



## Coaction versus reciprocity in continuous-time models of cooperation

G. Sander van Doorn<sup>a,b,\*</sup>, Thomas Riebli<sup>a</sup>, Michael Taborsky<sup>a,1</sup><sup>a</sup> Institute of Ecology and Evolution, University of Bern, Wohlenstrasse 50a, CH-3032 Hinterkappelen, Switzerland<sup>b</sup> Centre for Ecological and Evolutionary Studies, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands

## H I G H L I G H T S

- Coordination and information exchange are prominent in animal social interactions.
- We study how these features affect the evolution of cooperation.
- The ability of players to respond to each other in real time supports cooperation.
- Delays in information exchange (inherent to reciprocal altruism) favour selfishness.
- Cooperative coaction therefore evolves more readily than reciprocal cooperation.

## A R T I C L E I N F O

## Article history:

Received 2 December 2013

Received in revised form

3 March 2014

Accepted 11 March 2014

Available online 12 April 2014

## Keywords:

Cooperative coordination

Reciprocal altruism

Information transfer

Social dilemmas

Evolutionary game theory

## A B S T R A C T

Cooperating animals frequently show closely coordinated behaviours organized by a continuous flow of information between interacting partners. Such real-time coaction is not captured by the iterated prisoner's dilemma and other discrete-time reciprocal cooperation games, which inherently feature a delay in information exchange. Here, we study the evolution of cooperation when individuals can dynamically respond to each other's actions. We develop continuous-time analogues of iterated-game models and describe their dynamics in terms of two variables, the propensity of individuals to initiate cooperation (altruism) and their tendency to mirror their partner's actions (coordination). These components of cooperation stabilize at an evolutionary equilibrium or show oscillations, depending on the chosen payoff parameters. Unlike reciprocal altruism, cooperation by coaction does not require that those willing to initiate cooperation pay in advance for uncertain future benefits. Correspondingly, we show that introducing a delay to information transfer between players is equivalent to increasing the cost of cooperation. Cooperative coaction can therefore evolve much more easily than reciprocal cooperation. When delays entirely prevent coordination, we recover results from the discrete-time alternating prisoner's dilemma, indicating that coaction and reciprocity are connected by a continuum of opportunities for real-time information exchange.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

The iterated prisoner's dilemma (IPD) (Axelrod and Hamilton, 1981) is the most frequently used game-theoretical paradigm to study the evolution of cooperation among unrelated individuals. It is fundamental to theories of reciprocal altruism based on direct (Trivers, 1971; Axelrod and Hamilton, 1981), indirect (Nowak and Sigmund, 1998) and generalized reciprocity (Pfeiffer et al., 2005).

Yet, among biologists interested in the evolution of animal social behaviour, dissatisfaction with the IPD as a theoretical framework for studying cooperation has grown (Clutton-Brock, 2009), and several reviews have questioned its empirical relevance (Connor, 1995; Hammerstein, 2003; Noë, 2006; Raihani and Bshary, 2011).

One aspect of the IPD that has attracted much criticism is that players are assumed to decide on their actions independently of each other, without having access to information about the choice being made in the same round by the other player (Noë, 2006; Clutton-Brock, 2009). Though natural in the context of the discrete time structure of the IPD, this assumption is problematic for several reasons. First, the lack of information exchange between players can be exploited, posing a danger to the maintenance of cooperation. Second, establishing reciprocal exchange in the IPD is difficult (Stephens et al., 2002), because it requires players to pay in advance for an uncertain future benefit. In fact, all animals tested so far (including humans)

\* Corresponding author at: Centre for Ecological and Evolutionary Studies, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands. Tel.: +31 50 3638097; fax: +31 50 3635205.

E-mail addresses: [g.s.van.doorn@rug.nl](mailto:g.s.van.doorn@rug.nl) (G.S. van Doorn), [thomas.riebli@iee.unibe.ch](mailto:thomas.riebli@iee.unibe.ch) (T. Riebli), [michael.taborsky@iee.unibe.ch](mailto:michael.taborsky@iee.unibe.ch) (M. Taborsky).

<sup>1</sup> Tel.: +41 31 631 9156.

show temporal discounting, which means that they devalue future relative to immediate rewards (Chung and Herrnstein, 1967; Kalenscher and Pennartz, 2008). Moreover, when behavioural decisions are separated in time from the feedback about the consequences of those decisions, the effectiveness of simple associative learning mechanisms will be undermined, forcing players to rely on some sort of memory of past interactions to infer the optimal choice in the current round (Brosnan et al., 2010). Only species with highly advanced cognitive abilities are considered capable of solving decision problems of this kind (Stevens et al., 2005). Still, cooperation is observed in some organisms clearly lacking such abilities (e.g., Milinski et al., 1990; Rutte and Taborsky, 2008; Krams et al., 2008), suggesting that cooperative interactions between animals differ in important aspects from the IPD.

Tellingly, one of the main challenges in designing empirical tests of the IPD is to restrict the flow of information between players such that they make their decisions independently of each other (Noë, 2006). Successful tests have been performed in humans (Milinski and Wedekind, 1998; Gintis et al., 2003) and other animal species (Clements and Stephens 1995; Hauser et al., 2003; Rutte and Taborsky, 2008), but not without forcing subjects to interact under highly artificial conditions. In contrast, natural cooperative behaviours, such as predator inspection (Milinski et al., 1990; Pitcher, 1992), cooperative hunting (Boesch and Boesch, 1989) or joint territory defence (Krams et al., 2008), typically rely on social information exchange. Individuals may observe the current actions of their partners and respond immediately to changes in their behaviour, or they may actively communicate with each other while they establish or maintain cooperation. As a result, cooperation between animals generally involves coercion or more advanced forms of behavioural coordination (Boesch and Boesch, 1989; Schuster, 2002). ‘Acting together’ has in fact been proposed as a simple operational definition of cooperation (Taborsky, 2007) and it has been emphasized that the ‘achievement [of cooperation] requires collective action’ (Dugatkin, 1997, p. 14). In humans, it has been shown experimentally that synchronous action can foster cooperation, partly because it may help to mitigate the free-rider problem (Wiltermuth and Heath, 2009).

The implications of social information exchange and behavioural coordination for the evolution of cooperation are not well understood, because in biology few theoretical models have strayed from the elementary game structure of the IPD. In the present paper, we therefore analyse a continuous-time model of cooperation that allows individuals to respond in real time to the behaviour of their partner. In this model, pairs of players are able to establish cooperation by coordinating current behaviour, as opposed to reciprocating favours that are separated in time. Previous work on dynamic games in economics suggests that the time-structure of interactions is of critical importance for the establishment of cooperation between human players in a public-goods game (e.g., Marx and Matthews, 2000; Duffy et al., 2007). The reason is that information exchange allows for smaller history-contingent contributions to the public good, enabling each of the players to try the other’s good faith for a small price (Schelling, 1960). We show here that this effect also has implications for the evolution of cooperation. In fact, our analysis indicates that restricting the flow of information between players is equivalent to increasing the cost of cooperation. The exchange of social information, which accompanies natural cooperative interactions between animals, therefore creates conditions that are much more favourable for the evolution of cooperation than one would predict from theory based on the IPD.

## 2. The model

The distinguishing feature of our model is that individuals can switch between actions in continuous time rather than in discrete

**Table 1**  
Payoff parameters and model variants.

Payoff to focal	Partner plays	D	C
Focal plays		$P=0$	$T=b$
		$S=-c$	$R=b-c+h$
Stag-hunt game <sup>a</sup>	$b-c+h > b > 0 > -c \Rightarrow$	$k_D > 0$	$k_C < 0$
Snowdrift game <sup>b</sup>	$b > b-c+h > -c > 0 \Rightarrow$	$k_C - 1 < k_D < 0$	$0 < k_C < 1$
Prisoner’s dilemma <sup>c</sup>	$b > b-c+h > 0 > -c \Rightarrow$	$k_D > 0$	$0 < k_C < 1$

<sup>a</sup> In the stag-hunt game, players prefer mutual defection over unilateral cooperation, disfavoured the initial establishment of cooperation. However, once cooperation has been established between a pair of players, neither has an incentive to cheat.

<sup>b</sup> In the snowdrift game (also known as the hawk-dove game or the game of chicken), unilateral cooperation is preferred over mutual defection, but players achieve the highest payoff if they defect when their partner cooperates.

<sup>c</sup> The prisoner’s dilemma combines the social dilemmas of the snowdrift and stag-hunt game. Unilateral defection is preferred over mutual cooperation, and mutual defection is preferred over unilateral cooperation, such that  $D$  is the dominant strategy in a one-shot game.

<sup>d</sup> The dimensionless parameter combinations  $k_D=c/b$  and  $k_C=(c-h)/b$  measure the relative cost of playing  $C$  when the partner defects or cooperates, respectively. These parameters appear in Figs. 2, 3, 5 and S2.

rounds. Other than that, we closely follow the assumptions of standard iterated cooperation games (Macy and Flache, 2002). The results presented here focus almost exclusively on the prisoner’s dilemma game, but our analysis extends to the strategically different situations embodied by the snowdrift and the stag-hunt game (Table 1). Each player interacts with the same partner over an extended period of time, allowing for repeated interactions. During this time, the momentary rate of increase of a player’s payoff is dependent on its own action and that of its partner. Players can choose between two actions, labelled ‘cooperate’ ( $C$ ) and ‘defect’ ( $D$ ), such that a pair of two players  $i$  and  $j$  can be in one of four discrete states at any moment in time. The state of the pair will be denoted by a combination of two letters,  $DD$ ,  $DC$ ,  $CD$  or  $CC$ , indicating the action currently played by the focal individual  $i$ , followed by the action currently played by its partner, individual  $j$ .

### 2.1. Pair-state dynamics

Each player’s strategy is specified by four parameters that determine the rate at which the individual switches between cooperation and defection, depending on the current state of the pair. The four switching rates are assumed to be bounded away from zero by a small constant  $0 < \varepsilon \ll 1$  (Selten, 1975). We use  $\sigma_i=(p_i, q_i, r_i, s_i)$  and  $\sigma_j=(p_j, q_j, r_j, s_j)$  to denote the strategies of individual  $i$  and  $j$ , respectively, and associate the strategic parameters with transitions between the states of the pair in the following way:

$$\begin{aligned} i \text{ changes state : } & DD \xrightleftharpoons[q_i]{p_i} CD \quad DC \xrightleftharpoons[s_i]{r_i} CC \\ j \text{ changes state : } & DD \xrightleftharpoons[q_j]{p_j} DC \quad CD \xrightleftharpoons[s_j]{r_j} CC \end{aligned} \quad (1)$$

From here it is straightforward to derive ordinary differential equations for the probability distribution of pairs over the different states. For example,  $f_{CD}(t)$ , the probability of finding a pair in state  $CD$  at time  $t$ , changes through time according to the equation  $(d/dt) f_{CD}(t) = p_i f_{DD}(t) + s_j f_{CC}(t) - (q_i + r_j) f_{CD}(t)$ . Similar equations for the other pair states give rise to a system of linear ordinary

differential equations that can be solved explicitly for its equilibrium (Eq. (4)). When the switching rates are bounded away from zero, the equilibrium is unique and stable (§S1.1; throughout, labels starting with ‘S’ will refer to items in the Supplementary online material). In addition, the stochastic process of pair-state switching is then ergodic, meaning that the equilibrium describes both the long-term time-averaged distribution over pair states of an individual pair, and the ensemble average frequency of pair states once equilibrium has been reached.

## 2.2. Payoffs from cooperation

Players receive  $b$  fitness units per unit of time when they defect while their partner cooperates; they lose  $c$  fitness units per unit of time when they cooperate while their partner defects, and gain  $b - c + h$  fitness units per unit of time when they cooperate simultaneously with their partner. No payoff is accumulated when both partners defect. Throughout, we assume that  $b - c + h > 0$ , such that bilateral cooperation yields a higher payoff than bilateral defection. In terms of the usual payoff parameters  $P$ ,  $R$ ,  $S$  and  $T$  of the prisoner's dilemma game (Axelrod and Hamilton, 1981),  $b = R - P$  represents the benefit of an act of altruism to the receiver,  $c = P - S$  measures its cost to the actor, and  $h = P + R - S - T$  can be interpreted as the benefit of coordinating actions with the partner. To ensure that  $T > R > P > S$  (the payoff relationships characteristic of a prisoner's dilemma), it is necessary that  $c > 0$  and  $h < c$ . The parameter regimes corresponding to the snowdrift ( $T > R > S > P$ ) and stag-hunt game ( $R > T > P > S$ ) are listed in Table 1.

Assuming that individuals  $i$  and  $j$  interact for an extended period of time  $\tau \gg \epsilon^{-1}$ , we calculate the payoff per time unit to individual  $i$  as

$$\pi(\sigma_i, \sigma_j) = f_{DC} b - f_{CD} c + f_{CC} (b - c + h), \quad (2)$$

where  $f_k$  is the long-term time-averaged frequency of pairs in state  $k$  ( $k = DD, DC, CD$  or  $CC$ ). The payoff function  $\pi$  depends on the focal's strategy and that of its partner, because both individuals affect the distribution of pair states by their actions.

## 2.3. Lifetime reproductive success

The lifetime reproductive success of an individual is assumed to increase linearly with the payoff from cooperation. Accordingly, the average fitness of an individual with strategy  $\sigma_i$  is written as  $w(\sigma_i) = S(\sigma_i)(1 + E[\pi(\sigma_i, \sigma_j)])$ , where  $S(\sigma_i)$  denotes the survival probability of the focal individual. Assuming that individuals are paired at random, the expectation  $E[\ ]$  in the second term of the fitness function is calculated over the distribution of  $\sigma_j$  in the population. We ensure that fitness is positive by taking the payoff parameters  $b$ ,  $c$  and  $h$  to be small relative to the baseline reproductive success, which is scaled to one without loss of generality. The survival probability of an individual is allowed to depend on its strategy in order to enable the incorporation of switching costs in the model. Players may in principle switch between actions at an arbitrarily high rate, but we assume that speeding up information processing and the execution of behaviour trades off negatively with survival. As a result, viability selection is stabilizing, with the strategy  $\sigma_0 = (0, 0, 0, 0)$  maximizing survival (scaled such that  $S(\sigma_0) = 1$ ).

If evolution is mutation-limited and, hence, the population variance in  $\sigma$  is small, the effects of viability selection and the cooperative interactions become additive. In order to show this, we first approximate the expected payoff from cooperative interactions by the first two terms of its Taylor series:  $E[\pi(\sigma_i, \sigma_j)] \approx \pi(\sigma_i, \bar{\sigma}) + E[\sigma_j - \bar{\sigma}](\partial/\partial\sigma)\pi(\sigma_i, \bar{\sigma})$ , where  $\bar{\sigma} = (\bar{p}, \bar{q}, \bar{r}, \bar{s})$  is the mean strategy in the population. This approximation has a truncation error of  $O(\|\Sigma\|^2)$ , where  $\|\Sigma\|$  is the norm of the phenotypic

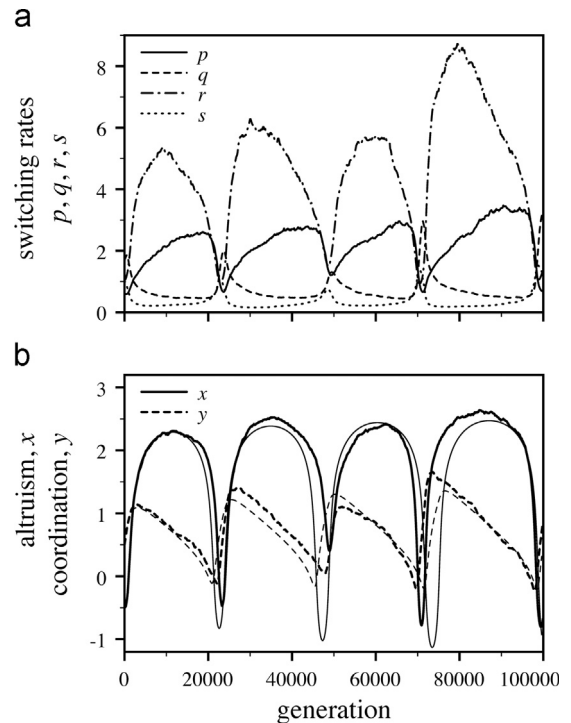
variance-covariance matrix of the strategic variables (a quantity that is guaranteed to be small under mutation-limited evolution). Next, after realizing that  $E[\sigma_j - \bar{\sigma}] = 0$  and applying another Taylor expansion around  $\sigma_i = \bar{\sigma}$ , the relative fitness of strategy  $\sigma_i$  is approximated as follows:

$$\begin{aligned} \frac{w(\sigma_i, \bar{\sigma})}{w(\bar{\sigma}, \bar{\sigma})} &\approx \frac{(S(\bar{\sigma}) + (\sigma_i - \bar{\sigma})(\partial/\partial\sigma)S(\bar{\sigma}))_{|\sigma=\bar{\sigma}}(1 + \pi(\bar{\sigma}, \bar{\sigma}) + (\sigma_i - \bar{\sigma})(\partial/\partial\sigma)\pi(\bar{\sigma}, \bar{\sigma}))_{|\sigma=\bar{\sigma}}}{S(\bar{\sigma})(1 + \pi(\bar{\sigma}, \bar{\sigma}))} \\ &\approx 1 + (\sigma_i - \bar{\sigma})(\partial/\partial\sigma)(\ln S(\bar{\sigma}) + \ln(1 + \pi(\bar{\sigma}, \bar{\sigma})))_{|\sigma=\bar{\sigma}}, \end{aligned} \quad (3)$$

which also has a truncation error of  $O(\|\Sigma\|^2)$ .

## 3. Results

The model was analysed by a combination of mathematical techniques and individual-based simulations. The latter incorporate the effects of genetic drift and do not rely on the approximations (weak selection, mutation-limited evolution) needed in the formal analysis. Fig. 1a shows the outcome of an individual-based simulation run. In this example, the switching rates  $p$ ,  $q$ ,  $r$  and  $s$  display a regular pattern of oscillations over evolutionary time. The population goes through three phases, comparable to what has been observed in discrete-time models of the iterated prisoner's dilemma (Nowak and Sigmund, 1989; Imhof et al., 2005). First, populations with low levels of cooperation are invaded by conditional cooperators, who are willing to coordinate actions and maintain mutual cooperation with a cooperating partner. The establishment of conditional cooperation paves the way for the invasion of more generous cooperative strategies, which subsequently invite the



**Fig. 1.** Evolution of the rates of switching between cooperation and defection in an individual-based simulation. Cooperative interactions between pairs of players were simulated in a population of 100,000 individuals. Each individual carried 4 sets of diploid unlinked autosomal loci, coding for the switching rates  $p$ ,  $q$ ,  $r$  and  $s$ . Mutation occasionally introduced novel alleles at these loci. (a) For  $b = 0.5$ ,  $c = 0.2$ ,  $h = 0.15$  and weak stabilizing viability selection ( $\gamma_1 = 1 \cdot 10^{-4}$  and  $\gamma_2 = \gamma_1 \times 2 = 1 \cdot 10^{-3}$ ; see §S1.2) the evolution of the switching rates exhibits regular oscillations with a period of approximately 25,000 generations. (b) Thick black lines show the dynamics of altruism ( $x = \ln[(pr)/(qs)]/2$ ) and coordination ( $y = \ln[(qr)/(ps)]/2$ ) in the simulation shown in (a). Corresponding predictions based on the analytical model (Eqs. (4)–(6)) are shown as thin grey lines. Here, we fitted only the time-scale parameter  $\kappa$  (Eq. (6)) in order to match the period of the oscillations (§S2).

evolution of cheating behaviour (manifested by a slow erosion of  $r$ , followed by a rapid increase of  $s$ ). Eventually, cooperation breaks down, and the population evolves back to a state characterized by high levels of mutual defection.

Supporting the above interpretation, our mathematical analysis traces the cause of the evolutionary cycles of the switching rates  $p$ ,  $q$ ,  $r$  and  $s$  to an interaction between two factors, the propensity of individuals to cooperate unconditionally and their tendency to establish and maintain bilateral cooperation. In fact, a substitution of variables confirms that the essential features of the evolutionary dynamic of the four switching rates are captured by a dynamical model with only two variables (Fig. 1b; §S1.3). The new coordinates,  $x$  and  $y$ , are related to the original switching rates by  $x = (1/2) \ln((\bar{p}\bar{r})/(\bar{q}\bar{s}))$  and  $y = (1/2) \ln((\bar{q}\bar{r})/(\bar{p}\bar{s}))$ , which can be interpreted as measures of the tendencies to cooperate unconditionally ('altruism') and coordinate actions with the partner ('coordination'), respectively. The equilibrium frequencies of pair states are represented in terms of these new variables as

$$\begin{aligned} f_{DD} &= \frac{e^{-x}}{e^{-x} + 2e^{-y} + e^x}, \\ f_{DC} &= f_{CD} = \frac{e^{-y}}{e^{-x} + 2e^{-y} + e^x} \text{ and} \\ f_{CC} &= \frac{e^x}{e^{-x} + 2e^{-y} + e^x}. \end{aligned} \quad (4)$$

### 3.1. Analytical results

The gradient of the fitness function (3) with respect to the new strategic variables  $x$  and  $y$  consists of terms that derive from three different sources of selection (§S1.2–3)

$$\begin{aligned} w_x &= \left. \frac{\partial \ln(w(\sigma, \bar{\sigma}))}{\partial \sigma} \right|_{\sigma = \bar{\sigma}} \frac{\partial \bar{\sigma}}{\partial x} = \underbrace{\frac{\pi_{CC} - \bar{\pi}}{2} f_{CC} - \frac{\pi_{DD} - \bar{\pi}}{2} f_{DD}}_{\text{maximization of mean fitness}} - \underbrace{\frac{b+c}{2} f_{CD}}_{\text{incentives for unilateral defection}} - \underbrace{\gamma_1 x}_{\text{stabilizing viability selection}}, \\ w_y &= \left. \frac{\partial \ln(w(\sigma, \bar{\sigma}))}{\partial \sigma} \right|_{\sigma = \bar{\sigma}} \frac{\partial \bar{\sigma}}{\partial y} = \underbrace{\frac{\pi_{CC} - \bar{\pi}}{2} f_{CC} + \frac{\pi_{DD} - \bar{\pi}}{2} f_{DD}}_{\text{maximization of mean fitness}} - \underbrace{\frac{b+c}{2} f_{CD} \tanh\left(\frac{x}{2}\right)}_{\text{incentives for unilateral defection}} - \underbrace{\gamma_{1 \times 2} y}_{\text{stabilizing viability selection}}. \end{aligned} \quad (5)$$

The first component acts to maximize the mean fitness of pairs in response to the differences between the mean payoff  $\bar{\pi} = f_{DC} b - f_{CD} c + f_{CC}(b - c + h)$  and the payoffs in the coordinated pair states,  $\pi_{DD} = 0$  and  $\pi_{CC} = b - c + h$ . The mean fitness of pairs is maximized in state  $CC$  (i.e., at large positive values of  $x$ ), unless the payoffs from unilateral cooperation and defection on average exceed  $\pi_{CC}$  (this requires  $h < -(b - c)/2$ , in which case selection favours anti-coordination, i.e., large negative values of  $y$ ). A second source of selection captures the selfish interests of players to exploit their partner. The difference  $b - (-c) = b + c$  between the payoffs for a player in states  $DC$  and state  $CD$  creates an incentive for individuals to (1) become less altruistic (i.e., selfish interests invariably favour lower values of  $x$ ), (2) coordinate actions with a partner who defects, and (3) anti-coordinate with a cooperating partner. Accordingly, the direction of selection on  $y$  is determined by the sign of  $x$  (this feature is captured by the factor  $\tanh(x/2)$ ). The final terms on the right-hand side of Eq. (5) incorporate the stabilizing effect of viability selection, which pushes  $x$  and  $y$  in the direction of  $x = 0$  and  $y = 0$ , respectively.

The particularly simple form of the stabilizing selection terms in Eq. (5) is obtained for a selection function  $S$  that is a product of Gaussian stabilizing selection functions (§S1.2). The underlying assumption motivating this choice is that the fitness cost for rapid switching can be decomposed into three independent components

related to response to changes in one's own state, the state of the partner and the coordination state of the pair. The marginal fitness costs associated with each of these components are measured by the parameters  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_{1 \times 2}$ . One would typically expect  $\gamma_1 \leq \gamma_2 \leq \gamma_{1 \times 2}$ , since responding to changes in the behaviour of the partner is likely to require more complex information processing mechanisms than responding to changes in one's own actions, yet both of these might be easier to achieve than rapidly coordinating with the partner. We assume throughout that the marginal cost parameters are small such that stabilizing selection is weak.

The strategic variables  $x$  and  $y$  change over evolutionary time under the influence of selection in a direction that is positively correlated with the fitness gradient (Hofbauer and Sigmund, 1998). If evolution is mutation-limited, the dynamics of this process can be modelled as

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \kappa \mathbf{G} \begin{pmatrix} w_x \\ w_y \end{pmatrix}, \quad (6)$$

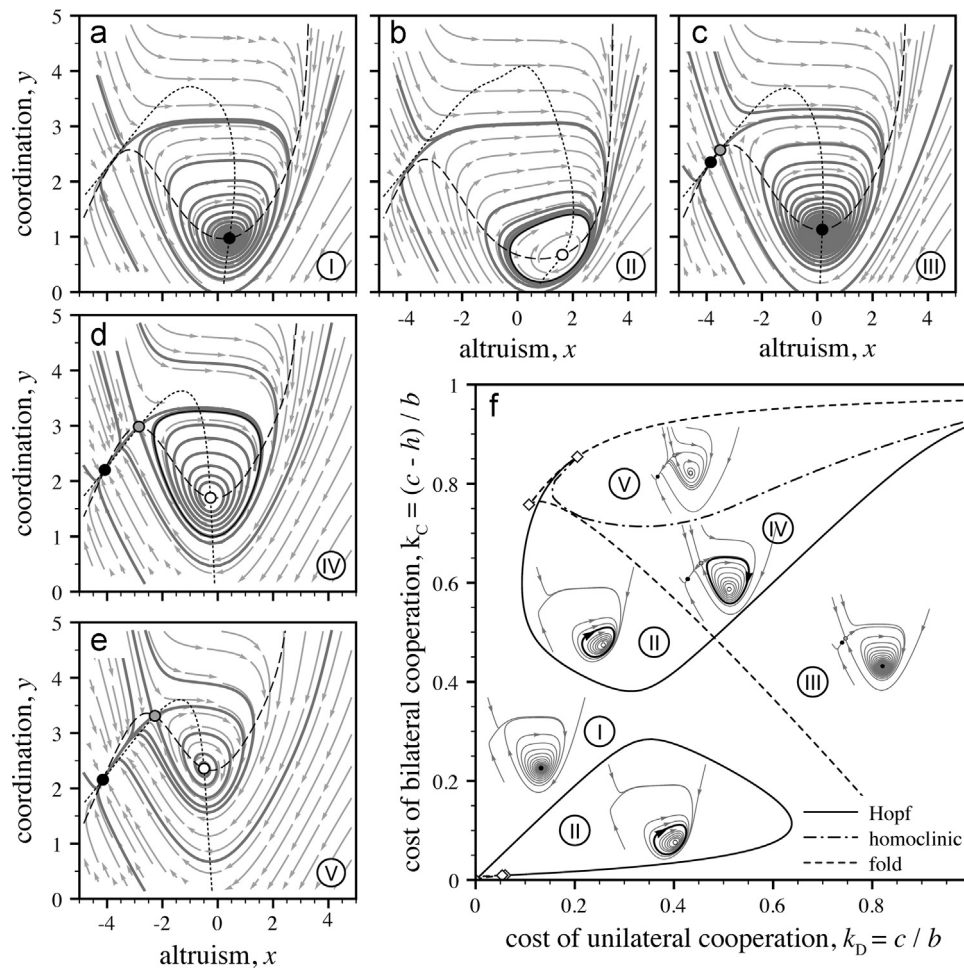
where  $\kappa$  is a constant positive prefactor that depends on the rate of mutations and the population size, and  $\mathbf{G}$  is a  $2 \times 2$  matrix with mutational variances  $V_x$  and  $V_y$  on the diagonal and mutational covariance  $C_{xy}$  off the diagonal (Dieckmann and Law, 1996). We treat the mutational (co)variance matrix as constant, which amounts to assuming that mutations of the original rate parameters  $p$ ,  $q$ ,  $r$ , and  $s$  have multiplicative effects on the phenotype drawn from a fixed distribution. This is the simplest biologically reasonable mutation model for rate parameters, which have a natural lower bound at zero.

Potential endpoints of evolution can be found by solving for combinations of  $x$  and  $y$  at which  $w_x$  and  $w_y$  vanish. If the switching costs are negligible, there is at most one evolutionary

equilibrium with equilibrium values given by  $x^* = \ln(c/(c - h))$  and  $y^* = \ln((b + c)/(b - c + h))$ . The equilibrium exists only if  $c$  and  $c - h$  have the same sign, implying that the payoff relationships must obey the parameter constraints of the prisoner's dilemma (Table 1). In the snowdrift game ( $-b < c < 0$ ), selection invariably leads to negative levels of altruism and coordination, with  $y \approx x + \ln(-(b + c)/c)$  (and  $x \rightarrow -\infty$  in the absence of viability costs). In this case, pairs switch between periods of mutual defection and unilateral cooperation. Each individual cooperates for a fraction  $-c/(b - c)$  of the time, during which time the other individual defects. The pair spends the remaining time in state  $DD$ , and  $f_{DD} = (b + c)/(b - c)$ . In the stag-hunt game, selection favours high values of altruism and coordination, allowing individuals to coordinate effectively on bilateral cooperation ( $f_{CC} \rightarrow 1$ ).

The stability of the equilibrium ( $x^*$ ,  $y^*$ ) is determined by linearizing equation system (6) around the equilibrium point. If mutations have no pleiotropic effects ( $C_{xy} = 0$ ) and the mutational variances are equal ( $V_x = V_y$ ; see §S1.4 and Fig. S1 for more general results) the equilibrium changes its stability at  $h = 0$  (condition 1) or at  $b - 2c + h = (b - c)/2$  (condition 2). Both of these conditions are associated with the origin or disappearance of a limit cycle. Condition 2 is reminiscent of the well-known parameter constraint  $b - c + h > (b - c)/2$  (or,  $R > (S + T)/2$ ) in the discrete-time IPD, which ensures that establishing bilateral cooperation yields a





**Fig. 2.** Five qualitatively different dynamical regimes for the evolution of altruism ( $x$ ) and coordination ( $y$ ). (a)–(e) Each phase diagram shows the vector field with fixed points (black: stable node or spiral; grey: saddle point; white: unstable node or spiral), null-clines (dashed and dotted black lines) and some illustrative evolutionary trajectories in grey. Stable periodic orbits (present in (b) and (d)) are shown as black solid lines. (f) The evolutionary dynamics depend qualitatively on the relative costs of unilateral and bilateral cooperation (measured by  $k_p$  and  $k_c$ , respectively; Table 1). Each of the dynamical regimes (labelled as I–V) occurs in its own region of parameter space which is separated from other regions by lines of Hopf, homoclinic or fold bifurcations (§S1.6 for additional technical details and a characterization of the codim-2 bifurcation points (diamonds)). Parameters are  $b=0.1$ ,  $\gamma_1=2 \cdot 10^{-5}$  and  $\gamma_1 \times 2=4 \cdot 10^{-5}$ ; in panel (a)–(e),  $c=0.055$  and  $h=0.02, 0.045, 0.01, -0.015$  and  $-0.03$ , correspondingly.

higher payoff than alternating between  $cd$  and  $dc$ . Note that, in a continuous-time model, mutual cooperation (yielding payoff  $b - c + h$ ) can only be established after a phase of unilateral cooperation (yielding payoff  $-c$ ). A qualitative change in the evolutionary dynamics occurs when the net fitness effect of this two-step transition ( $b - 2c + h$ ) equals the mean payoff in states  $cd$  and  $dc$ .

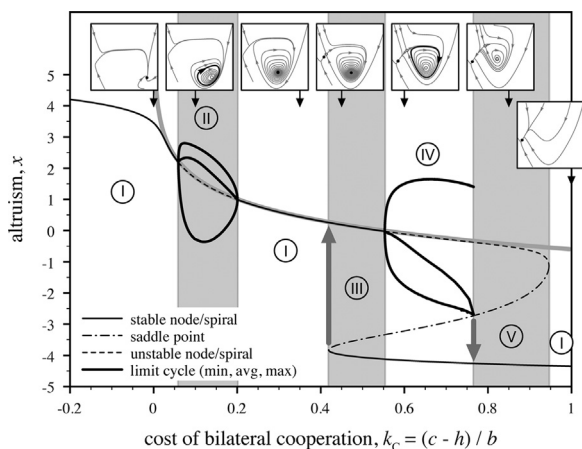
Whenever the equilibrium  $(x^*, y^*)$  is stable with respect to the dynamics described by Eq. (6), it is also stable against invasion by all mutants that differ slightly from the equilibrium strategy, i.e., convergence stability and local evolutionary stability coincide (§S1.5, but see the discussion of Figs. 4 and 5).

### 3.2. Numerical bifurcation analysis

In order to investigate the evolution of altruism ( $x$ ) and coordination ( $y$ ) in more detail, we ran simulations of equation system (6) for different combinations of the payoff and viability selection parameters. Here, we concentrate on the effects of two dimensionless combinations of the payoff parameters (Table 1), the relative cost of unilateral cooperation ( $k_p = c/b$ ), and the relative cost of bilateral cooperation ( $k_c = (c-h)/b$ ). Results for the viability selection parameters are shown in Fig. S2b and c. We observed five qualitatively different outcomes (henceforth regime I–V, represented in Fig. 2a–e), that differ in the number of equilibria and/or the nature of the evolutionary attractors (fixed

point or stable cycle). The five regimes are separated from each other by bifurcation lines (Fig. 2f), which can be traced through parameter space by numerical methods (§S1.5). Each point on a bifurcation line identifies a special combination of parameters at which the dynamical behaviour of the model changes from one regime to another. We observed three different types of such qualitative changes (Fig. 2f): ‘fold’ bifurcations (marking the sudden emergence of a pair of equilibria), ‘Hopf’ bifurcations (associated with the birth of a limit cycle) and ‘homoclinic’ bifurcations (where a limit cycle disappears as it collides with a saddle point). Additional details are provided in Fig. S2 and the accompanying text in the Supplementary material (§S1.6).

Fig. 3 illustrates the realized levels of altruism along a transect through parameter space ( $k_p=0.55$ ;  $0 \leq k_c \leq 1$ ) that passes through each of the five dynamical regimes described in Fig. 2. Small insets show simplified phase-diagrams that characterize the evolutionary dynamics in each of the regions (indicated by alternating background colours). Bifurcations occur on the boundaries between regions. A broad observation is that the realized level of altruism decreases with the cost of bilateral cooperation, as one would expect. However, in some cases, larger costs lead to an increase in altruism, due to the emergence of evolutionary oscillations with a time-average dominated by cooperative strategies (as in Fig. 1). The time-average of the oscillations may also be biased towards non-cooperative strategies. This occurs close to the homoclinic bifurcation (at  $k_c=0.77$  in Fig. 3),



**Fig. 3.** Realized levels of altruism as a function of the cost of cooperation. The various black curves indicate the stability and location of fixed points along a transect through parameter space, as well as the minimum, maximum and time-average of  $x$  in cases where evolution converges on a stable periodic orbit. Small insets show simplified phase diagrams (with axes and scale as in Fig. 2a–e) at specific points along the transect (indicated by arrows). For these parameter conditions ( $b=0.1$ ,  $\gamma_1=2 \cdot 10^{-5}$ ,  $\gamma_1 \times 2=4 \cdot 10^{-5}$  and  $c=0.055$ , corresponding to  $k_0=0.55$ ), the system passes through each of the dynamical regimes identified in Fig. 2 as the costs of initiating bilateral cooperation are varied. The different regimes are indicated by the labels I–V and alternating background colour. Qualitative changes in the number and type of evolutionary attractors, which occur on the boundaries between grey and white areas, can be associated with discontinuous jumps in the realized level of altruism (grey arrows). The thick grey line coinciding with part of the equilibrium curve represents the analytical prediction for the location of the equilibrium in the absence of response costs.

as the evolutionary dynamics come nearly to a stop in the vicinity of the saddle point.

There is always an equilibrium close to the point  $(x^*, y^*)$  predicted by the analytical results (Fig. 3, thick light-grey curve) if the response cost parameters  $\gamma_1$  and  $\gamma_1 \times 2$  are small relative to the payoff parameters. Yet, there may be up to two additional equilibria that do not show up in the analysis for negligible costs. The presence of these equilibria leads to bistability over a wide range of parameters. As a result, hysteresis and discontinuous changes in the frequency of cooperation (indicated by arrows) may occur in response to fluctuations of the payoff parameters. Viability selection also has consequences in the snowdrift and the stag-hunt game, where response costs prevent the strategic variables from diverging to  $\pm \infty$ . In both games, evolution converges on a unique stable equilibrium, either at low (snowdrift game) or high (stag-hunt game; Fig. 3,  $k_c < 0$ ) levels of altruism. The realized distribution of pair states at these equilibria is nevertheless close to what one would predict based on the analysis for negligible costs.

### 3.3. Comparison with discrete-time models of cooperation

In discrete-time models of cooperation such as the IPD, players cannot react immediately to changes in the behaviour of their partner. Instead, they are forced to rely on information of past interactions. The lack of real-time information allows for a discrepancy between inferred and actual behaviour that can be exploited by cheaters. In order to illustrate how this problem unfolds as the exchange of information between players is restricted, we extended our model by introducing an information delay. Specifically, we assumed that it takes some time before an individual becomes aware of changes in the behaviour of its partner. This delay allows for an information asymmetry between players, which is a characteristic feature of discrete-time models of cooperation. Indeed, for delays comparable to the decision time of individuals, we recover results of

discrete-time models of the alternating (or asynchronous) prisoner's dilemma (Frean, 1994; Nowak and Sigmund, 1994).

In the extended model, the state of a pair is represented by a combination of four letters, successively designating the actual state of focal individual ( $D$  or  $C$ ), the perceived state of its partner (denoted by lowercase  $d$  or  $c$ ), the actual state of its partner ( $D$  or  $C$ ) and the state of the focal individual as perceived by its partner (lowercase  $d$  or  $c$ ). Consider, for example, a pair of conditional cooperators who establish mutual cooperation after a period of mutual defection. If the delay in information exchange between the players is small and focal individual  $i$  is the player initiating cooperation, then the pair is likely to pass through the following sequence of states during the transition

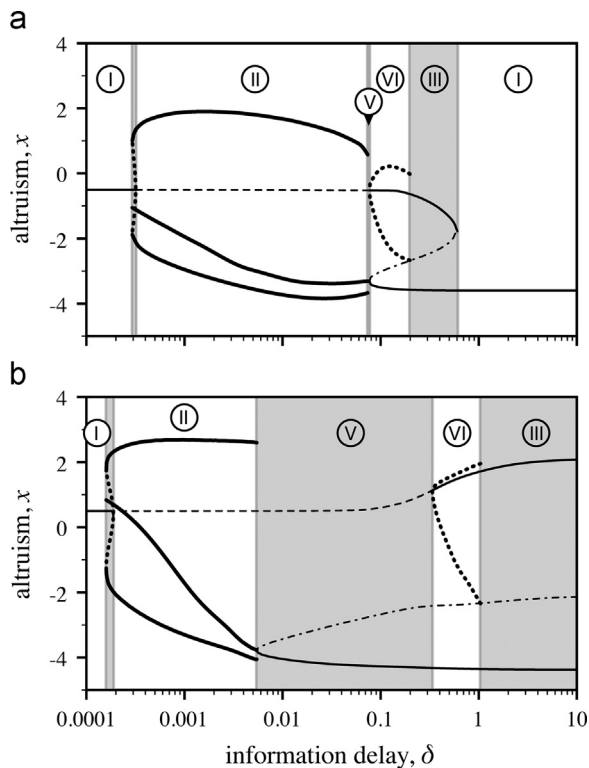
$$DdDd \xrightarrow{p_i} CdDd \xrightarrow{1/\delta} CdDc \xrightarrow{r_j} CdCc \xrightarrow{1/\delta} CcCc$$

here, changes in actual state (in the first and third step of the sequence) are mirrored by changes in perceived state (second and fourth step) at rate  $\delta^{-1}$ , in such a way that it takes, on average,  $\delta$  time units for an individual to become aware of a change in its partner's behaviour. In order to understand how such a delay might interfere with the establishment of cooperation, it is useful to consider alternative events that might occur along the pathway leading towards mutual cooperation. For a pair of conditional cooperators, the most likely alternatives leading away from the original pathway (highlighted in boldface) are

$$\begin{array}{ccccc} DdDd & \xleftarrow{q_i} & CdDd & \xleftarrow{q_j} & CdCd \\ & & \downarrow 1/\delta & & \\ DdDc & \xleftarrow{q_i} & CdDc & \xleftarrow{r_j} & CdCc & \xleftarrow{q_i} & DdCc \\ & & \downarrow 1/\delta & & \\ CcDc & \xleftarrow{r_j} & CcCc & \xleftarrow{s_i} & DcCc \end{array}$$

(with the length of arrows indicating the relative magnitude of forward and reverse transition rates). Conditional cooperators are more likely to cooperate with a cooperating partner than to initiate unilateral cooperation with a partner who defects, implying that the first step of the transition process is rate-limiting. Note also that it takes until the last step in the sequence before individual  $i$  becomes aware of the change in its partner's behaviour. Before that step, individual  $i$  assumes that its partner is defecting, meaning that it has a high probability of switching back to defection itself. In a similar way, individual  $j$  is unlikely to switch to cooperation before it becomes aware of the change in its partner's behaviour in the second step of the sequence. As a result of these asymmetries (i.e.,  $q_i > s_i$  and  $r_j > p_j$ ), delays in the exchange of information make it more difficult for conditional cooperators to coordinate on bilateral cooperation. In fact, a player must be willing to cooperate unilaterally (to the best of its knowledge) for  $2\delta$  time units, before it can reasonably expect to find out whether the help provided is being reciprocated.

The fitness function in the model version with delays depends on the stable frequency distribution of the pair over the actual (not perceived) states of the individuals. We could not solve analytically for this frequency distribution, preventing us from performing a full bifurcation analysis. Yet, we were able to numerically track the location of the equilibrium points, and used numerical integration methods to detect the stable attractors for a range of values of  $\delta$ . Fig. 4 shows results for two different combinations of the payoff parameters  $b$ ,  $c$  and  $h$ . In both cases, small delays ( $\delta < 0.001$ ) destabilize cooperation, resulting in the emergence of evolutionary oscillations. Well before the delay in information transfer becomes comparable to the typical time between changes in an individual's actual state (at  $\delta \approx 1$ ), the evolutionary oscillations



**Fig. 4.** Cooperation breaks down with delays in the transfer of information between interacting partners. Fixed points and periodic orbits were located across a range of values for  $\delta$ , the information delay in the extended model (attractor type and stability are indicated as in Fig. 3; dotted lines show the approximate extreme values of  $x$  along an unstable limit cycle). Restricting the flow of information between players has effects strikingly similar to increasing the cost of cooperation. A qualitative difference with Fig. 3 is that a new stable cooperative equilibrium may emerge at higher values of  $\delta$ . This equilibrium is initially separated from the other stable attractor by an unstable limit cycle. In (a), the novel cooperative equilibrium disappears at longer delays; in (b), it persists. The novel dynamical regime VI (featuring two stable fixed points, a saddle point and an unstable cycle) is connected to regime III by a homoclinic bifurcation. Parameters are  $b=0.05$ ,  $\gamma_1=\gamma_2=1 \cdot 10^{-5}$ ,  $\gamma_1 \times 2=2 \cdot 10^{-5}$ ; in panel (a),  $c=0.01$  and  $h=-0.005$ ; in panel (b),  $c=0.025$  and  $h=0.01$ .

are lost again as a result of the appearance of a new stable fixed point with high rates of defection. Populations that initially track the evolutionary cycle will end up at this attractor and stay there when the information delay is increased further. This outcome is comparable to the effect of increasing the cost of cooperation and caused by a similar sequence of bifurcation events (cf. Fig. 4a and the left part of Fig. 3).

A novel cooperative equilibrium appears in Fig. 4a (at  $\delta=0.076$ ) and Fig. 4b (at  $\delta=0.33$ ), which is initially separated from the other stable equilibrium by an unstable limit cycle. This solution is similar to the winning strategy in tournaments of the alternating prisoner's dilemma ('firm but fair'; Frean, 1994), the discrete-time analogue of our extended model. The novel equilibrium is locally evolutionarily stable (§1.5), i.e., it is resistant against invasion by mutants that are similar to the equilibrium strategy, but not globally stable. Specifically, a mutant playing the non-cooperative equilibrium strategy (which may immigrate from another population, for example), can increase in frequency and replace the cooperative resident strategy. The results shown in Fig. 5 indicate this to be a general outcome.

Fig. 4a is illustrative for other parameter combinations with  $h < 0$  (Fig. 5, above the main diagonal). Here, equilibrium solutions supporting more than a minimal level of cooperation are eventually lost as individuals are increasingly unable to coordinate their actions. Likewise, the bifurcation diagram shown in Fig. 4b is

typical for much of the remaining parameter space. When  $h > 0$  (Fig. 5, below the main diagonal), there is an equilibrium with high levels of cooperation to which evolution will converge from a range of initial conditions (Fig. 5a–d and insets in panel e–g). This equilibrium is locally evolutionarily stable, but it is typically not resistant against invasion by defectors at low information transfer rates. In most of the parameter space, the only equilibrium that is evolutionarily stable as well as a stable endpoint of the evolutionary dynamics of mutation and selection (Eq. (6)) features minimal levels of cooperation (Fig. 5a–g, main panels). In other words, with long delays in the transfer of information between players (Fig. 5e–g) high evolutionarily stable levels of cooperation can only be maintained under a restricted range of parameters.

#### 4. Discussion

The conditions of the iterated prisoner's dilemma are rarely met in animal societies under natural conditions (Connor, 1995; Hammerstein, 2003; Noë, 2006; Clutton-Brock, 2009; Raihani and Bshary, 2011). Instead of a discrete sequence of decisions to cooperate or not, made by two interacting partners with limited possibilities to predict their partner's response, it seems more realistic to think of cooperation as an outcome of social partners adjusting their decisions to a continuous flow of information between them (Pitcher, 1992; McNamara et al., 1999; Mendres and de Waal, 2000; Schuster, 2002). Cooperative hunting (e.g., Boesch and Boesch, 1989) provides a striking example where two or more individuals constantly adjust their behaviour and effort to the actions of other participants in the hunt.

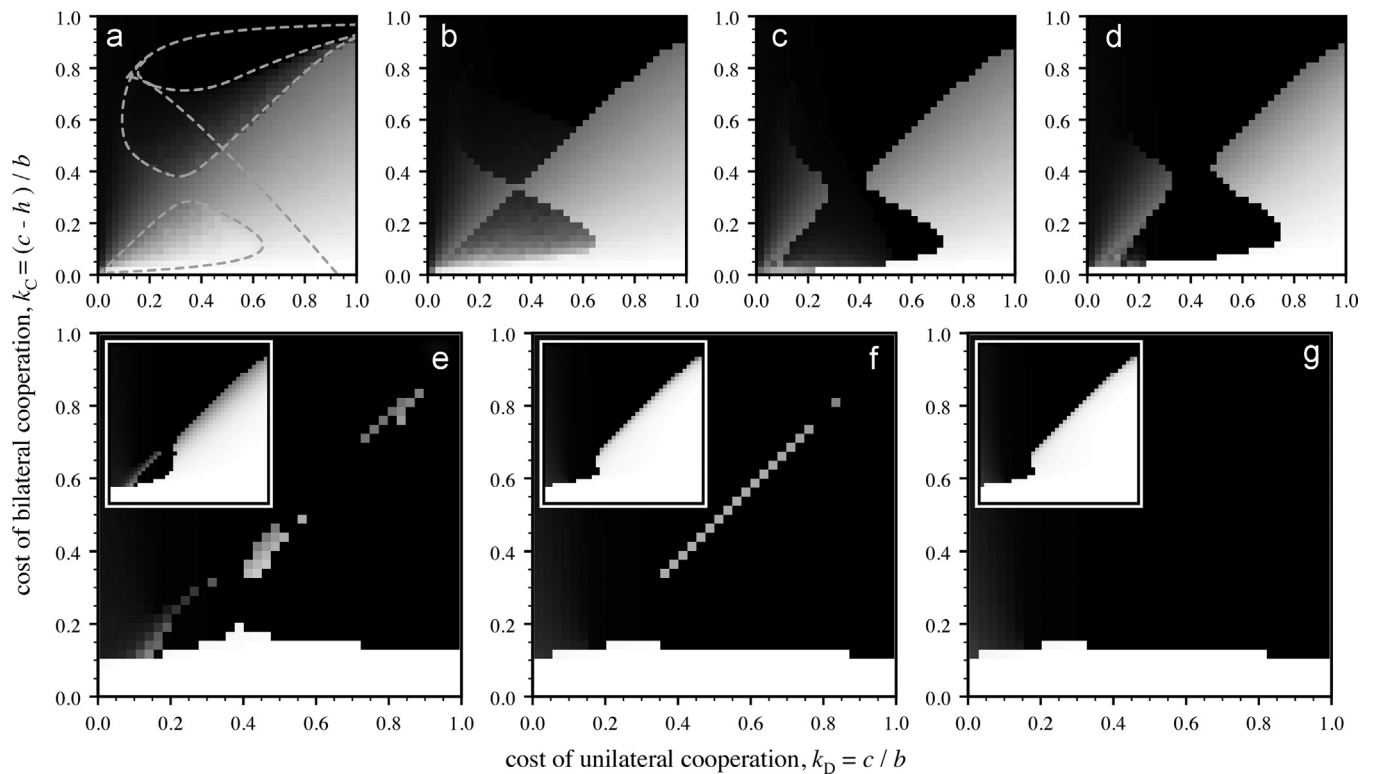
To capture the aspect of close behavioural coordination in natural cooperative interactions, we analysed a continuous-time cooperation game where players could observe each other's actions and respond in real time to changes in the behaviour of their partner. Two variables are sufficient to describe the long-term evolution of the strategies in such a coaction game (Fig. 1). The first variable,  $x$ , reflects the propensity of individuals to cooperate unconditionally and initiate unilateral cooperation ('altruism'); the second variable,  $y$ , measures the tendency to adopt the same behaviour as the partner ('coordination').

High levels of altruism and coordination evolve when players benefit from maintaining bilateral cooperation, whereas establishing cooperation with a defecting partner is costly (the social dilemma embodied by the stag-hunt game). Alternatively, if unilateral cooperation is preferred over mutual defection, but each player has an incentive to cheat when its partner cooperates (the social dilemma embodied by the snow-drift game), the outcome of evolution is that players alternate between periods of bilateral defection and unilateral cooperation.

Intermediate levels of altruism and coordination are observed for payoff parameters characteristic of the prisoner's dilemma. Evolution may then either converge on an evolutionarily stable fixed point, or on a limit cycle. When cooperation is costly, the evolutionary dynamics are bi-stable (Figs. 2c and d and 3), due to the presence of a second attractor at low levels of altruism. This equilibrium remains as the sole attractor if the benefit-to-cost ratio of providing help becomes prohibitively unfavourable (Figs. 2e and 3).

Our results show several parallels to discrete-time models of iterated social dilemma games. First, depending on the payoff relationships, the observed equilibrium strategies exhibit similar properties as successful strategies in the IPD (generous Tit-for-Tat and Pavlov) (Nowak and Sigmund, 1993), or they are similar to known evolutionarily stable solutions for the iterated snowdrift game (Van Doorn et al., 2003). Second, we observe evolutionary oscillations as in discrete-time models of cooperation (Nowak and Sigmund, 1989; Imhof et al., 2005) with the same characteristic





**Fig. 5.** Maximum level of evolutionarily stable cooperation across equilibria. The procedure followed for Fig. 4 was repeated for a range of values of  $c$  and  $h$ . For all stable equilibria detected, we evaluated the realized level of cooperation and evolutionary stability. Panels (a)–(g) show the highest the realized level of cooperation (calculated as the fraction of time an individual was cooperating, ranging from zero (black) to one (white)) for different values of  $\delta$ , across equilibria that were both convergence stable (i.e., stable attractors of the evolutionary dynamics) and evolutionarily stable (resistant against invasion by alternative equilibrium strategies). If the delay in information exchange between partners is not too high (panel (a)–(d);  $\delta=10^{-4}$ ,  $\delta=10^{-3}$ ,  $\delta=0.01$  and  $\delta=0.1$ , correspondingly), the equilibrium with the highest realized level of cooperation is evolutionarily stable. For  $\delta \geq 1$  (panel (e)–(g);  $\delta=1$ ,  $\delta=10$  and  $\delta=100$ , correspondingly), there is still a convergence-stable equilibrium with a high frequency of cooperation (shown in the insets). However, unless the relative cost of bilateral cooperation is small ( $k_C < 0.15$ ), this equilibrium can be invaded by defectors playing the alternative convergence-stable equilibrium strategy. The alternative equilibrium (shown in the main panel) is evolutionarily stable and features minimal levels of cooperation. Parameters are as in Fig. 4.

sequence of events: cautious conditional cooperators first outperform defectors, but are then gradually replaced by more generous cooperative strategies, which eventually pave the way for the next wave of defectors.

Notwithstanding these qualitative similarities, the conditions for the evolution of cooperation differ widely between discrete-time models of reciprocity and a continuous-time coaction model. If individuals must base their decisions on past actions, populations exhibiting high levels of cooperation become more vulnerable to invasion by defectors (Fig. 5). In addition, introducing delays in the exchange of information between players rapidly destabilizes cooperation and leads to the emergence of an alternative equilibrium with high rates of defection (Fig. 4). These effects are qualitatively similar to those of increasing the cost of cooperation, because obstacles to social information exchange increase the risk of being exploited while attempting to establish or maintain cooperation. Hence, our results fit with the observation that the best suggested examples of reciprocal cooperation generally concern rather inexpensive behaviours and/or repeated interactions between social partners within a narrow time frame (Dugatkin, 1997; Clutton-Brock, 2009; but see Cheney et al., 2010). This applies, for example, to the exchange of body care (grooming or preening) between two or more group members in many birds and mammals (see Clutton-Brock, 2009 for review).

Several recent theoretical studies have addressed aspects of information exchange in social interactions. For example, conformist social learning in various forms has been shown to promote prosocial behaviour in structured populations (Molleman et al.,

2013; Szolnoki and Perc, 2013a). Similarly, Szolnoki and Perc (2013b) have shown that the invasion of defectors is disfavoured when individuals update their strategy rapidly relative to the timescale at which their actions affect their payoff (e.g., when the distribution of a public good in a cooperative group is delayed). Other studies have come to qualitatively different conclusions. For example, delayed and inconsistent information has no appreciable effect on the evolution of trust and trustworthiness in a trust game (Manapat and Rand, 2012) and in the snow-drift game, a lack of information about payoff-relevant asymmetries between players can even increase the tendency of individuals to cooperate, relative to a situation in which players are aware of such asymmetries (Mesterton-Gibbons and Sherratt, 2011). Hence, it is important to recognize that the relationship between information exchange and cooperation is multi-faceted, depending not only on the type of information that is accessible to players (e.g., information about strategies, current or past actions, or payoff-relevant asymmetries) but also on the strategic context of their interaction (e.g., whether or not it is beneficial to be predictable).

As a final point, we note that, in our extended model, coaction in continuous time and reciprocity in discrete time are points of a continuum along which social partners are more or less constrained in exchanging information in real time (Fig. 4). This captures an important aspect of cooperative interactions under natural conditions, which typically depend on a mix of real-time information and predictions based on past experiences. For example, in a cooperatively breeding fish species where subordinates in groups cooperate to maintain and defend a territory, focal individuals are more likely



to invest in digging out a common shelter if their respective social partner does so as well during the same time interval, but also if it has done so shortly before (Riebli and Taborsky, submitted for publication). In the same study, the effect of a cooperating partner was greatest when the collaboration occurred concurrently (coaction), while the effect of previous helpful behaviour (reciprocity) vanished with increasing time intervals between experience and test. Many natural cooperative interactions are similar to this example in that it is difficult to classify the mechanism of cooperation as either pure coaction or pure reciprocity. This problem has been discussed extensively for the predator inspection behaviour of fishes, which may explore a perceived threat in pairs or groups of conspecifics (Milinski et al., 1990; Pitcher, 1992; Dugatkin, 1997). It has been argued that this behaviour reflects an IPD, with two social partners taking turns in their approaches towards the perceived threat (Milinski et al., 1990). However, in the initial experiments with a mirror, predator-inspecting sticklebacks were actually responding to their simultaneously moving mirror image rather than taking turns with a true social partner while approaching the perceived threat (Milinski, 1987). The interpretation of predator inspection as an IPD has therefore been criticized (cf. Noë, 2006 for review), because individuals have access to real-time information about the behaviour of their partner.

In conclusion, cooperation by coaction is more easily obtained and stabilised against exploitation by defectors than cooperation based on reciprocity, which involves a time lag between investment and compensation. Simultaneous and successive contingent cooperation should, however, not be interpreted as alternative categories, because the time axis underlying experience and response is continuous rather than discrete (cf. Shapiro et al., 2008 for a similar perspective in a non-cooperative context). We therefore argue against imposing an artificial distinction between coaction and reciprocity in empirical studies that aim to explain the contingent responses of social partners to each other's cooperation propensity in a natural context, even if this distinction is of conceptual value in the context of theoretical models of cooperation.

## Acknowledgements

The authors thank Ronald Noë, Franjo Weissing and three anonymous reviewers for their comments on this manuscript. This study was financially supported by the Swiss National Science Foundation (SNF-Grant 3100A0-122511 to MT), the European Research Council (ERC starting Grant 30955 to GSvD) and the Netherlands Organization for Scientific Research (NWO Vidi Grant 864.11.012 to GSvD).

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.jtbi.2014.03.019>.

## Reference

- Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. *Science* 211, 1390–1396.
- Boesch, C., Boesch, H., 1989. Hunting behaviour of wild chimpanzees in the Tai national park. *Am. J. Phys. Anthropol.* 78, 547–573.
- Brosnan, S.F., Salwiczek, Bshary, R., 2010. The interplay of cognition and cooperation. *Phil. Trans. R. Soc. B* 365, 2699–2710.
- Cheney, D.L., Moscovice, L.R., Heesen, M., Mundry, R., Seyfarth, R.M., 2010. Contingent cooperation between wild female baboons. *Proc. Natl. Acad. Sci. USA* 107, 9562–9566.
- Chung, S.H., Herrnstein, R.J., 1967. Choice and delay of reinforcement. *J. Exp. Anal. Behav.* 10, 67–74.
- Clements, K.C., Stephens, D.W., 1995. Testing models of non-kin cooperation: mutualism and the Prisoner's Dilemma. *Anim. Behav.* 50, 527–535.
- Clutton-Brock, T., 2009. Cooperation between non-kin in animal societies. *Nature* 462, 51–57.
- Connor, R.C., 1995. Altruism among non-relatives: alternatives to the 'Prisoners' Dilemma. *Trends Ecol. Evol.* 10, 84–86.
- Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* 34, 579–612.
- Duffy, J., Ochs, J., Vesterlund, L., 2007. Giving little by little: dynamic voluntary contribution games. *J. Publ. Econ.* 91, 1708–1730.
- Dugatkin, L.A., 1997. *Cooperation Among Animals. An Evolutionary Perspective*. Oxford University Press, Oxford.
- Frean, M.R., 1994. The Prisoner's Dilemma without synchrony. *Proc. R. Soc. B* 257, 75–79.
- Gintis, H., Bowles, S., Boyd, R., Fehr, E., 2003. Explaining altruistic behaviour in humans. *Evol. Hum. Behav.* 24, 153–172.
- Hammerstein, P., 2003. Why is reciprocity so rare in social animals? A protestant appeal. In: Hammerstein, P. (Ed.), *Genetical and cultural evolution of cooperation*. MIT Press, Cambridge, pp. 83–93.
- Hauser, M.D., Chen, M.K., Chen, F., Chuang, E., 2003. Give unto others: genetically unrelated cotton-top tamarin monkeys preferentially give food to those who altruistically give food back. *Proc. R. Soc. B* 270, 2362–2370.
- Hofbauer, J., Sigmund, K., 1998. *Adaptive dynamics*. Chapter 9 in *evolutionary games and population dynamics*. Cambridge University Press, Cambridge, pp. 101–112.
- Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci. USA* 102, 10797–10800.
- Kalenscher, T., Pennartz, C.M.A., 2008. Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Progr. Neurobiol.* 84, 284–315.
- Krams, I., Krama, T., Igaune, K., Mand, R., 2008. Experimental evidence of reciprocal altruism in the pied flycatcher. *Behav. Ecol. Sociobiol.* 62, 599–605.
- Macy, M.W., Flache, A., 2002. Learning dynamics in social dilemmas. *Proc. Natl. Acad. Sci. USA* 99, 7229–7236.
- Manapat, M.L., Rand, D.G., 2012. Delayed and inconsistent information and the evolution of trust. *Dyn. Games Appl.* 2, 401–410.
- Marx, L., Matthews, S., 2000. A dynamic voluntary contribution to a public project. *Rev. Econ. Stud.* 67, 327–358.
- McNamara, J.M., Gasson, C.E., Houston, A.I., 1999. Incorporating rules for responding into evolutionary games. *Nature* 401, 368–371.
- Mendres, K.A., de Waal, F.B.M., 2000. Capuchins do cooperate: the advantage of an intuitive task. *Anim. Behav.* 60, 523–529.
- Mesterton-Gibbons, M., Sherratt, T.N., 2011. Information, variance and cooperation: minimal models. *Dyn. Games Appl.* 1, 419–439.
- Milinski, M., 1987. Tit-for-tat in sticklebacks and the evolution of cooperation. *Nature* 325, 433–435.
- Milinski, M., Pflüger, D., Külling, D., Kettler, R., 1990. Do sticklebacks cooperate repeatedly in reciprocal pairs. *Behav. Ecol. Sociobiol.* 27, 17–21.
- Milinski, M., Wedekind, C., 1998. Working memory constrains human cooperation in the Prisoner's Dilemma. *Proc. Natl. Acad. Sci. USA* 95, 13755–13758.
- Molleman, L., Pen, I., Weissing, F.J., 2013. Effects of conformism on the cultural evolution of social behaviour. *PLoS One* 8, e68153.
- Noë, R., 2006. Cooperation experiments: coordination through communication versus acting apart together. *Anim. Behav.* 71, 1–18.
- Nowak, M., Sigmund, K., 1989. Oscillations in the evolution of reciprocity. *J. Theor. Biol.* 137, 21–26.
- Nowak, M., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364, 56–58.
- Nowak, M.A., Sigmund, K., 1994. The alternating Prisoner's dilemma. *J. Theor. Biol.* 168, 219–226.
- Nowak, M.A., Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. *Nature* 393, 573–577.
- Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., Bonhoeffer, S., 2005. Evolution of cooperation by generalized reciprocity. *Proc. R. Soc. B* 272, 1115–1120.
- Pitcher, T., 1992. Who dares, wins – the function and evolution of predator inspection behavior in shoaling fish. *Neth. J. Zool.* 42, 371–391.
- Raihani, N.J., Bshary, R., 2011. Resolving the iterated prisoner's dilemma: theory and reality. *J. Evol. Biol.* 24, 1628–1639.
- Riebli, T., Taborsky, M., 2014. Coaction and Reciprocal Cooperation Among Unrelated Individuals in the Highly Social Cichlid *Neolamprologus pulcher* (submitted for publication).
- Rutte, C., Taborsky, M., 2008. The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): direct vs generalised reciprocity. *Behav. Ecol. Sociobiol.* 62, 499–505.
- Schelling, T.C., 1960. *The Strategy of Conflict*. Harvard University Press, Cambridge MA, pp. 45–46.
- Schuster, R., 2002. Cooperative coordination as a social behaviour – experiments with an animal model. *Hum. Nat.* 13, 47–83.
- Selten, R., 1975. A reexamination of the perfectness concept for equilibrium points in extensive games. *Int. J. Game Theor.* 4, 25–55.
- Shapiro, M.S., Siller, S., Kacelnik, A., 2008. Simultaneous and sequential choice as a function of reward delay and magnitude: normative, descriptive and process-based models tested in the European Starling (*Sturnus vulgaris*). *J. Exp. Psychol. – Anim. Behav. Process.* 34, 75–93.
- Stephens, D.W., McLinn, C.M., Stevens, J.R., 2002. Discounting and reciprocity in an iterated prisoner's dilemma. *Science* 298, 2216–2218.
- Stevens, J.R., Cushman, F.A., Hauser, M.D., 2005. Evolving the psychological mechanisms for cooperation. *Annu. Rev. Ecol. Syst.* 36, 499–518.

- Szolnoki, A., Perc, M., 2013a. Information sharing promotes prosocial behaviour. *New J. Phys.* 15, 053010.
- Szolnoki, A., Perc, M., 2013b. Decelerated invasion and waning-moon patterns in public goods games with delayed distribution. *Phys. Rev. E* 87, 054801.
- Taborsky, M., 2007. Cooperation built the tower of Babel. *Behav. Process.* 76, 95–99.
- Trivers, R.L., 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- Van Doorn, G.S., Hengeveld, G.M., Weissing, F.J., 2003. The evolution of social dominance I: two-player models. *Behaviour* 140, 1305–1332.
- Wiltermuth, S.S., Heath, C., 2009. Synchrony and cooperation. *Psychol. Sci.* 20, 1–5.