Recall that $r < s$, $i < j$ and $r, s \in [1, m]$. Let us consider for how many values of $(r, s)$ each of the above cases occur:

- $(r, s) = (i, j)$: This occurs for exactly one value of $(r, s)$.
- $r = i, s \neq j$: Combining the fact that $r < s$ and $r = i$ we get $i < s$. Thus $s \in (i, j) \cup (j, m]$, and there are $(j - i - 1) + (m - j) = (m - i - 1)$ possible values of $s$. As there is only one possible value of $r$ this means that there are also $(m - i - 1)$ possible values of $(r, s)$.
- $s = j, r \neq i$: Combining the fact that $r < s$ and $s = j$ we get $r < j$. Thus $r \in [1, i) \cup (i, j)$, and there are $(i - 1) + (j - i - 1) = (j - 2)$ possible values of $(r, s)$.
- $s = i$: Here we want $r \neq j$, however $r < s = i < j$, so explicitly stating $r \neq j$ is redundant. Combining the fact that $r < s$ and $s = i$ we get $r < i$. Hence $r \in [1, i]$ and there are $i - 1$ possible values for $(r, s)$.
- $r = j$: Here we want $s \neq i$, however $i < j = r < s$, so explicitly stating that $r \neq j$ is redundant. From here on we will not state redundant inequalities. Combining the fact that $r < s$ and $r = j$ we get $j < s$. Hence $s \in (j, m]$ and there are $m - j$ possible values for $(r, s)$.

Hence,

$$
\begin{aligned}
\sum_{1 \le r < s \le m} \Gamma_{ij,ij}(r,s) &= (1)(1) + ((m - i - 1) + (j - 2))\left(\frac{1}{3}\right)^2 + ((i - 1) + (m - j))\left(\frac{-1}{3}\right)^2 && (42)\\
&= 1 + (m + j - i - 3)\left(\frac{1}{9}\right) + (m + i - j - 1)\left(\frac{1}{9}\right) && (43)\\
&= (9 + (m + j - i - 3) + (m + i - j - 1))/9 && (44)\\
&= \frac{2m + 5}{9}. && (45)
\end{aligned}
$$

Let us consider now the case $i = k,\ j \neq l$. Then

$$\Gamma_{ij,il}(r,s) = \Omega_{ij,rs}\Omega_{rs,il} = \begin{cases} 1\Omega_{rs,il} & \text{if} \quad (r,s) = (i,j), \\ 1/3\Omega_{rs,il} & \text{if} \quad r = i, s \neq j \text{ or } s = j, r \neq i, \\ -1/3\Omega_{rs,il} & \text{if} \quad s = i \text{ or } r = j, \\ 0 & \text{if} \quad i,j,r,s \text{ are all distinct.} \end{cases} \tag{46}$$

more precisely,

$$\Gamma_{ij,il}(r,s) = \begin{cases} (1)(1/3) = 1/3 & \text{if} \quad (i,j) = (r,s), \\ (1/3)(1) = 1/3 & \text{if} \quad r = i, s = l \neq j, \\ (1/3)(1/3) = 1/9 & \text{if} \quad r = i, s \neq j, s =\neq l, \\ (1/3)(0) = 0 & \text{if} \quad s = j \neq l, r \neq i, \\ (-1/3)(-1/3) = 1/9 & \text{if} \quad s = i, \\ (-1/3)(1/3) = -1/9 & \text{if} \quad r = j, s = l, \\ (-1/3)(0) = 0 & \text{if} \quad r = j, s \neq l, \\ 0 = 0 & \text{if} \quad i,j,r,s \text{ are all distinct,} \end{cases} \tag{47}$$

hence,

$$\sum_{1 \le r < s \le m} \Gamma_{ij,il}(r,s) = \frac{1}{3} + \frac{1}{3} + \sum_{1 \le r < s \le m, r=i, s \neq j, s =\neq l} \frac{1}{9} + \sum_{1 \le r < s \le m, s=i} \frac{1}{9} - \frac{1}{9} \tag{48}$$

$$= \frac{1}{3} + \frac{1}{3} + \sum_{i < s \le m} \frac{1}{9} - \frac{2}{9} + \sum_{1 \le r < i} \frac{1}{9} - \frac{1}{9} \tag{49}$$

$$= \frac{1}{3} + (m-i)\frac{1}{9} + (i-1)\frac{1}{9} \tag{50}$$

$$= \frac{m+2}{9}. \tag{51}$$

Similarly for $i \neq k,\ j = l$, we may show $(\Omega^2)_{ij,kj} = \frac{m+2}{9}$. If $j = k$ then

$$(\Omega^2)_{ij,kl} = -\frac{1}{3} - \frac{1}{3} + \frac{1}{9} - \sum_{1 \le r < i, r \neq i} \frac{1}{9} - \sum_{j < s \le m, s \neq l} \frac{1}{9}, \tag{52}$$

$$= -\frac{m+2}{9}, \tag{53}$$

similarly for $l = i$. If $i, j, k, l$ are all distinct, $(\Omega^2)_{ij,kl}$ equals 0. Consequently

$$\Omega^2 = \left(\frac{m+2}{3}\right)\Omega - \left(\frac{m+1}{9}\right)I. \tag{54}$$

Since the matrix $\Omega$ satisfies $\Omega^2 = \alpha\Omega + \beta I$ with $\beta \neq 0$ it has an inverse, hence $\Omega$ is not singular. □

**Theorem 5.15** *The probability that the Tideman winner and Dodgson winner coincide converges asymp-*

*totically to* $1$ *as* $n \to \infty$.

Proof. We will prove that the Tideman winner asymptotically coincides with the Dodgson Quick winner. The Tideman winner is the alternative $a \in A$ with the minimal value of

$$G(a) = \sum_{b \in A} \text{adv}(b, a), \tag{55}$$

while the DQ-winner has minimal value of

$$F(a) = \sum_{b \in A} \left\lceil \frac{\text{adv}(b, a)}{2} \right\rceil. \tag{56}$$

Let $a_T$ be the Tideman winner and $a_Q$ be the DQ-winner. Note that $G(c) - m \leq 2F(c) \leq G(c)$ for every alternative $c$. If for some $b$ we have $G(b) - m > G(a_T)$, then $2F(b) \geq G(b) - m > G(a_T) \geq 2F(a_T)$ and so $b$ is not a DQ-winner. Hence, if $G(b) - m > G(a_T)$ for all alternatives $b$ distinct from $a_T$, then $a_T$ is also the DQ-winner $a_Q$. Thus,

$$P(a_T \neq a_Q) \leq P(\exists_{a \neq b} |G(a) - G(b)| \leq m) \tag{57}$$

$$= P\left(\exists_{a \neq b} \left| \frac{G(a) - G(b)}{\sqrt{n}} \right| \leq \frac{m}{\sqrt{n}}\right), \tag{58}$$

thus for any $\epsilon > 0$ and sufficiently large $n$, we have

$$P(a_T \neq a_Q) \leq P\left(\exists_{a \neq b} \left| \frac{G(a) - G(b)}{\sqrt{n}} \right| \leq \epsilon\right). \tag{59}$$

We will show that the right-hand side of the inequality above converges to 0 as $n$ tends to $\infty$. All probabilities are non-negative so $0 \leq P(a_T \neq a_Q)$. From these facts and the sandwich theorem it will follow that $\lim_{n \to \infty} P(a_T \neq a_Q) = 0$.

Let

$$G_j = \sum_{i<j} (Y_{ij})^+ + \sum_{k>j} (-Y_{jk})^+, \tag{60}$$

where variables $Y_{ij}$ come from the matrix (32) to which $\sum_{i=1}^{n} M_i/\sqrt{n}$ converges by Lemma 5.13 . Thus,

$$\lim_{n \to \infty} P\left(\exists_{a \neq b} \left| \frac{G(a) - G(b)}{\sqrt{n}} \right| \leq \epsilon\right) = P\left(\exists_{i \neq j} |G_i - G_j| \leq \epsilon\right) \tag{61}$$

Since $\epsilon > 0$ is arbitrary,

$$\lim_{n \to \infty} P(a_T \neq a_Q) \leq P(\exists_{i \neq j} G_i = G_j). \tag{62}$$

For fixed $i < j$ we have

$$G_i - G_j = -Y_{ij} + \sum_{k<i} (-Y_{ki})^+ + \sum_{k>i, k\neq i} (Y_{ik})^+ - \sum_{k<j, k\neq i} (Y_{kj})^+ - \sum_{k>j} (-Y_{jk})^+ . \quad (63)$$

Define $v$ so that $G_i - G_j = -Y_{ij} + v$. Then $P(G_i = G_j) = P(Y_{ij} = v) = E[P(Y_{ij} = v \mid v)]$. Since $Y$ has a multivariate normal distribution with a non-singular covariance matrix $\Omega$, it follows that $P(Y_{ij} = v \mid v) = 0$. That is, $P(G_i = G_j) = 0$ for any $i, j$ where $i \neq j$. Hence $P(\exists_{i\neq j} G_i = G_j) = 0$. As discussed previously in this proof, we may now use the sandwich theorem to prove that $\lim_{n\to\infty} P(a_T \neq a_Q) = 0$. □

## 6 Numerical Results

In this section we present tables demonstrating the rate of convergence to Dodgson of the Dodgson Quick rule introduced in this paper in comparison to the Tideman rule. These tables show that the convergence of the Tideman winner to the Dodgson Winner occurs much slower than the exponential convergence of the DQ-Winner. We also study the asymptotic limit of the probability that the Simpson winner is the Dodgson winner as we increase the number of agents.

Table 1: Number of occurrences per 10,000 Elections with 5 alternatives that the Dodgson Winner was not chosen

| Voters | 3 | 5 | 7 | 9 | 15 | 17 | 25 | 85 | 257 | 1025 |
|---|---|---|---|---|---|---|---|---|---|---|
| DQ | 1.5 | 1.9 | 1.35 | 0.55 | 0.05 | 0.1 | 0 | 0 | 0 | 0 |
| Tideman | 1.5 | 2.3 | 2.7 | 3.95 | 6.05 | 6.85 | 7.95 | 8.2 | 5.9 | 2.95 |
| Simpson | 57.6 | 65.7 | 62.2 | 57.8 | 48.3 | 46.6 | 41.9 | 30.2 | 23.4 | 21.6 |

In these 10,000 simulations we were breaking ties according to the preferences of the first agent. In Table 2 we present the results of another 10,000 simulations in which we consider the rules as social choice correspondences and do not break ties.

Table 2: Number of Occurrences per 10,000 Elections with 5 Alternatives that the Set of Dodgson Winners is Not Chosen

| Voters | 3 | 5 | 7 | 9 | 15 | 17 | 25 | 33 | 85 | 257 | 1025 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| DQ | 4.31 | 4.41 | 3.21 | 1.94 | 0.27 | 0.08 | 0.04 | 0 | 0 | 0 | 0 |
| Tideman | 4.31 | 5.57 | 7.31 | 8.43 | 12.73 | 13.15 | 15.46 | 16.35 | 15.18 | 10.2 | 5.4 |

Another question is how well does Dodgson Quick approximate the Dodgson rule when the number of alternatives is different from 5 or when the number of agents is not large in comparison to the number of agents. From Table 3, it appears that the DQ-approximation is still reasonably accurate under these conditions. This table was generated by averaging 10,000 simulations, and splitting ties according to the preferences of the first agent.

Table 3: Frequency that the DQ-Winner is the Dodgson Winner

| # Alternatives \ # Agents | 3 | 5 | 7 | 9 | 15 | 25 | 85 |
|---|---|---|---|---|---|---|---|
| 3 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 5 | 0.9984 | 0.9976 | 0.9980 | 0.9992 | 0.9999 | 1.0000 | 1.0000 |
| 7 | 0.9902 | 0.9875 | 0.9879 | 0.9933 | 0.9980 | 0.9995 | 1.0000 |
| 9 | 0.9792 | 0.9742 | 0.9778 | 0.9837 | 0.9924 | 0.9978 | 0.9999 |
| 15 | 0.9468 | 0.9327 | 0.9338 | 0.9412 | 0.9571 | 0.9743 | 0.9988 |
| 25 | 0.8997 | 0.8718 | 0.8661 | 0.8731 | 0.8971 | 0.9265 | 0.9840 |

To give meaning to these figures, let us compare them with the figures in Tables 4 and 5. We see that even where the number of agents is not very large, the Dodgson Quick rule seems to do a slightly better job of approximating the Dodgson rule than Tideman's approximation. We also see that Simpson's rule does a particularly poor job of approximating the Dodgson winner when the number of alternatives is large.

Table 4: Frequency that the Tideman winner is the Dodgson winner

| # Alternatives \ # Agents | 3 | 5 | 7 | 9 | 15 | 25 | 85 |
|---|---|---|---|---|---|---|---|
| 3 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 5 | 0.9984 | 0.9974 | 0.9961 | 0.9972 | 0.9936 | 0.9917 | 0.9930 |
| 7 | 0.9902 | 0.9864 | 0.9852 | 0.9868 | 0.9845 | 0.9805 | 0.9847 |
| 9 | 0.9792 | 0.9730 | 0.9724 | 0.9731 | 0.9718 | 0.9760 | 0.9815 |
| 15 | 0.9468 | 0.9292 | 0.9263 | 0.9273 | 0.9379 | 0.9485 | 0.9649 |
| 25 | 0.8997 | 0.8691 | 0.8620 | 0.8625 | 0.8833 | 0.9113 | 0.9534 |

Table 5: Frequency that the Simpson Winner is the Dodgson Winner

| # Alternatives \ # Agents | 3 | 5 | 7 | 9 | 15 | 25 | 85 |
|---|---|---|---|---|---|---|---|
| 3 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 5 | 0.9433 | 0.9307 | 0.9339 | 0.9398 | 0.9493 | 0.9575 | 0.9714 |
| 7 | 0.8734 | 0.8627 | 0.8689 | 0.8786 | 0.9018 | 0.9153 | 0.9404 |
| 9 | 0.8256 | 0.8153 | 0.8167 | 0.8251 | 0.8562 | 0.8808 | 0.9124 |
| 25 | 0.5895 | 0.5772 | 0.6147 | 0.6322 | 0.7114 | 0.7529 | 0.7957 |

It appears that Simpson's rule is not a very accurate approximation of Dodgson's Rule. The probability that the Simpson winner does not equal the Dodgson winner is much greater than for

Tideman or DQ. We may ask, does the Simpson rule eventually converge to the Dodgson rule as we increase the number of voters, and, if not, how close does it get?

From Lemma 4.2 and Theorem 5.15 we know that the Dodgson winner, Dodgson Quick winner, and Tideman winner all asymptotically converge as we increase the number of agents. Hence we may compute the asymptotic probability that the Simpson winner is equal to the Dodgson winner, by computing the asymptotic probability that the Simpson winner equals the Tideman winner.

From Lemma 5.13 we know that the matrix of advantages converges to a multivariate normal distribution as we increase the number of agents. If we had a multivariate normal random vector generator, we could use this model to perform simulations and count in how many simulations the Simpson winner is equal to the Tideman winner. We decided to use a slightly different model so that we could use a univariate normal random number generator.

Let $\mathbf{v}$ be a linear order on $A$. As usual, by $\bar{\mathbf{v}}$ we denote the linear order on $A$ where all preferences are reversed. Let $a, b \in A$ be two distinct alternatives. Let $V = (v_1, v_2, \ldots, v_{m!/2})$ be the set of all linear orders from $\mathcal{L}(A)$ where $a$ is ranked above $b$. Note that $\{v_1, \bar{v}_1, v_2, \bar{v}_2, \ldots, v_{m!/2}, \bar{v}_{m!/2}\}$ is the set $\mathcal{L}(A)$ of all possible linear orders of $A$. Given a random linear order $\mathbf{v}$ chosen from the uniform distribution on $\mathcal{L}(A)$, we define an $(m!/2)$-dimensional random vector $X = (X_1, X_2, \ldots, X_{m!/2})$ so that

$$X_i = \begin{cases} 1 & \text{if } \mathbf{v} = v_i \\ -1 & \text{if } \mathbf{v} = \bar{v}_i \\ 0 & \text{otherwise} \end{cases} \tag{64}$$

Let $\mathcal{P} = (P_1, P_2, \ldots, P_n) \in \mathcal{L}(A)^n$ be a profile with $m$ alternatives and $n$ agents. Let $X^i$ be the random vector corresponding to $P_i$. These random vectors are independently identically distributed (i.i.d.) with means of $0$, and covariance matrix $\Omega = rI_m$ where $r$ is some real number greater than $0$ and $I_m$ is the identity $m \times m$ matrix. By the multivariate central limit theorem, we know that $Y = \sum_{i=1}^{n} X^i/\sqrt{n}$ converges to an $N(0, rI_m)$ multivariate normal distribution. Hence we may model $Y_1, Y_2, \ldots, Y_n$ as i.i.d. univariate normally distributed variables.

Using this model we performed 100,000 simulations and generate Table 6.

Table 6: Number of Occurrences per 1000 Elections that the Simpson Winner is Not the Dodgson Winner. (Limit as $n \to \infty$)

| #Alternatives | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|
| #(DO $\neq$ SI) per 1000 | 0 | 6.81 | 17.18 | 27 | 39.33 | 50.18 |

Note that as the number of agents approaches infinity, the probability of a tie approaches $0$, and so tie breaking is irrelevant in this table. In Table 6, we see that even with an infinite number of voters, the Simpson rule is not especially close to the Dodgson rule.

## 7 Conclusion

In this paper we showed that under the Impartial Culture assumption the Tideman rule approximates Dodgson's rule and converges to it, when the number of agents tends to infinity. However we discovered that a new rule, which we call Dodgson Quick, approximates Dodgson's rule much better and converges to it much faster. We also show that Simpson's rule does not converge to Dodgson's rule asymptotically despite often selecting the same winner. The Dodgson Quick rule is computationally very simple however in our simulations it picked the Dodgson winner in all of 1,000,000 elections with 85 agents and 5 alternations. We give numerical results illustrating the rate of convergence of Dodgson Quick to Dodgson.

We proved a generalisation of McGarvey's theorem from ordinary tournaments to weighted tournaments and used it as a tool for constructing examples where the aforementioned rules select different winners.

The most interesting question for further research that this paper rises is whether or not the Dodgson Quick rule approximates Dodgson's rule under the Impartial Anonimous Culture assumption and other models for the population.

## References


Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. John Wileys and Sons, Brisbane, 2nd edition, 1984.

Bartholdi, III., Tovey, C. A., and Trick, M. A. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare: Springer-Verlag*, 6:157–165, 1989. Industrial and Systems Engineering, Georgia Institute of Technology.

Berg, S. Paradox of voting under an urn model: The effect of homogeneity. *Public Choice*, 47:377–387, 1985.

Black, D. *Theory of commitees and elections*. Cambridge University Press, Cambridge, 1958.

Debord, B. *Axiomatisation de procédures d'agrégation de préférences*. Ph.D. thesis, Université Scientifique, Technologique et Médicale de Grenoble, 1987.

Dembo, A. and Zeitouni, O. *Large deviations techniques*. Johns and Barlett, 1993.

Dodgson, C. L. *A method for taking votes on more than two issues*. Clarendon Press, Oxford, 1876. Reprinted in (Black, 1958) with discussion.

Hemaspaandra, E., Hemaspaandra, L., and Rothe, J. Exact analysis of dodgson elections: Lewis carroll's 1876 voting system is complete for parallel access to np. *Journal of the ACM*, 44(6):806–825, 1997.

Homan, C. M. and Hemaspaandra, L. A. Guarantees for the success frequency of an algorithm for finding dodgson-election winners. 2005.

Laslier, J.-F. *Tournament solutions and majority voting*. Springer, Berlin - New York, 1997.

M^c^Cabe-Dansted, J. C. and Slinko, A. Exploratory analysis of similarities between social choice rules. *Group Decision and Negotiation*, 15:1–31, 2006. http://dx.doi.org/10.1007/s00355-005-0052-4.

M^c^Garvey, D. C. A theorem on the construction of voting paradoxes. *Econometrica*, 21:608–610, 1953.

Simpson, P. B. On defining areas of voter choice: Professor tullock on stable voting. *The Quarterly Journal of Economics*, 83(3):478–90, 1969.

Tideman, T. N. *Social Choice and Welfare*, 4:185–206, 1987.

Vidu, L. An extension of a theorem on the aggregation of separable preferences. *Social Choice and Welfare*, 16(1):159–67, 1999.

Walpole, R. E. and Myers, R. H. *Probability and Statistics for Engineers and Scientists*. Maxwell Macmillian International, Sydney, 1993.