



**HAL**  
open science

## On following behaviour as a mechanism for collective movement

Aina Astudillo Fernandez, Jean Louis Deneubourg

► **To cite this version:**

Aina Astudillo Fernandez, Jean Louis Deneubourg. On following behaviour as a mechanism for collective movement. *Journal of Theoretical Biology*, 2011, 284 (1), pp.7. 10.1016/j.jtbi.2011.06.001 . hal-00720865

**HAL Id: hal-00720865**

**<https://hal.science/hal-00720865>**

Submitted on 26 Jul 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Author's Accepted Manuscript

On following behaviour as a mechanism for collective movement

Aina Astudillo Fernandez, Jean Louis Deneubourg

PII: S0022-5193(11)00299-2  
DOI: doi:10.1016/j.jtbi.2011.06.001  
Reference: YJTBI6502

To appear in: *Journal of Theoretical Biology*

Received date: 7 December 2010  
Revised date: 11 May 2011  
Accepted date: 1 June 2011

Cite this article as: Aina Astudillo Fernandez and Jean Louis Deneubourg, On following behaviour as a mechanism for collective movement, *Journal of Theoretical Biology*, doi:[10.1016/j.jtbi.2011.06.001](https://doi.org/10.1016/j.jtbi.2011.06.001)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



[www.elsevier.com/locate/jtbi](http://www.elsevier.com/locate/jtbi)



24 follow each other leads a large fraction of the individuals to the non pre-  
25 ferred side. Moreover, this fraction increases with the total population size.  
26 Finally, we include the possibility to follow  $N$  individuals. This provides the  
27 link between our model and other collective migration models. If enough in-  
28 dividuals are perceived, the results shift from symmetrical (low cohesion) to  
29 asymmetrical (high cohesion) distribution of the individuals. All in all, our  
30 results suggest that following the direct predecessor must be complemented  
31 with other cohesive behaviours (involving the perception of more individuals  
32 or a navigation system) to guarantee its efficiency. We discuss our findings  
33 in the context of the different following behaviours covered in the literature.

34 *Keywords:*

35 migration, collective decision making, group cohesion, aggregation, Markov  
36 chains

---

## 37 1. Introduction

38 Social decision making is common among a wide range of species. In  
39 those species, when an individual has to make a choice between a set of  
40 options, its decision is influenced by the decision of conspecifics. When imi-  
41 tation<sup>1</sup> in the broad sense is involved, it can lead to complex group behaviour  
42 (Deneubourg and Goss, 1989). This often results in some level of cohesion  
43 and/or synchronisation (Sumpter, 2006). The range of perception of each

---

<sup>1</sup>By imitation we mean that an individual's probability of choosing one option is in-  
creased if other individuals already chose that option. Depending on the specific context  
it can be referred to as allelomimesis, social facilitation, public information... but here we  
use it in a very broad sense.

44 individual determines in the number of individuals by whom it will be influ-  
45 enced. This factor can therefore be of key importance in the resulting group  
46 cohesion.

47 Many types of decisions can be socially influenced. An example is tem-  
48 poral decisions (when to do something). For instance, pedestrians are influ-  
49 enced by other pedestrians as they decide when to cross the road (Faria et al.,  
50 2010). Sometimes, the imitation of others in temporal decisions can lead  
51 to synchronisation. There are countless examples ranging from flashing in  
52 fireflies (Buck, 1988, Ramirez Ávila et al., 2003) to synchronised alternation  
53 between feeding and vigilance bouts in sheep (Michelena et al., 2006, Gau-  
54 trais et al., 2007) or to synchronized clapping in human audiences (Neda et  
55 al., 2000a, 2000b). Social influence has also been widely studied in the con-  
56 text of spatial decisions, such as where to forage, where to settle or where to  
57 breed (Danchin et al., 2004). Finally, there are directional decisions (which  
58 direction to take). Sometimes, directional decisions and temporal decisions  
59 are not dissociated. Sheep for example, tend to move when an other sheep  
60 moves, in the same direction as the initiator (Pillot et al., 2010). In this  
61 paper, however, we focus only on directional decisions, with no particular  
62 regard for the context of the displacement (where are they going to or why).  
63 More precisely, our interest lies in the mechanisms involved in the decision  
64 at the individual level, and their consequences at the collective level.

65 This ambition to link individual behaviour with collective movement pat-  
66 terns lies at the heart of a number of mathematical modelling studies. The  
67 behaviour at stake is, on the whole, following the other individuals. Under-  
68 standably, the different biological realities require different types of models

69 (for a review see Giardina, 2008). In flocks of birds or schools of fish, each  
70 individual repeatedly updates its position according to the position of others  
71 (Couzin and Krause, 2003). The majority of models of schooling or flock-  
72 ing have been individual based, self-propelled particle (SPP) models (Okubo  
73 1986, Viscek et al., 1995, Czirok et al., 1997, Couzin et al. 2002, Grégoire  
74 and Chaté, 2004).

75 Another important section of the literature on collective movement con-  
76 cerns situations in which the directional decisions are taken sequentially by  
77 each individual, rather than simultaneously. This is typically the case of  
78 species that tend to move in single file such as social caterpillars, ants or army  
79 worms. Some fish have also been observed moving heads-to-tails (Gudger,  
80 1944). Finally, single files occur in any species constrained to move in a se-  
81 quence due to the geometry of the environment, such as sheep in mountain  
82 paths or cows in cattle-handling facilities (Grandin, 1980). The question of  
83 social influence on directional choices is typically studied with binary choice  
84 set-ups in which a bifurcation presents two identical options. The outcome  
85 of such experiments is either a binomial bell-shaped distribution (no social  
86 influence) or a U-shaped distribution, with most experiments resulting in the  
87 collective selection of one of the two options (social influence). This simple  
88 experiment has put forward collective movement in species as diverse as ants  
89 (Deneubourg et al., 1990), fish (Ward et al., 2008), dust mites (Mailleux  
90 et al., 2011), sheep (Michelena et al., 2010), caterpillars (Dussutour et al.,  
91 2008), spiders (Jeanson et al., 2004, Saffre et al., 1999) and earth worms  
92 (Zirbes et al., 2010).

93 The mathematical models used to describe such collective choices are

94 often individual-based simulations coding for the particular mechanism in-  
95 volved, for example dragline attachment in the case of spiderlings (Jeanson  
96 et al., 2004). General models covering a wide range of situations have seldom  
97 been formulated. One of these general models, hereafter referred to as the  
98 trail model (Deneubourg et al., 1990) expresses the probability of choosing  
99 one option as an increasing (potentially non-linear) function of the number  
100 of individuals who previously chose that option (see section 4 for more details  
101 on this probability function). The major prediction of this model is that in-  
102 dividuals will distribute themselves asymmetrically between the two options,  
103 if the response function is non-linear. Another, perhaps more abstract model  
104 is Polya's urn model, which makes similar predictions to the trail model (for  
105 given parameter sets). One of the assumptions of these models is that indi-  
106 viduals are able to perceive the choice of every preceding individual. This  
107 assumption is justified in the case of many species, which use modification  
108 of the environment (e.g., laying a physical or chemical trail) as means of  
109 communication. As long as the modification is permanent at the time-scale  
110 considered (no or little evaporation of the trail), then the assumption holds.  
111 However, with means of communication such as sight or hearing (e.g. in  
112 birds or mammals), the assumption ceases to be valid as soon as the total  
113 number of preceding individuals exceeds the range of perception.

114 The extreme case is when perception is limited to one individual, for ex-  
115 ample when animals in a sequence perceive others through tactile cues (con-  
116 tact). This is the case of processionary caterpillars (Fitzgerald, 2003), earth-  
117 worms (Zirbes et al., 2010), spiny lobsters (Herrnkind, 1969, Berill 1975),  
118 or some self-assembling robots (Levi and Kernbach, 2010). In these species

119 each animal follows its direct predecessor. Despite being the simplest case  
120 possible of following behaviour, it has never been studied from a theoretical  
121 point of view. Some car traffic models, called “Follow the leader” (Gazis et  
122 al., 1960), consider the influence of the direct predecessor in decisions about  
123 speed and acceleration, but not about direction.

124 The question we address here, is therefore how does the number of in-  
125 dividuals perceived affect the cohesion of the moving group? We begin by  
126 developing a model for the simplest case: following a single preceding indi-  
127 vidual. We then extend the model to cover situations in which individuals  
128 follow an environmental cue as well as the preceding individual. Finally, the  
129 link between our model and the trail model is provided by a general model  
130 allowing the perception of  $N$  preceding individuals. Our results are discussed  
131 in terms of the efficiency of the different following behaviours covered in the  
132 literature.

## 133 **2. Following the previous individual**

134 The theoretical set-up is a path leading to a bifurcation.  $n$  individuals  
135 have to take the path and choose between left and right, one after the other.  
136 Each individual’s information is limited to its direct predecessor’s choice. It  
137 has a probability  $p$  of following its predecessor, and a probability  $1 - p$  of  
138 taking a different path from its predecessor ( $p > 0.5$  by definition).

139

140 The process is memoryless, in the sense that every choice depends ex-  
141 clusively on the choice of the predecessor. The choices of everyone else that  
142 passed before that previous individual do not have any influence on the focal



143 individual's choice. Therefore, this system can be assimilated to a two-state  
 144 Markov chain, where the two states are the possible choices of the individuals  
 145 (left or right). Since the probability of following the predecessor whether he  
 146 took left or right is the same, the transition matrix is symmetric:

$$\mathbf{T} = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix} \quad (1)$$

147 The entry  $T(i, j)$  in the transition matrix represents the probability of  
 148 transition from state  $i$  to state  $j$ . In other words, it is the probability of  
 149 choosing side  $j$ , given that the previous individual took side  $i$ .

150 The question we address here is whether this mechanism is efficient in  
 151 terms of ensuring cohesion. We suggest two ways of answering this question.  
 152 Firstly, we calculate the proportion of individuals expected to end up on the  
 153 side with the most individuals (hereafter referred to as the "winning side").  
 154 A very cohesive species would have an expected fraction of individuals on  
 155 the winning side that is relatively high, whereas a poorly cohesive species  
 156 would have an expected fraction approaching 0.5. Secondly, we calculate the  
 157 number of individuals that are expected to take the same side successively,  
 158 thereby forming a chain of followers. The expected length of those chains  
 159 represents an accurate estimate of cohesion because all of those successive  
 160 followers will end up in the same place.

161 *Expected fraction of individuals on the winning side*

162

163 Let  $n$  be the total number of individuals that face the binary choice.  
 164 First, we need to calculate the expected absolute number of individuals on

165 the winning side. This can be done by using the definition of the expected  
 166 value in probability theory  $E(X) = \sum_{k=n/2}^n k \cdot P(X = k)$ , where  $k$  can be  
 167 any possible number of individuals on the winning side and ranges from  $n/2$   
 168 to  $n$ .  $P(X = k)$  is the probability of having  $k$  individuals on the winning  
 169 side.

170 This probability is difficult to calculate, because all of the possible se-  
 171 quences leading to  $k$  individuals on a given side do not necessarily have the  
 172 same probability (see an example in Fig. 1). Therefore, we calculate this  
 173 probability  $P(X = k)$  by grouping all of those sequences in groups with the  
 174 same number of switches (an individual taking the opposite side to his prede-  
 175 cessor). The reason for grouping the sequences in such a way is simple: the  
 176 probability of any given sequence depends only on the number of switches  $a$   
 177 and on the total number of individuals  $n$ . Therefore all of the sequences in  
 178 each group have the same probability, that is  $P(X = k) = 0.5 \cdot p^{a-1-s}(1-p)^a$ .  
 179 The sum of the probabilities of all those possible sequences gives us the prob-  
 180 ability of having  $k$  individuals on the winning side:

$$P(X = k) = \sum_{a=0}^{2(n-k)} 0.5 \cdot p^{n-1-a}(1-p)^a f(n, k, a) \quad (2)$$

181 The function  $f(n, k, a)$  gives the number of sequences leading to  $k$  out of  $n$   
 182 individuals on the winning side in which  $a$  switches took place. We propose a  
 183 way to calculate this function  $f(n, k, a)$ . The sequences can have two possible  
 184 configurations depending on the side chosen by the first individual (Fig. 2).  
 185 In the first configuration (a), the number of individuals who chose the left are  
 186 divided into three subgroups, and in configuration (b) into two subgroups.  
 187 For the individuals who took the right, it is the opposite.

188 Fig. 3 shows the exhaustive enumeration of all the sequences in a concrete  
 189 example involving 4 switches. In that example, the number of sequences of  
 190 type (a) is simply the number of different ways there are to make 3 subgroups  
 191 with 3 black elements multiplied by the number of ways there are to make 2  
 192 subgroups out of 4 white elements (same reasoning for type (b)).

193 If we define  $G(x, y)$  as the number of ways to make  $x$  subgroups out of  
 194  $y$  elements, then the total number of sequences  $f(n, k, a)$  can be calculated  
 195 with the expression in Eq. 3. We give some details about this  $G(x, y)$  function  
 196 in appendix A. This expression was verified with simulations (appendix C).

$$\begin{aligned}
 f(n, k, a) = & \overbrace{2 G\left(\left\lceil \frac{a+1}{2} \right\rceil, n-k\right) \cdot G\left(\left\lfloor \frac{a+1}{2} \right\rfloor, k\right)}^{\text{sequences of type (a)}} \\
 & + \overbrace{2 G\left(\left\lfloor \frac{a+1}{2} \right\rfloor, n-k\right) \cdot G\left(\left\lceil \frac{a+1}{2} \right\rceil, k\right)}^{\text{sequences of type (b)}} \quad (3)
 \end{aligned}$$

197 With this formula, we calculated the expected fraction of individuals on  
 198 the winning side, and how it varies with the total number of individuals pass-  
 199 ing through the binary choice set-up. This was done for different values of  
 200 the probability of following. As is shown in Fig. 4, the system tends towards  
 201 a symmetric distribution of the individuals between the two sides, for prob-  
 202 abilities of following that are different from 1. In other words, even though  
 203 individuals follow one another, the fraction of individuals on either side tends  
 204 towards 0.5. In Markovian terms, this is easily verified by calculating the sta-  
 205 tionary distribution of the system, that is, when the probabilities of finding  
 206 the system in any given state do not change over time. At the stationary  
 207 distribution the probability of finding the system in state  $i$  is called  $\pi(i)$  and

208 is given by the formula  $\pi(i) = T(j, i)/(T(j, i) + T(i, j))$ , where  $T(i, j)$  is the  
 209 entry of the transition matrix (Eq. 1) containing the probability of taking  
 210 side  $j$ , given that the previous individual took side  $i$ , ( $i \neq j$ ). In our case,  
 211  $\pi(\text{left}) = \pi(\text{right}) = 0.5$ .

212 *Expected number of consecutive followers*

213 The individuals that follow each other uninterruptedly form a “chain” of  
 214 individuals that all end up on the same side. The mean length of those chains  
 215 gives us an estimate of the cohesiveness of the species: the larger the mean  
 216 chain length, the greater the cohesion. Note that this is true in general, even  
 217 for choices that are not binary, which are more representative of the actual  
 218 nature of most species’ environment. The expected size  $E(C)$  (or mean size)  
 219 of these chains of followers is:

$$E(C) = \sum_{c=1}^{n-1} cp^{c-1}(1-p) + np^{n-1} \quad (4)$$

220 As is shown in Fig. 5A, this expected chain length increases with the  
 221 total number of individuals up to a certain point. From there onwards it is  
 222 independent of the total number of individuals. For an infinite number of  
 223 individuals, it depends exclusively on the probability of following, according  
 224 to the function  $E(C) = 1/(1-p)$  (Fig. 5B). It is noteworthy that the ex-  
 225 pected chain length is hardly affected by following the predecessor, except for  
 226 probabilities of following superior to 0.85. For example, in a group of indi-  
 227 viduals that follow each other with a probability of 0.85, the expected number  
 228 of followers would be of 6.7, whereas in a group moving independently the  
 229 expected number would be 2. Chains of more than 20 individuals are only

230 obtained with probabilities of following superior to 0.95.

231

### 232 3. Following the previous individual and an environmental cue

233 So far, we have considered the probability of following to be indepen-  
 234 dent of the direction taken by the preceding individual. However, in a more  
 235 general case, we can include a bias or an environmental cue that causes the  
 236 individuals to prefer one side over the other. The preferred side will hereafter  
 237 be referred to as *the biased side*. For the first individual, let the probabil-  
 238 ity of taking the biased side be  $b$ . From then onwards, we can consider a  
 239 two-state Markov process in which the Markovian transition matrix becomes  
 240 asymmetric:

$$\mathbf{T} = \begin{pmatrix} p_B & 1 - p_B \\ 1 - p_{NB} & p_{NB} \end{pmatrix} \quad (5)$$

241 In this transition matrix,  $p_B$  and  $p_{NB}$  are the probabilities of following the  
 242 preceding individual towards the biased and the non-biased side respectively.  
 243 For the system to be consistent with the scenario we propose, the parame-  
 244 ters require some constraints: (1) the probability of following an individual  
 245 towards the biased side  $p_B$  is the highest of all probabilities (including  $b$ );  
 246 (2) the probability of taking the non-biased side, even though the previous  
 247 individual took the biased side  $p_{NB}$  is the lowest of all probabilities (3) if the  
 248 probability of following towards the non-biased side is higher than 0.5, the  
 249 attraction to others is higher than the attraction to the environmental cue  
 250 causing the bias (and vice-versa).

251 In this situation, the expected number of individuals on the biased side  
 252 is:

$$E(X) = \sum_{k=0}^n k \cdot P(K = k) \quad (6)$$

253 The probability of having  $k$  individuals on the biased side  $P(K = k)$  is  
 254 explained in appendix B.

255 We calculated the expected proportion of individuals on the biased side  
 256 for different values of the three parameters  $p_B$ ,  $p_{NB}$  and  $b$ . Unsurprisingly,  
 257 the majority of individuals always ends up on the biased side. The expected  
 258 proportion tends towards a steady value as the total number of individu-  
 259 als increases (Fig. 6). This steady value depends on  $p_B$ ,  $p_{NB}$  (Fig. 7) and  
 260 equals the probability of finding the system in state  $B$ , once the stationary  
 261 distribution is reached, that is  $\pi(B)$ :

$$\pi(B) = (1 - p_{NB}) / (2 - p_B - p_{NB}) \quad (7)$$

262 This value can be greater than the bias (amplification of the bias) or lower  
 263 than the bias (alleviation of the bias), depending on the values of  $p_B$  and  $p_{NB}$   
 264 (Fig.6). This means that for certain parameter sets (satisfying  $\pi(B) < b$ ),  
 265 the tendency to follow each other can lead a large fraction of the popula-  
 266 tion to the non-preferred side. At least, this fraction is larger than it would  
 267 have been if the individuals were not influenced by one another. In these  
 268 situations, the fraction on the non-preferred side increases with the total  
 269 population. Another interesting feature is that for certain values of  $p_B$  and  
 270  $p_{NB}$ , satisfying  $p_{NB} = (2b - bp_B - 1) / (b - 1)$ , the expected fraction of in-  
 271 dividuals on the biased side is  $b$ . This result is identical to what it would

272 be, had they chosen independently from one another in a biased set-up. In  
 273 other words, they can be very strongly influenced by each other, and to the  
 274 observer who looks only at the end result, rather than at the dynamics, it  
 275 may seem as if they had just followed a bias.

276

#### 277 4. Following the $N$ previous individuals

278 So far, we have assumed that the range of perception of individuals was  
 279 limited to their direct predecessor. In order to allow for larger perception  
 280 ranges, we generalise our model by defining  $N$ , the number of predecessors  
 281 perceived by the focal individual. The situation we have discussed so far is  
 282 in fact an extreme case where  $N = 1$ , the other extreme being the ability to  
 283 perceive all of the preceding individuals ( $N = n$ ), which is the case of the  
 284 trail model when the trail does not evaporate.

285 Let us go back to a non-biased environment. Each individual perceives  
 286  $N$  preceding individuals, and chooses preferentially the side where most of  
 287 them went. The choice is made according to the choice function of the trail  
 288 model (Deneubourg et al., 1990):

$$p_L = \frac{(\kappa + N_L)^\epsilon}{(\kappa + N_L)^\epsilon + (\kappa + N_R)^\epsilon} \quad \text{and} \quad p_R = \frac{(\kappa + N_R)^\epsilon}{(\kappa + N_L)^\epsilon + (\kappa + N_R)^\epsilon} \quad (8)$$

289 In this choice function,  $p_L$  and  $p_R$  are the probabilities of choosing the  
 290 left and the right, respectively.  $N_L$  and  $N_R$  are the number of perceived  
 291 individuals that went to the left and to the right, respectively. Parameter  $\kappa$   
 292 represents the inherent attractiveness of each side. For high values of  $\kappa$ , the

293 branches are so attractive that the attractiveness added by the social term  
 294 ( $N_L$  or  $N_R$ ) does not greatly influence the probability. Therefore, when  $\kappa$   
 295 tends towards infinity, both probabilities tend towards 0.5, and the resulting  
 296 probability distribution is the binomial distribution. The exponent  $\epsilon$  gives  
 297 an estimate of the system's amplification potential. If  $\epsilon = 0$ , the probability  
 298 distribution is also binomial. This choice function was previously used to  
 299 model trail following in ants (Deneubourg et al., 1990) and in dust mites  
 300 (Mailleux et al. 2011). Actually, in the latter study the function used is a  
 301 slightly modified version, but it shares the same properties, namely the gener-  
 302 ation of great asymmetries in the distribution of individuals when  $\epsilon > 1$ . For  
 303 this trail model, it is possible to calculate the exact probability distribution,  
 304 that is the probability of all possible outcomes, expressed as the probability  
 305  $p(l, r)$  of having  $l$  individuals to the left and  $r$  individuals to the right. The  
 306 initial probabilities are  $p(0, 1) = p(1, 0) = 0.5$ . The probabilities of having a  
 307 negative number of individuals on either side ( $p(x, y)$  with  $x < 0$  or  $y < 0$ )  
 308 are naturally equal to zero. From there onwards the probability distribution  
 309 is given by:

$$p(l, r) = p(l - 1, r)p_L + p(l, r - 1) * p_R \quad (9)$$

310 For numbers of perceived individuals lower than the total number of in-  
 311 dividuals, we used Monte Carlo simulations to generate the probability dis-  
 312 tributions. This was done for various values of the parameters  $N$ ,  $\kappa$ , and  $\epsilon$ .  
 313 Fig. 8 shows a sample of the the parameter sets we used. As expected, the  
 314 lower the inherent attractiveness of the branches  $\kappa$ , the higher the asymme-  
 315 try. Also, when  $\epsilon = 1$ , the average number of individuals on the winning



316 side is relatively low, even when many preceding individuals are perceived  
317 (high values of  $N$ ). Very low values of  $\kappa$  ( $\kappa \leq 1$ ) might make the expected  
318 number on the winning side higher, but this is due to a high probability of  
319 every individual taking the same side. In fact, for amplification to appear,  
320 the exponent  $\epsilon$  has to be higher than one. As Fig. 9 shows, the asymmetries  
321 increase with the number of perceived individuals ( $N$ ).

## 322 5. Discussion

323 In this paper we study the efficiency (in terms of group cohesion) of fol-  
324 lowing the preceding individual. This question is largely absent from the  
325 theoretical literature on collective movements and socially influenced deci-  
326 sions. As we show in this paper, the question is far less trivial to answer  
327 than to ask. Our main results may seem somewhat counter-intuitive: cohe-  
328 sive behaviour at the individual level (following the preceding individual)  
329 does not lead to cohesion at the collective level. Eventually, it leads to a  
330 homogeneous dispersal of the group. Some other results are more intuitive,  
331 such as the positive relationship between cohesiveness and the probability of  
332 following, or the cohesion obtained from increasing the number of perceived  
333 individuals, but they remain worth quantifying. By providing a quantitative  
334 description of the relationship between parameters and variables, our model  
335 offers a set of predictions that experimenters can compare their results with.

336 This study concerns binary-choice experiments made on groups, which  
337 are extremely popular in ethology and psychology. The set-up of these ex-  
338 periments (reviewed in Jeanson and Deneubourg, 2009) can either present  
339 two identical options (as in Maillieux et al. 2011, or Zirbes et al., 2010) or

340 two different options (as in Canonge et al., 2009 or Ward et al. 2008). The  
341 former is an elegant way of putting forward social influence on the choice  
342 of individuals. The latter can be used to study the interplay between social  
343 interaction and individual preference. Our predictions cover both situations,  
344 and can provide insight into the interpretation of such experimental results.  
345 For example, a homogeneous distribution of individuals between two identi-  
346 cal options is typically interpreted as an absence of social interaction. Here,  
347 we show that a homogeneous distribution can happen, even in the presence  
348 of strong social influence, provided that the perception range is limited (by  
349 the nature of the communication used, or by the experimental set-up) to  
350 a relatively small number of predecessors. Therefore, questions on percep-  
351 tion mechanisms and population density in the set-up must be taken into  
352 account. Furthermore, the models we propose can be used to validate (or  
353 invalidate) hypotheses on the mechanisms by which animals influence one  
354 another. Take for example the hypothesis that individuals follow each other  
355 by contact. It can be tested by calculating the probability of following the  
356 direct predecessor, based on the behaviour of the second individual. From  
357 that probability and our model it is possible to predict the behaviour of the  
358 subsequent individuals (Fig. 4). The comparison between this prediction  
359 and the experimental results could confirm (or invalidate) the hypothesis.  
360 Another conclusion that can be drawn from our results is that it is useful  
361 to calculate the mean length of chains of consecutive followers. Indeed, this  
362 variable can be crucial to distinguish between a situation in which there is  
363 no social interaction and one in which each individual follows its the direct  
364 predecessor. It can also be used to make the difference between following the

365 predecessor and following a trail.

366 The model we study in this paper is general enough to cover a wide range  
367 of situations. Animals that use tactile cues to perceive and follow their pre-  
368 decessor, such as processionary caterpillars (Fitzgerald, 2003), spiny lobsters  
369 (Herrnkind, 1969) or earthworms (Zirbes et al., 2010) are the obvious exam-  
370 ples. But our predictions can extend to self-assembling robots designed to  
371 form chains, or even to humans in situations where information is limited to  
372 one's direct predecessor, such as some situations in car traffic. In fact, the  
373 model can apply to socially influenced decisions that are not directional deci-  
374 sions. For example, it is well known that conformism can make humans very  
375 influenced by other people's decisions (Asch, 1951). Therefore, situations  
376 in which humans have to make decisions sequentially and can perceive the  
377 decisions of others (typically consumer behaviour) can be studied through  
378 the theoretical perspective we propose.

379 More generally, the mathematical questions raised in this study are rele-  
380 vant to any situation that can be assimilated to a two-state Markov chain.  
381 These situations can be found even outside the context of decisions. For ex-  
382 ample, a classical model to predict precipitation is a two-state Markov chain  
383 (Todorovic et al., 1975). Other studies have tackled some of the mathemat-  
384 ical questions raised here in completely different contexts. The question of  
385 the exact expression for the probability distribution discussed in the first  
386 section (Eqs. 2 and 3) is explored from a purely mathematical point of view  
387 in the context of a correlated random walk through the integers (Renshaw  
388 and Henderson 1981). The more general case of the asymmetric environment  
389 was studied from a purely abstract perspective by Gabriel (1959). The exact

390 expressions derived in these two studies are in essence very similar to ours.  
391 However, our objective here is to use these mathematical explorations to infer  
392 properties to a biological system, that is to a group of animals in motion.

393 One interesting feature of our results is the fragility of the cohesion of  
394 groups of followers. Indeed, we show that for probabilities of following  
395  $p < 0.85$ , the expected chains are not drastically longer than they would  
396 be if individuals acted independently of one another (6.7 versus 2). Fur-  
397 thermore, what we define as the probability of following  $p$ , contains in fact  
398 the probability of perceiving multiplied by the probability of following the  
399 preceding individual. In our theoretical set-up, we have implicitly considered  
400 the probability of perceiving the preceding individual to be equal to one. In  
401 natural situations, where the linear density of individuals can be low, this  
402 assumption does not necessarily hold. This potentially decreases the cohe-  
403 sion of groups. In addition to this, even for high probabilities of perceiving  
404 and of following (different from one), large groups will end up homogeneously  
405 distributed across the options. Cohesion actually decreases with the size of  
406 the group.

407 Not only does cohesion decrease with group size, but in some cases ac-  
408 curacy also decreases with group size. Indeed we show that in situations in  
409 which individuals have a preference for one of the options, a large fraction  
410 can end up choosing the non-preferred option, if the probability of following  
411 is large enough. More importantly, we show that this fraction increases with  
412 group size. This is directly opposed to the outcome of other types of collec-  
413 tive decision making models, such as the trail model (Deneubourg 1990) or  
414 quorum responses (Sumpter and Pratt, 2009). In the latter, cohesiveness and

415 most often accuracy increase with group size. The key difference is that in  
416 these models, there are non-linearities that arise from positive feedbacks. In  
417 our memoryless situation, it is impossible for amplification to happen, hence  
418 the importance of perceiving more than one individual, and of non-linear re-  
419 sponse functions to maintain cohesion. Indeed, as we show in section 4, stable  
420 asymmetries in the distribution happen only when the number of perceived  
421 individuals increases and when the exponent of the probability function is  
422 higher than one.

423 Nevertheless, the theoretical results we present must be put into perspec-  
424 tive with the biological reality that they represent. First of all, the cohesive  
425 species that we have talked about display behaviours that potentially increase  
426 the probability of following, or more precisely the probability of perceiving  
427 the predecessor. For example, they can synchronise departures, which in-  
428 creases the potential linear density. It is true that the larger the group,  
429 the lower the cohesion. Nevertheless, the groups in some cases are small  
430 enough to ensure sufficient cohesion. The smallest group size are groups of  
431 two, which is the case of ant recruitment tandems. This type of recruitment  
432 requires constant antennal contact between the following ant and the lead-  
433 ing ant's abdomen (Franks and Richardson, 2006, Möglich et al 1974). It is  
434 very efficient, and most of the time both ants end up in the same food source.  
435 Caterpillars move in larger groups: Edwards (1910) reports processions of up  
436 to 300 individuals. Spiny lobsters also seem to migrate in large groups: every  
437 autumn, in the Bahamas, thousands of Spiny lobsters (Herrnkind, 1969) mi-  
438 grate to the South-West in parallel chains of 3 to 30 individuals (Herrnkind  
439 and Cummings, 1964). It is necessary for them to rely on complementary

440 mechanisms to ensure cohesion. caterpillars use silk and chemical cues as a  
 441 trail (Fitzgerald, 2003), and spiny lobsters are capable of true navigation in  
 442 a magnetic field (Boles and Lohmann, 2003). However it has been shown  
 443 in both cases that tactile cues are the most important cues for maintaining  
 444 cohesion (Fitzgerald 2003, Herrnkind, 1969).

445 In this paper we formulate a model that provides answers to questions  
 446 concerning the efficiency of following behaviour. By studying the simplest  
 447 case possible of following behaviour, we put forward three important factors  
 448 in ensuring cohesion: the number of perceived individuals, the potential for  
 449 amplification, and group size.

#### 450 **Acknowledgements**

451 A.A. is supported by a grant from the FRIA (Fonds pour la Formation  
 452 à la Recherche dans Industrie et l'Agriculture). J.L.D. is research associate  
 453 from the Belgian National Funds for Scientific Research. We also wish to  
 454 thank O.Astudillo Fernandez and D.Lalor for their helpful comments and  
 455 suggestions.

#### 456 **Appendix A. The number of combinations resulting in $A$ switches**

457 Although intuitive, the  $G(x, y)$  function requires a few definitions in order  
 458 to be calculated correctly:

- 459 - There is one way to make 0 subgroups with 0 elements
- 460 - There is no way to make  $x$  subgroups with  $y$  elements if  $x > y$
- 461 - There is no way to make 0 subgroups with  $y$  elements if  $y \neq 0$

- 462 - If  $x$  and/or  $y$  are negative,  $G(x, y) = 0$
- 463 - Otherwise,  $G(x, y) = C_{x-1}^{y-1} = \frac{(y-1)!}{(y-x)!(x-1)!}$

## 464 Appendix B. Generalisation to any two-state Markov chain

465 Let  $P_{(a)}(K = k)$  be the probability of having  $k$  individuals on the biased  
 466 side given that the first one chose the biased side and  $P_{(b)}(K = k)$  the  
 467 probability, given that the first one chose the non-biased side. The probability  
 468  $P(K = k)$  of having  $k$  individuals on the biased side after  $n$  individuals have  
 469 passed is the sum of those two probabilities.

$$\begin{aligned}
 P(K = k) &= P_{(a)}(K = k) + P_{(b)}(K = k) \\
 P_{(a)} &= b \sum_{a=0}^{2(n-k)} p_B^{k - \lceil \frac{a+1}{2} \rceil} p_{NB}^{n-k - \lfloor \frac{a+1}{2} \rfloor} (1-p_B)^{\lceil \frac{a}{2} \rceil} (1-p_{NB})^{\lfloor \frac{a}{2} \rfloor} \\
 &\quad G\left(\left\lceil \frac{a+1}{2} \right\rceil, k\right) \cdot G\left(\left\lfloor \frac{a+1}{2} \right\rfloor, n-k\right) \\
 P_{(b)} &= (1-b) \sum_{a=0}^{2(n-k)} p_B^{k - \lfloor \frac{a+1}{2} \rfloor} p_{NB}^{n-k - \lceil \frac{a+1}{2} \rceil} (1-p_B)^{\lfloor \frac{a}{2} \rfloor} (1-p_{NB})^{\lceil \frac{a}{2} \rceil} \\
 &\quad G\left(\left\lfloor \frac{a+1}{2} \right\rfloor, k\right) \cdot G\left(\left\lceil \frac{a+1}{2} \right\rceil, n-k\right)
 \end{aligned}$$

470 The validity of this expression was verified (appendix C).

## 471 Appendix C. Verification of the exact expressions for the probabil- 472 ity distributions

473 We conducted Monte-Carlo simulations in order to verify that the exact  
 474 expressions for the probability distributions in sections 1 and 2 are correct. In

475 each simulation  $n$  virtual individuals were sequentially faced with the choice  
476 between left and right. The probability of taking left and right depended on  
477 the choice of the preceding individual, and on the environment (homoge-  
478 neous or heterogeneous) according to the transition matrices (Eqs. 1 and 5).  
479 For each individual, a random number was generated. If the random number  
480 was lower than the probability of taking left, the individual's choice would be  
481 left, otherwise it would be right. Repeating this simulation a million times  
482 gave a good approximation of the frequency distribution of each possible out-  
483 come (analogous to a probability distribution). Fig. 10 shows that the exact  
484 expression and the simulations converge to the same distribution.

#### 485 **References**

- 486 Asch, S. E., 1951. Effects of group pressure on the modification and distortion  
487 of judgments. In: H. Guetzkow (Ed.), Groups, leadership and men. Carnegie  
488 Press, Pittsburgh, pp.177-190
- 489 Berill, M., 1975. Gregarious behavior of juveniles of the spiny lobster, *Pan-*  
490 *ulirus argus* (Crustacea: Decapoda). Bulletin of Marine Science 25, 515-522
- 491 Boles, L.C., Lohmann, 2003. True navigation and magnetic maps in spiny  
492 lobsters. Nature 421,60-63
- 493 Buck, J., 1988. Synchronous rhythmic flashing of fireflies.II. The Quarterly  
494 Review of Biology 63,265-289
- 495 Canonge, S., Sempo, G., Jeanson, R., Detrain, C., Deneubourg, J.L., 2009.



- 496 Self-amplification as a source of interindividual variability: shelter selection  
497 in cockroaches. *Journal of Insect Physiology* 55,976-982 doi:10.1016/j.jinsphys.2009.06.011
- 498 Couzin, I.D., Krause, J., James, R., Ruxton, G.D., Franks, N.R., 2002. Col-  
499 lective memory and spatial sorting in animal groups. *Journal of Theoretical*  
500 *Biology* 218,1-11
- 501 Couzin, I.D., Krause, J., 2003. Self-Organization and Collective Behavior in  
502 Vertebrates. *Advances in the Study of Behaviour* 32,1-75 doi:10.1016/S0065-  
503 3454(03)01001-5
- 504 Czirok, A., Stanley, H.E., Vicsek, T., 1997. Spontaneously ordered motion  
505 of self-propelled particles. *Journal of Physics A: Mathematical and General*  
506 30,1375–1385
- 507 Danchin, E., Giraldeau, L-A., Valone, T.J., Wagner, R.H., 2004. Public  
508 information: from nosy neighbours to cultural evolution. *Science* 305,487-  
509 491
- 510 Deneubourg, J.L., Goss, S., 1989. Collective patterns and decision making.  
511 *Ethology Ecology and Evolution* 1,295-311
- 512 Deneubourg, J.L., Aron, S., Goss, S., Pasteels, J.M., 1990. The self-organizing  
513 exploratory pattern of the Argentine ant. *Journal of Insect Behavior* 3,159-  
514 168
- 515 Dussutour, A., Nicolis, S.C., Despland, E., Simpson, S.J., 2008. Individ-  
516 ual differences influence collective behaviour in social caterpillars. *Animal*

- 517 Behaviour 76,5-16, doi:10.1016/j.anbehav.2007.12.009
- 518 Edwards, T.G., 1910. On the procession and pupation of the larva of *Cnetho-*  
519 *campa pinivora*. Proceedings of the Cambridge Philosophical Society 15,431–436
- 520 Faria, J.J., Krause, S., Krause, J., 2010. Collective behavior in road crossing  
521 pedestrians: the role of social information. Behavioral Ecology 21,1236-1242
- 522 Fitzgerald, T.D., 2003. Role of trail pheromone in foraging and procession-  
523 ary behaviour of pine processionary caterpillars *Thaumetopoea pityocampa*.  
524 Journal of Chemical Ecology 29,513-532
- 525 Franks, N.R., Richardson, T., 2006. Teaching in tandem-running ants. Na-  
526 ture 439,153
- 527 Gabriel, K.R., 1959. The distribution of the number of successes in a se-  
528 quence of dependent trials. Biometrika 46,454-460
- 529 Gautrais, J., Michelena, P., Sibbald, A., Bon, R., Deneubourg, J.L., 2007.  
530 Allelomimetic synchronization in Merino sheep. Animal Behaviour 74,1443-  
531 1454, doi:10.1016/j.anbehav.2007.02.020
- 532 Grandin, T., 1980. Observations of cattle behavior applied to the design of  
533 cattle-handling facilities. Applied Animal Ethology 6,19-31
- 534 Gazis, D.C., Herman, R., Rothery, R., 1960. Non-linear follow-the-leader  
535 models of traffic flow. Operations Research 9,545-567

- 536 Giardina, I., 2008. Collective behavior in animal groups: theoretical models  
537 and empirical studies. *HFSP Journal* 2,205-219
- 538 Grégoire, G., Chaté, H., 2004. Onset of collective and cohesive motion.  
539 *Physical Review Letters* 92,025-702
- 540 Gudger, E.W., 1944. Fishes that swim heads to tails in single file. *Copeia*  
541 1944, 152-154
- 542 Herrnkind, W.F., 1969. Queuing behavior of spiny lobsters. *Science* 164,1425-  
543 1427
- 544 Herrnkind, W.F, Cummings, W.C., 1964. Single file migrations of the spiny  
545 lobster *Panulirus argus* (Latreille). *Bulletin of Marine Science of the Gulf*  
546 *and Caribbean* 14,123-125
- 547 Jeanson, R., Deneubourg, J.L., 2009. Positive feedback, convergent collective  
548 patterns and social transitions in arthropods. In: Gadau, Ü., Fewell, J. (Eds.)  
549 *Organization of insect societies: From genome to sociocomplexity*. Harvard  
550 University Press, pp.460-476
- 551 Jeanson, R., Deneubourg, J.L., Theraulaz., G., 2004. Discrete dragline at-  
552 tachment induces aggregation in spiderlings of a solitary species. *Animal*  
553 *Behaviour* 67,531-537, doi:10.1016/j.anbehav.2003.06.013
- 554 Levi, P., Kernbach, S. (Eds.), 2010. Symbiotic multi-robot organisms. *Reli-*  
555 *ability, Adaptability, Evolution Series: Cognitive Systems Monographs*, 480  
556 pp. Springer-Verlag Berlin Heidelberg

- 557 Mailleux, A-C., Astudillo Fernández, A., San Martin, G., Detrain, C., Deneubourg,  
558 J.L., 2011. Collective migration in house dust mites. *Ethology in press*
- 559 Michelena, P., Noël, S., Gautrais, J., Gerard, J-F., Deneubourg, J.L., Bon,  
560 R., 2006. Sexual dimorphism, activity budget and synchrony in groups of  
561 sheep. *Oecologia* 148,170-180
- 562 Michelena, P., Jeanson, R., Deneubourg, J.L., Sibbald, A., 2010. Personality  
563 and collective decision-making in foraging herbivores. *Proceedings of the*  
564 *Royal Society B* 207,1093-1099
- 565 Möglich, M., Maschwitz, U., Hölldobler, B., 1974. Tandem calling: a new  
566 kind of signal in ant communication. *Science* 186,1046-1047
- 567 Néda,Z., Ravasz, E.,Brechet, Y., Vicsek, T., Barabási A-L., 2000a. The  
568 sound of many hands clapping. *Nature* 403, 849
- 569 Néda,Z., Ravasz, E.,Vicsek, T., Brechet, Y., Barabási A-L., 2000b. Physics  
570 of the rhythmic applause. *Physical Review E* 61, 6987-6992
- 571 Okubo, A., 1986. Dynamical aspects of animal grouping. *Advances in Bio-*  
572 *physics* 22, 1-94.
- 573 Pillot, M-H., Gautrais, J., Gouello, J., Michelena, P., Sibbald, A., Bon, R.,  
574 2010. Moving together : Incidental leaders and naïve followers. *Behavioural*  
575 *Processes* 83, 235–241 doi:10.1016/j.beproc.2009.11.006
- 576 Ramirez Ávila, M., Guisset, J.L., Deneubourg, J.L., 2003. Synchronization

- 577 in light-controlled oscillators. *Physica D* 182,254-273
- 578 Renshaw, E., Henderson, R., 1981. The correlated random walk. *Journal of*  
579 *Applied Probability*, 18,403-414
- 580 Saffre, F. Furey, R., Krafft, B., Deneubourg, J.L., 1999. Collective decision-  
581 making in social spiders: dragline-mediated amplification process acts as a  
582 recruitment mechanism. *Journal of Theoretical Biology* 198,507-517  
583 doi:10.1006/jtbi.1999.0927
- 584 Sumpter, D.J.T., 2006. The principles of collective animal behaviour. *Philo-*  
585 *sophical Transactions of the Royal Society B* 361,5-22
- 586 Sumpter, D.J.T., Pratt, S.C., 2009. Quorum responses and consensus deci-  
587 sion making. *Philosophical Transactions of the Royal Society B* 364,743-753
- 588 Todorovic, P., Woolhiser, D.A., 1975. A stochastic model of n-day precipi-  
589 tation. *Journal of applied meteorology* 14,17-24
- 590 Vicsek, T., Czirok, A., Ben-Jacob, E., Cohen, I., and Shochet, O., 1995.  
591 Novel type of phase transition in a system of self-driven particles. *Physical*  
592 *Review Letters* 75,1226-1229
- 593 Ward, A.J.W., Sumpter, D.J.T., Couzin, I.D., Hart, P.J.B., Krause, J., 2008.  
594 Quorum decision-making facilitates information transfer in fish shoals. *Pro-*  
595 *ceedings of the National American Society* 105,6948-6953
- 596 Zirbes, L., Deneubourg, J.L., Brostaux, Y., Haubruge, E., 2010. A new

597 case of consensual decision: collective movement in earthworms. *Ethology*  
598 116,546–55

Accepted manuscript

Figure 1: Sequences of 5 individuals leading to 3 individuals on the right. The sequences of individuals are represented by a sequence of squares, where each square is the choice of one individual: black means it chose the left, white means it chose the right. The different sequences can have different probabilities ( $P$ ), depending on the order of the squares. More precisely, the probability depends on the number of individuals that took a different path than their predecessor (number of switches  $a$ ). the function  $f(n, k, a)$  gives the number of sequences in each subgroup.

Figure 2: If there are four switches, it means that the first one(s) go to the left, then some go to the right, then some more to the left again ending with some on the right, or the opposite starting with the right

Figure 3: All the sequences satisfying  $n = 7$ ,  $k = 4$ , and  $a = 4$ , in which the winning side is the right. It has to be multiplied by two, in order to obtain the total number of sequences  $f(n, k, a)$

Figure 4: Expected fraction of individuals on the winning side according to the total number of individuals. As the total number of individuals increases, the system tends towards symmetry (with the number on the winning side approaching 50%)

Figure 5: Mean length of the chains according to (A.) the total number of individuals and (B.) the probability of following the preceding individual. For every probability of following different from 1, after the total number of choosers exceeds a certain size, the mean length of the chain does no longer depend on the number of choosers. This dependency is strong with probabilities higher than 0.9.

Figure 6: Example of the evolution of the system as individuals pass. For a given value of the bias ( $b = 0.6$ ) and a given value of the tendency to follow the previous individual towards the biased side ( $p_B = 0.7$ ), the bias can be amplified (if  $\pi(B) > b$ , so  $p_{NB} < 0.55$ ) or weakened ( $p_{NB} > 0.55$ )

Figure 7: Proportion of individuals on the biased side for an infinite total number of passing individuals, for various values of  $p_B$  and  $p_{NB}$  and an environmental bias of  $b = 0.6$ .

Figure 8: The expected percentage of individuals on the winning side after 2000 individuals have passed. Different sets of parameters were used in the Monte Carlo simulations. For each parameter set, we ran 2500 simulations and what appears in the figure is the average of the 2500 simulations.

Figure 9: The distribution of the simulation results for a total number of individuals  $n = 200$ , and for a fixed attractivity of each branch ( $\kappa = 8$ ). The distributions change as we move from one extreme ( $N = 1$ ) to the other ( $N = n = 200$ ). For  $\epsilon = 1$ , the distribution remains bell shaped, but spreads around the most likely result (half of the individuals on each side). For  $\epsilon = 2$ , however, the distribution changes from bell-shaped to U-shaped. The most likely results become the asymmetric results (most individuals on one of the two sides).

Figure 10: Probability distributions calculated with exact expressions (solid line) and with simulations (histogram). The parameters taken were  $n = 40$ ,  $p = 0.75$ ,  $b = 0.6$ ,  $p_B = 0.8$  and  $p_{NB} = 0.7$ .



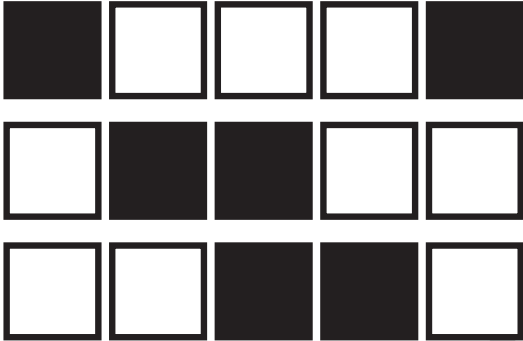
$a = 1$



$$P = 0.5 p (1 - p)^3$$

$$f(5,3,1) = 4$$

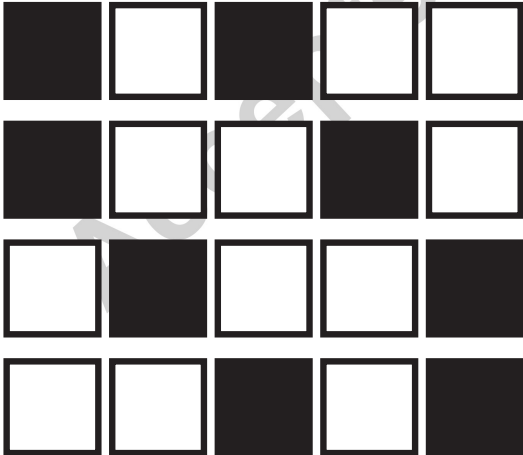
$a = 2$



$$P = 0.5 p^2 (1 - p)^2$$

$$f(5,3,2) = 6$$

$a = 3$



$$P = 0.5 p (1 - p)^3$$

$$f(5,3,3) = 8$$

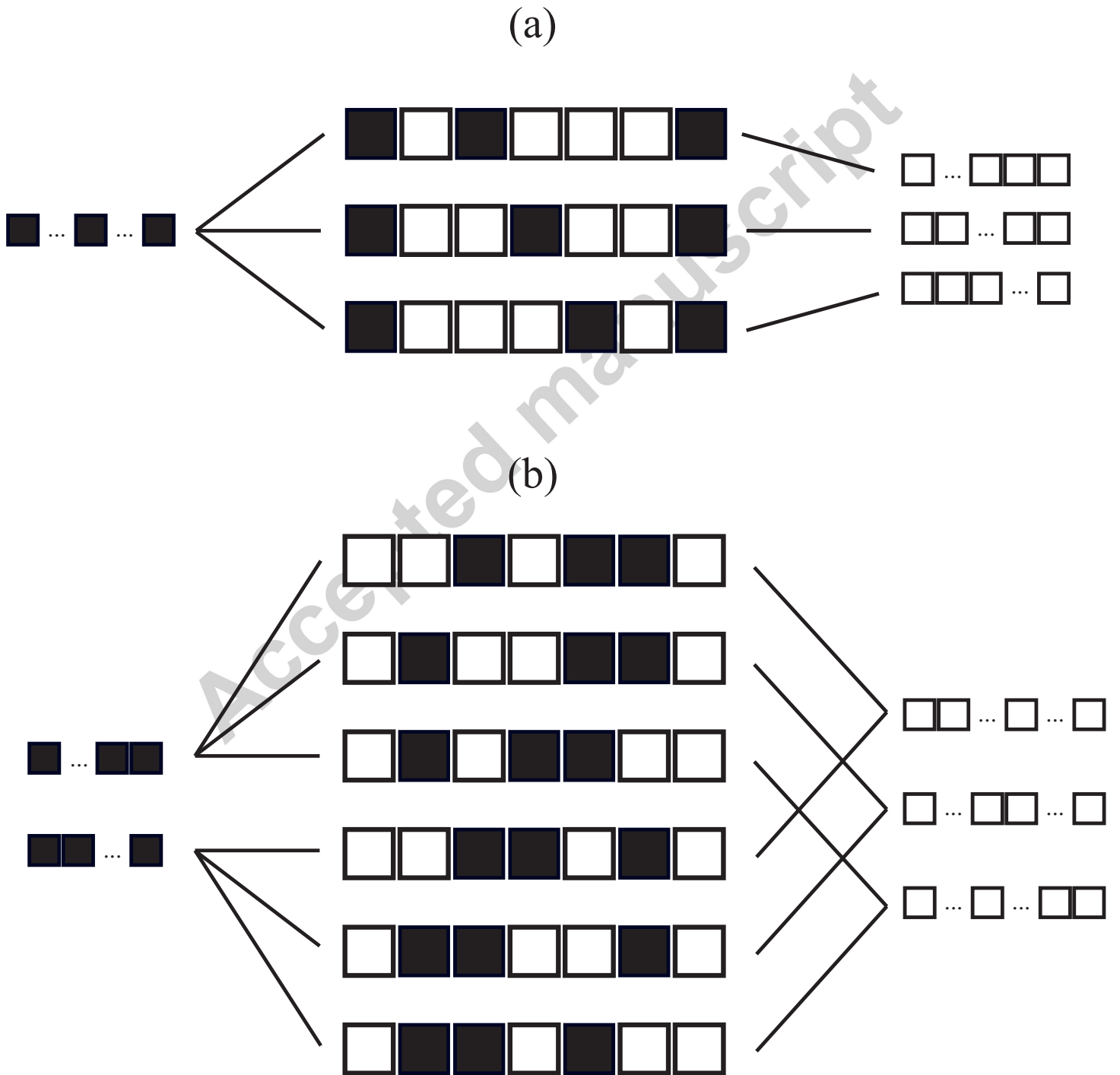
$a = 4$

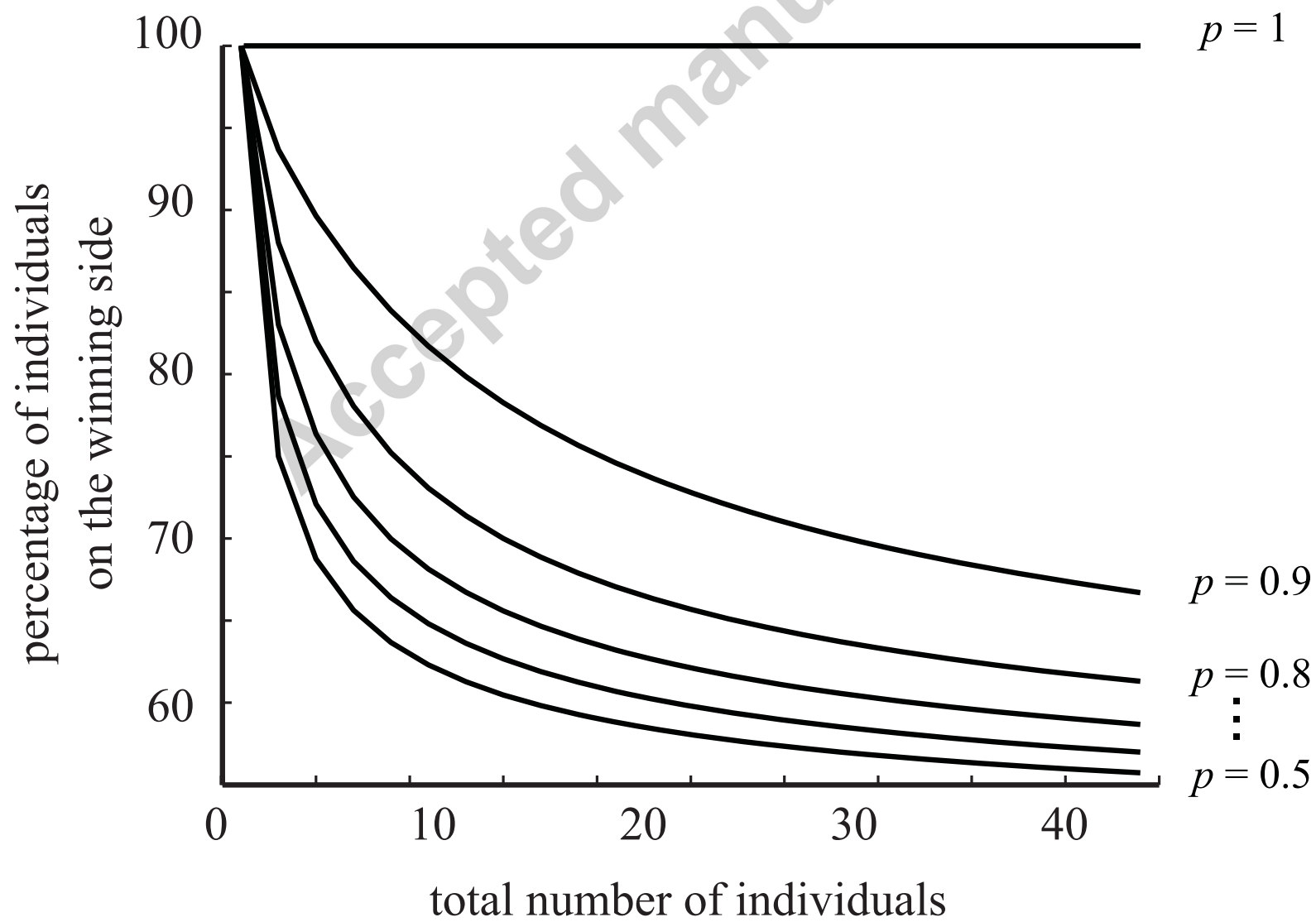


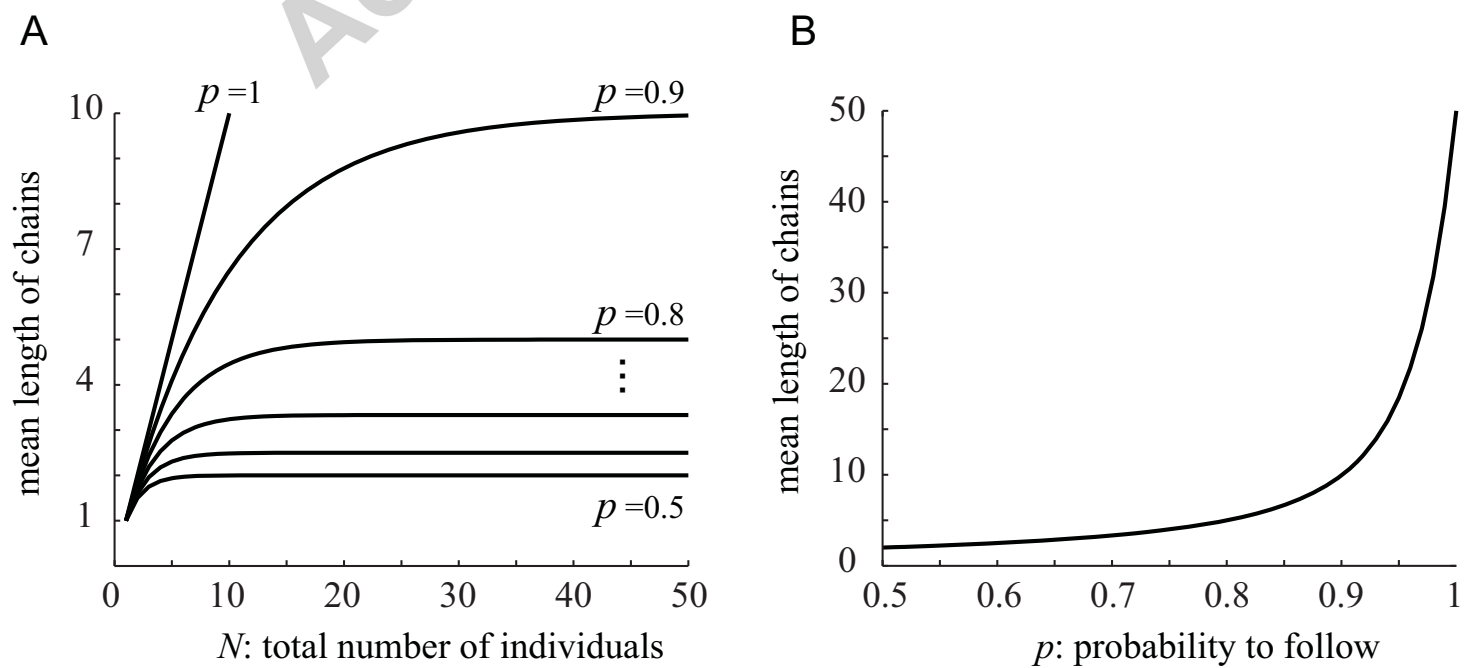
$$P = 0.5 (1 - p)^4$$

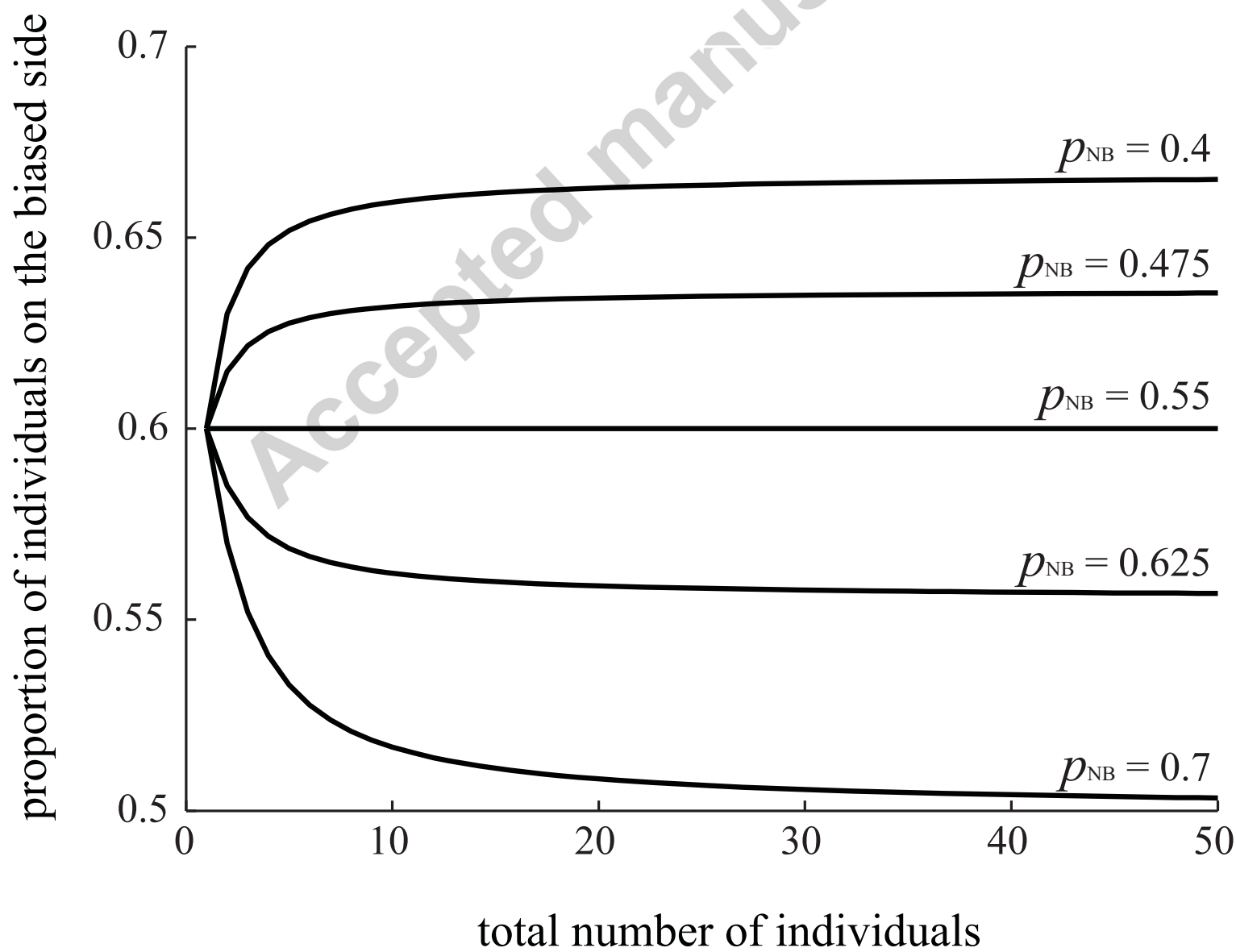
$$f(5,3,4) = 2$$

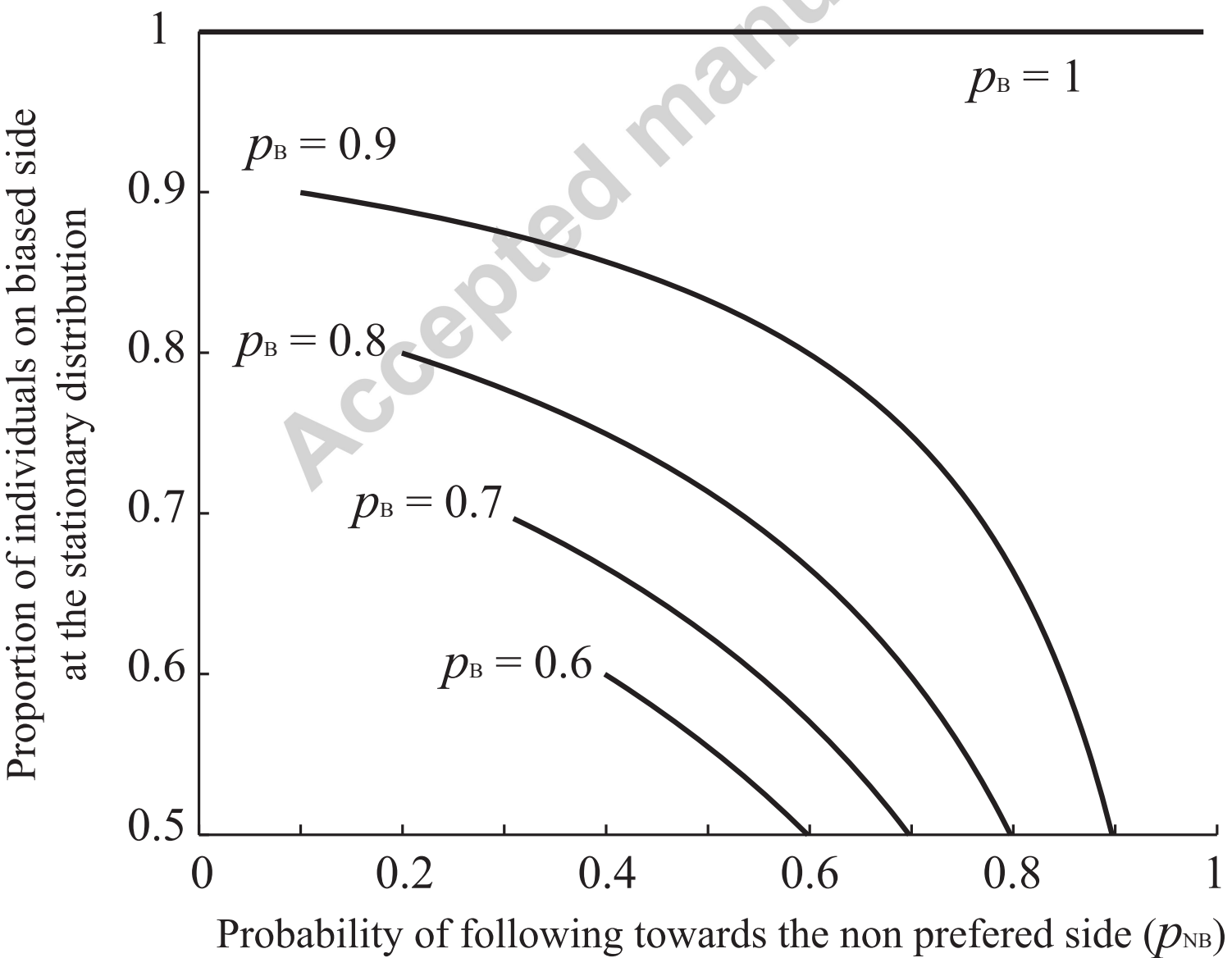




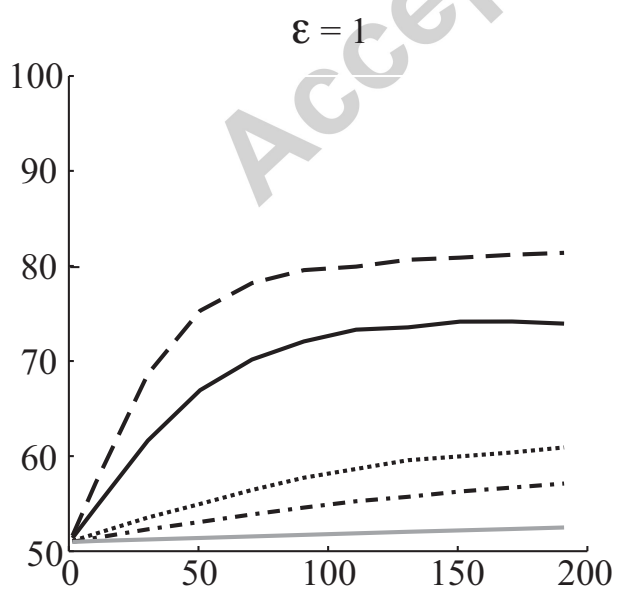




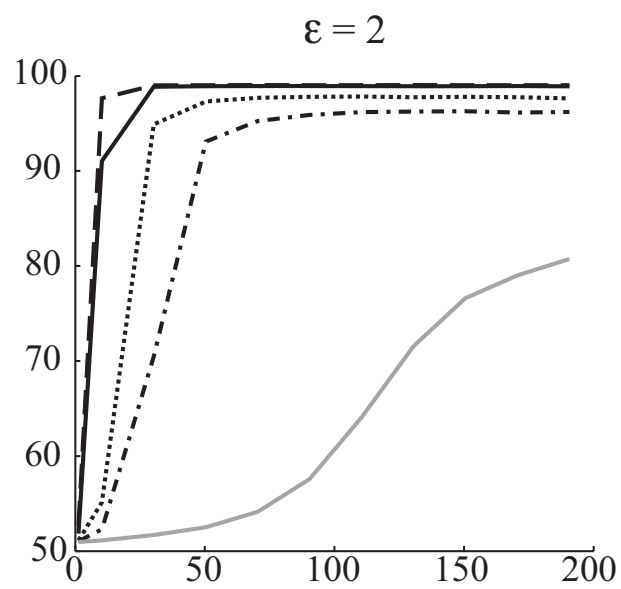




A

expected percentage of individuals  
on the winning side

B



---  $\kappa = 0.5$   
—  $\kappa = 1$   
.....  $\kappa = 5$   
-.-.-  $\kappa = 10$   
—  $\kappa = 50$

number of perceived individuals  $N$



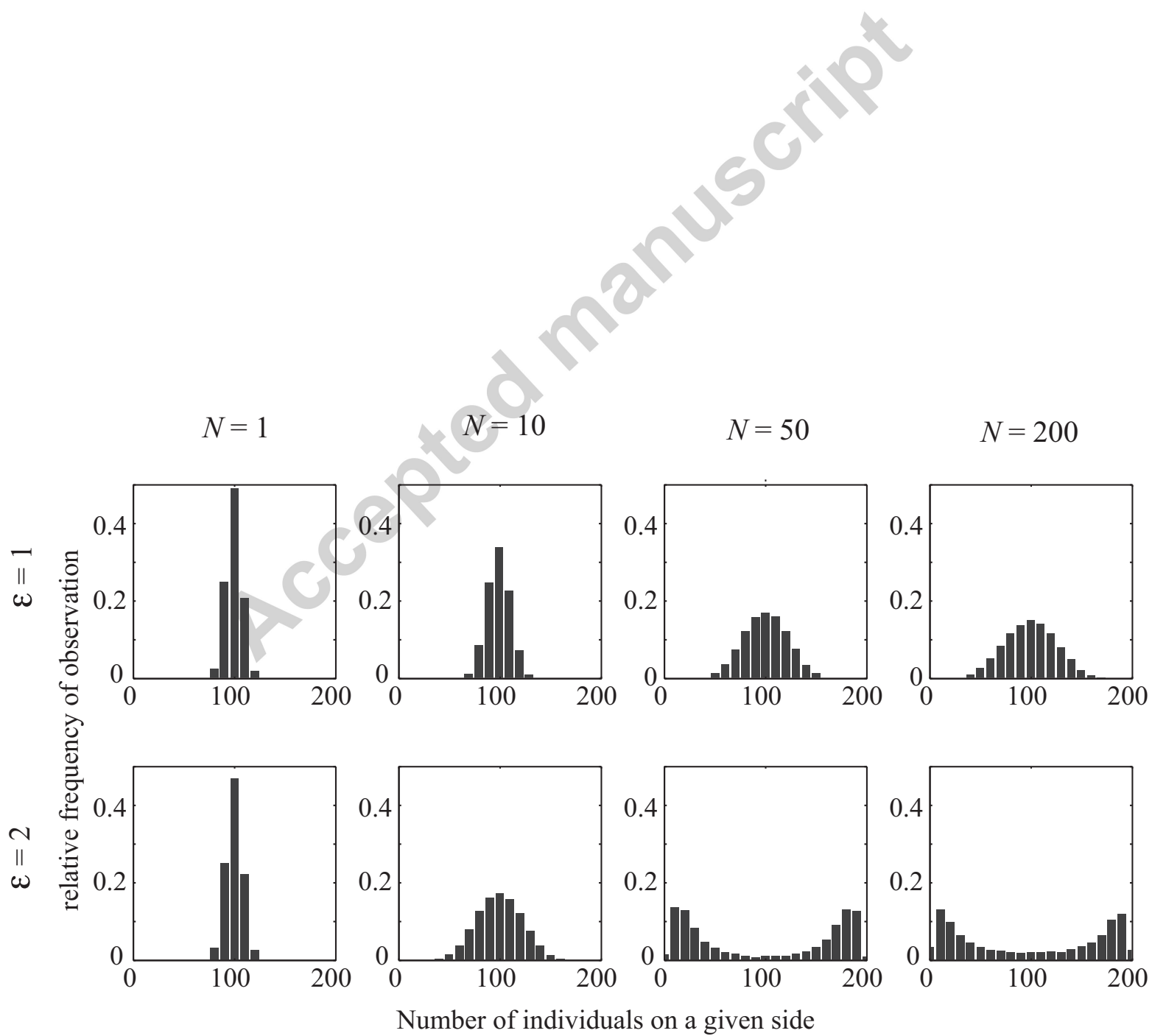
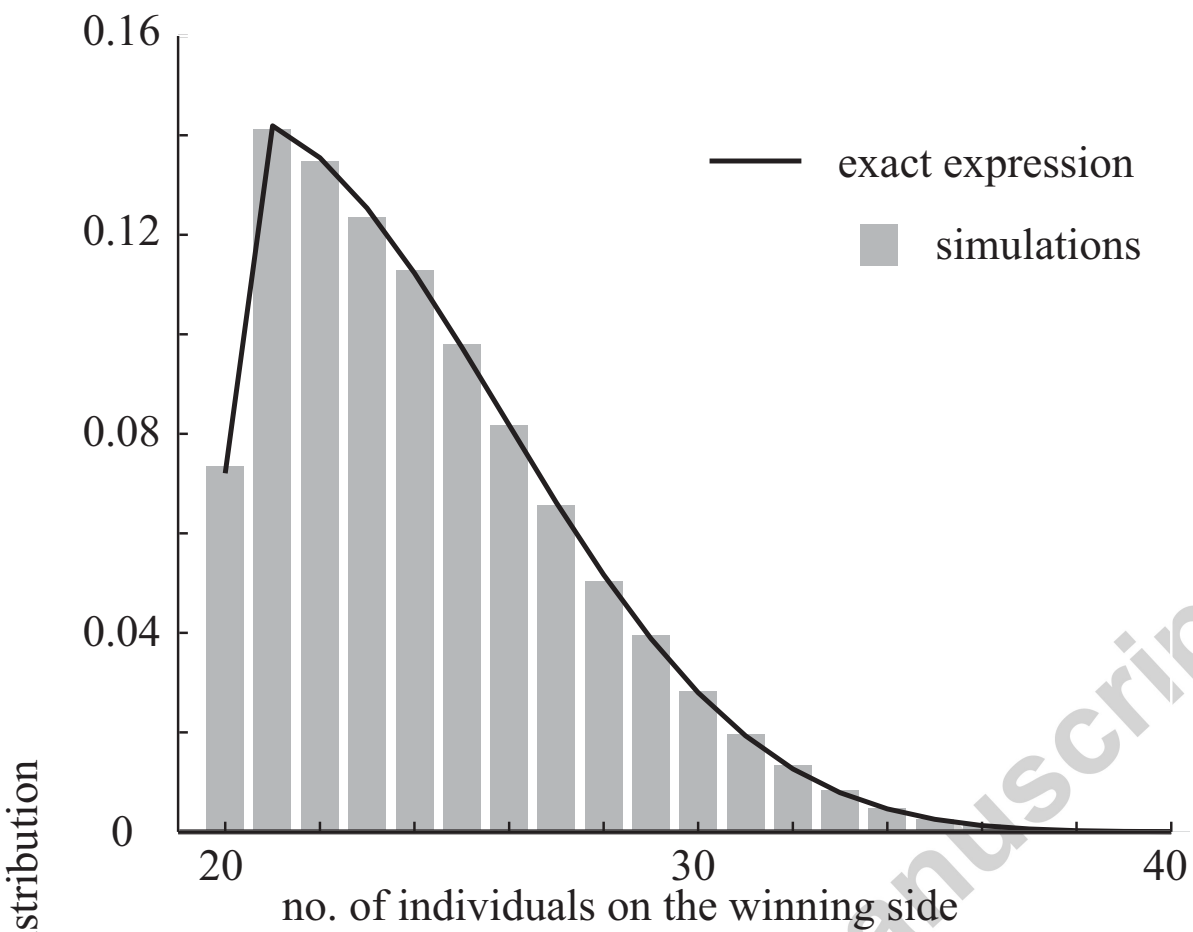
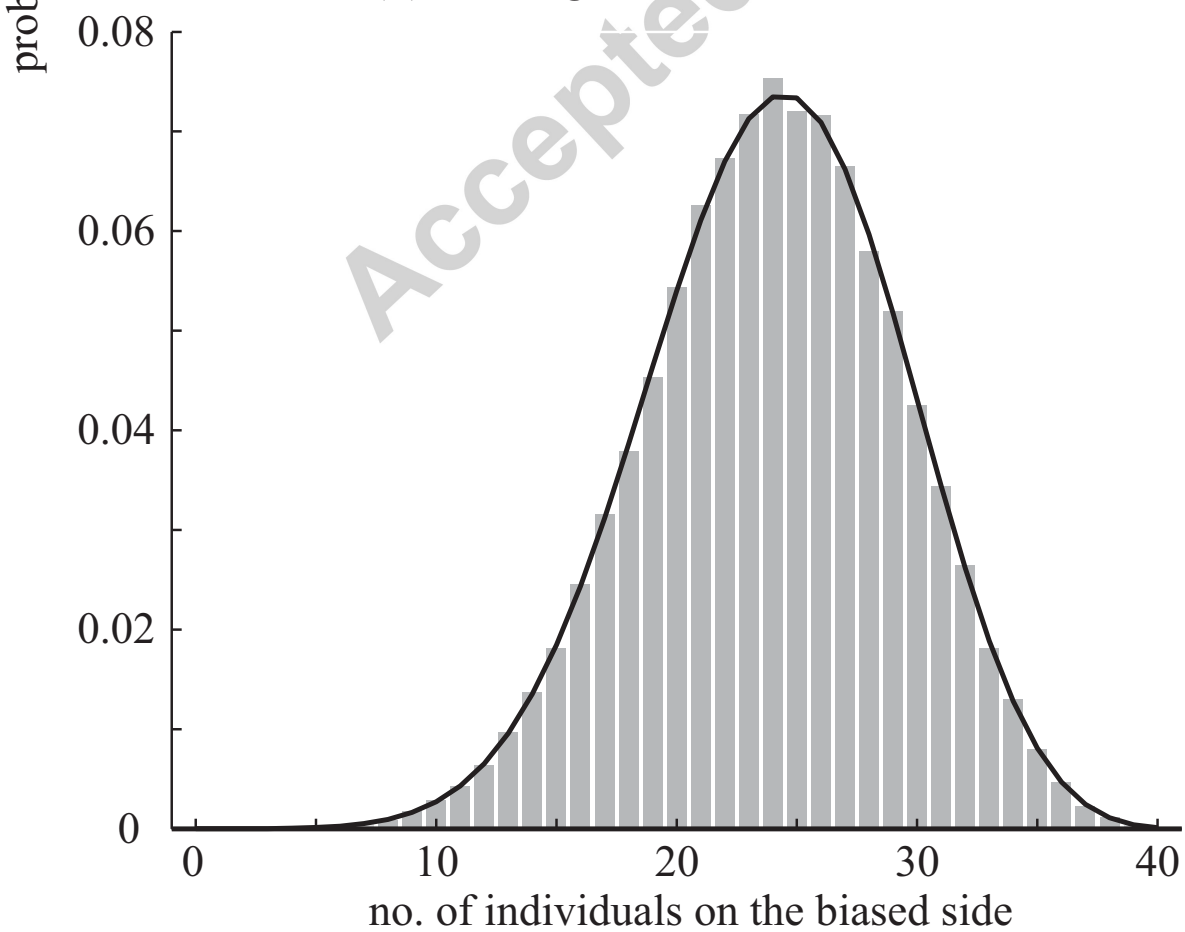


Figure 10

(a) homogeneous environment



(b) heterogeneous environment



## Highlights

- The potential for collective movement is evaluated in groups of individuals that tend to follow their direct predecessor.
- This behaviour is quantified with a two-state Markov chain model.
- The results show that the cohesion resulting from following the direct predecessor is weak.
- A generalized model shows that the perception of more predecessors is needed to ensure cohesion.

Accepted manuscript