ESCORT EVOLUTIONARY GAME THEORY

MARC HARPER

Abstract. A family of replicator-like dynamics, called the escort replicator equation, is constructed using information-geometric concepts and generalized information entropies and divergences from statistical thermodynamics. Lyapunov functions and escort generalizations of basic concepts and constructions in evolutionary game theory are given, such as an escorted Fisher’s Fundamental theorem and generalizations of the Shahshahani geometry.

1. Introduction

Recent interest in generalized entropies and approaches to information theory include evolutionary algorithms using new selection mechanisms, such as generalized exponential distributions in place of Boltzmann distributions [10,11] and the use of generalized logarithms and information measures in dynamical systems [19], statistical physics [29], information theory [6, 20, 22], and complex network theory [30]. This paper brings these generalizations to evolutionary dynamics, yielding generalizations of the replicator equation and the orthogonal projection dynamic, deeping the connection of the replicator equation and the relationship with information theory.

Information geometry yields fundamental connections between evolutionary stability and information theory, generating the replicator equation from information divergences [14]. Statistical thermodynamics describes generalized information divergences, such as the Tsallis divergence, using escort functions [22]. Lyapunov functions are constructed from generalized information divergences. Generalizations of Fisher’s fundamental theorem and Kimura’s maximal principle are formulated.

The motivation for these constructions comes from the fact that while Fisher information and the Kullback-Liebler divergences have...
nice uniqueness properties, they are not the only way to measure information. Many other divergences have been studied in other contexts and a natural question is to ask what differences arise in models of natural selection that measure information differently. Using alternative logarithms allows the generalization to a large class of divergences and produces new evolutionary dynamics following the framework established in [14].

2. Statistical Definitions

This section describes the necessary functions and definitions to define the escort replicator equation. The following definitions are due to or adapted from contemporary work in statistical thermodynamics and information geometry; see [21] and [4].

Assume a given function $\phi$, called an escort, that is strictly positive on $(0,1)$. An escort induces a mapping of the set of categorical probability distributions (the simplex) into itself by applying the escort to each coordinate and normalizing; a similar map is induced from $\mathbb{R}^n_+$ to $\Delta^n$. A common escort is $\phi(x) = x^q$, which yields constructions of the Tsallis and Rényi type. This is a $q$-deformation as the limit $q \to 1$ recovers the distribution.

Throughout, a function of a single variable applied to a vector is to be interpreted as applying to each coordinate, i.e. if $x = (x_1, \ldots, x_n)$ then $\log(x) = (\log(x_1), \ldots, \log(x_n))$. This abuse of notation will be clear from context and avoids excessive notation.

**Definition 1** (Partition Function, Escort Distribution). For a distribution $x = (x_1, x_2, \ldots, x_n)$, define the partition function

$$Z_\phi(x) = \sum_{i=1}^{n} \phi(x_i),$$

and the escort distribution

$$\hat{\phi}(x) = \frac{1}{Z_\phi(x)}(\phi(x_1), \phi(x_2), \ldots, \phi(x_n)) = \frac{\phi(x)}{Z_\phi(x)}.$$

The escort function has various effects depending on its form. In the case of a constant escort, all distributions are mapped to the uniform distribution and the mean is just the average. The induced endomorphism on the set of discrete probability distributions is terminal, mapping every point of the simplex to a single point, the barycenter. If the escort function does not have the property that $\phi(0) = 0$, such as for $\phi(x) = e^x$, then the escort changes an event with probability $x_i = 0$ to nonzero probability, affecting the computation of the mean if the escort
distribution is used. For such an escort, the induced endomorphism maps the boundary of the simplex into the interior of the simplex.

Notation for a generalized mean is convenient. The escort expectation is the expectation taken with respect to the escort distribution in place of the original distribution.

**Definition 2 (Escort Expectation).** For a distribution \( x = (x_1, x_2, \ldots, x_n) \) and a vector \( f \) define the escort expectation

\[
\mathbb{E}_x^\phi[f] = \mathbb{E}_{\hat{\phi}(x)}[f] = \frac{1}{Z_\phi(x)} \sum_{i=1}^n \phi(x_i) f_i = \frac{\phi(x) \cdot f}{Z_\phi(x)} = \hat{\phi}(x) \cdot f.
\]

Escort information divergences are defined by generalizing the natural logarithm using an escort function.

**Definition 3 (Escort Logarithm).** Define the escort logarithm

\[
\log_\phi(x) = \int_1^x \frac{1}{\phi(v)} \, dv
\]

For example, for the function \( \phi(x) = x^q \), the escort logarithm is

\[
\log_\phi(x) = \frac{x^{1-q} - 1}{1 - q}.
\]

The limit \( q \to 1 \) recovers the natural logarithm. A generalization of the logistic map of dynamical systems using this logarithm is given in [19]. This logarithm is used in Tsallis statistics, a generalized approach statistical thermodynamics, and has recently been used in complex network analysis.

The escort logarithm shares several properties with the natural logarithm. For instance, the escort logarithm is negative on \((0, 1)\) and positive on \((1, \infty)\); it is concave if \( \phi \) is strictly increasing. Define the function \( \exp_\phi \) to be the inverse function of \( \log_\phi \).

**Definition 4 (Escort Divergences).** Define the escort divergence (or \( \phi \)-divergence)

\[
D_\phi(x||y) = \sum_{i=1}^n \int_{x_i}^{y_i} \log_\phi(u) - \log_\phi(x_i) \, du.
\]

The identity function \( \phi(x) = x \) generates the usual logarithm and exponential with the Kullback-Liebler divergence. Setting \( \phi(x) = x^q \) generates divergences similar to Tsallis and Rényi divergences, and the \( \alpha \)-divergence of information geometry [4][9].
3. **Geometry and Dynamics**

Define the escort metric (or $\phi$-Fisher information metric)

$$g_{ij}^\phi(x) = \frac{1}{\phi(x_i)} \delta_{ij}$$

on the simplex. This is a Riemannian metric since the escort $\phi$ is strictly positive and so the metric is positive definite. The metric may be obtained as the Hessian of the escort divergence. The identity function $\phi(x) = x$ generates the Fisher information metric, also known as the Shahshahani metric in evolutionary game theory. Denote the simplex with the escort metric as the escort manifold.

The Shahshahani metric pulls back to the Fisher information metric

$$g_{ij}(x) = \mathbb{E} \left[ \frac{\partial \log p}{\partial x^i} \frac{\partial \log p}{\partial x^j} \right].$$

The escort metric can be obtained from a Fisher-like formula

$$g_{ij}(x) = \mathbb{E} \left[ \frac{\partial \log p}{\partial x^i} \frac{\partial \log \phi p}{\partial x^j} \right].$$

In coordinates,

$$g_{ij}(x) = \sum_k x_k \frac{1}{x_k} \delta_{ik} \ast \frac{1}{\phi(x_k)} \delta_{kj} = \frac{1}{\phi(x_i)} \delta_{ij}.$$
It is known that the replicator equation is the gradient flow of the Shahshahani manifold, with the right hand side of the equation a gradient with respect to the Shahshahani metric if the fitness landscape is a Euclidean metric \[15\, 26\]. This generalizes to the escort manifold.

**Proposition 1.** Let \( \hat{f}_{\phi, i} = \phi(x_i) \left( f_i(x) - \mathbb{E}_x^\phi [f(x)] \right) \); \( \hat{f}_\phi \) is a gradient with respect to the \( \phi \)-metric if \( f \) is a Euclidean gradient.

**Proof.** For notational convenience we denote \( \hat{f}_{\phi, i} \) by \( \hat{f}_i \) for this proof. Note first that for \( \hat{f}(x) \) to be in the tangent space of the \( n \)-simplex requires that \( \sum_i \hat{f}_{\phi, i}(x) = 0 \).

\[
\sum_i \hat{f}_i(x) = \sum_i \phi(x_i) \left( f_i(x) - \mathbb{E}_x^\phi [f(x)] \right) = \sum_i \phi(x_i)f_i(x) - \sum_i \phi(x_i) \left( \mathbb{E}_x^\phi [f(x)] \right) = \sum_i \phi(x_i)f_i(x) - \sum_j \phi(x_j)f_j(x) = 0
\]

To see that \( \hat{f} \) is a gradient with respect to the escort metric, recognize that the gradient is defined uniquely by the relation \(<\,\text{grad} \, f, z \,>_{x} = D_x V(z)\), where \( f \) is the euclidean gradient of \( V \).

\[
<\hat{f}, z >_{x} = \sum_i \frac{1}{\phi(x_i)} \hat{f}_i(x) z_i = \sum_i \frac{1}{\phi(x_i)} \phi(x_i) \left( f_i(x) - \mathbb{E}_x^\phi [f(x)] \right) z_i = \sum_i f_i(x) z_i - \sum_i z_i \left( \mathbb{E}_x^\phi [f(x)] \right) = \sum_i f_i(x) z_i - \left( \sum_i z_i \right) \left( \mathbb{E}_x^\phi [f(x)] \right) = \sum_i f_i(x) z_i = \sum_i \frac{\partial V}{\partial x_i}(x) z_i = D_x V(z)
\]

Define the escort replicator equation as

(1) \[ \dot{x}_i = \phi(x_i) \left( f_i(x) - \mathbb{E}_x^\phi [f(x)] \right) \]

Like the replicator equation, it is the gradient flow only when the fitness landscape is a Euclidean gradient.
The escort dynamic can be arrived at by deforming the replicator equation: noting that \( \frac{d}{dt} \log φ x_i = \frac{\dot{x}_i}{φ(x_i)} \), rewrite the escort dynamic as

\[
\frac{d}{dt} (\log φ x_i) = f_i(x) - \mathbb{E}_x [f(x)],
\]

which is a deformation of the replicator dynamic,

\[
\frac{d}{dt} (\log x_i) = f_i(x) - \mathbb{E}_x [f(x)].
\]

**Remark 1.** The escort dynamic shares many properties with the replicator dynamic. For instance, a Nash equilibrium of the fitness landscape \( f \) is a rest point of the escort dynamic. Indeed, if \( \hat{x} \) is a Nash equilibrium, there is a constant \( c \) such \( f(\hat{x}) = c \) for all \( i \in \text{supp}(\hat{x}) \), from which it follows that \( \mathbb{E}_x^{\phi} [f(\hat{x})] = c \). This gives a rest point of the dynamic since \( \dot{x}_i = \phi(x_i) \left( f_i(\hat{x}) - \mathbb{E}_x^{\phi} [f(\hat{x})] \right) = \phi(x_i)(c - c) = 0 \). Another method to understand that Nash equilibria are stationary is that the escort dynamic is an orthogonal projection with respect to the escort metric. Since the orthogonal complement to the tangent space of any point on the interior of the simplex is the span of \( (1, 1, \ldots, 1) \), a Nash equilibrium is already orthogonal.

### 4. Generalizations of Fundamental Results

A generalized version of Kimura’s Maximal Principle follows immediately from the fact that the escort dynamic is a gradient. A simple calculation gives an analog of Fisher’s Fundamental Theorem of Natural Selection.

**Theorem 2** (Escorted Fisher’s Fundamental Theorem of Natural Selection).

\[
\frac{d}{dt} V(x) = Z_\phi(x) \text{Var}_x^\phi [f(x)]
\]

**Proof.**

\[
\dot{V}(x) = D_x V(\dot{x}) = \langle \hat{f}_\phi(x), \hat{f}_\phi(x) \rangle = \sum_{i=1}^n \frac{1}{\phi(x_i)} [\phi(x_i)(f_i(x) - \mathbb{E}_x^{\phi} [f(x)])]^2 \\
= Z_\phi(x) \mathbb{E}_x^{\phi} [(f_i(x) - \mathbb{E}_x^{\phi} [f(x)])^2] = Z_\phi(x) \text{Var}_x^\phi [f(x)]
\]

□

The Kullback-Liebler divergence gives a Lyapunov function for the replicator equation, a result which can now be generalized.
Theorem 3. A state $\hat{x}$ is an ESS if and only if the escort divergence $\tilde{D}_{\phi}(\hat{x}||x)$ is a Lyapunov function for the escort dynamic.

Proof. Let $V(x) = D_{\phi}(\hat{x}||x)$. For non-decreasing $\phi$, $V$ is positive, and zero if and only if $\hat{x} = x$.

Taking the time derivative gives:

$$\dot{V} = -\sum_i (\hat{x}_i - x_i) \frac{1}{\phi(x_i)} \dot{x}_i$$
$$= -\sum_i (\hat{x}_i - x_i) \frac{1}{\phi(x_i)} \phi(x_i)(f_i(x) - E_{x}^{\phi} [f(x)])$$
$$= -\sum_i (\hat{x}_i - x_i)(f_i(x) - E_{x}^{\phi} [f(x)])$$
$$= -(\hat{x} - x) \cdot f(x) < 0,$$

where the last inequality is true if and only if the state $\hat{x}$ is an ESS. □

Remark 2. The escort dynamic is a special case of the construction of the adaptive dynamic. This immediately gives a local Lyapunov function from the geodesic approximation [15]. A second is given in the case of an escort gradient, if the landscape is a Euclidean gradient. A third is given above in terms of the dual divergence in the preceding theorem, which is global and independent of the form of the fitness landscape.

5. Examples

5.1. Replicator dynamic. Let $\phi(x) = x$ be the identity function. The $\log_{\phi}$ and escort distributions are then just the ordinary logarithm and distribution, and the escort replicator equation is the replicator dynamic. The induced metric is the Shahshahani metric, and the induced divergence is the Kullback-Liebler divergence. For a detailed exposition of this case see [14].

5.2. Replicator dynamic with selection intensity. Let $\phi(x) = \beta x$, where $\beta > 0$ is the inverse temperature. The induced escort mapping is the identity because the parameter $\beta$ cancels. The escort logarithm is $\frac{1}{\beta} \log x$ and the induced escort dynamic is

$$\dot{x}_i = \beta x_i (f_i(x) - \bar{f}(x)),$$

a form in which special cases have been derived using Fermi selection from stochastic differential equations (with parameter $\beta/2$) [24]. The parameter $\beta$ can be interpreted as the intensity of selection. It affects the velocity of selection but not the trajectories. This is related to the
fact that the Fisher information metric is the unique metric, up to a
constant multiple, on the simplex that respects sufficient statistics \[7\].
In general, given any escort function \( \psi \), a new escort \( \phi(x) = \beta \psi(x) \)
induces an escort dynamic with a leading intensity of selection factor.
The trajectories are the same up to a change in velocity.

5.3. **F-divergences.** Let \( F \) be a convex function on \((0, \infty)\) and consider the \( F \)-divergence

\[
D_F(x||y) = \sum_i x_i F\left(\frac{y_i}{x_i}\right).
\]

The \( F \)-divergence, when localized, produces a metric of the form \( g_{ij}(x) = F''(1) \frac{1}{x_i} \delta_{ij} \), so these divergences yield dynamics that are special cases of the previous example \[4\].

This example also informs the definition the escort divergence. The Kullback-Liebler divergence is often written as

\[
D_{KL}(x||y) = \sum_i x_i \log\left(\frac{x_i}{y_i}\right),
\]

which suggests the escort divergence take the form

\[
D_\phi(x||y) = \sum_i x_i \log_\phi\left(\frac{x_i}{y_i}\right).
\]

This definition would fall into the case of an \( F \)-divergence and so not produce interesting new dynamics, and follows from the fact that \( \log_\phi(xy) \neq \log_\phi x + \log_\phi y \) in general.

5.4. **\( q \)-deformed replicator dynamic.** Let \( \phi(x) = x^q \). This generates the \( q \)-metric \( x_i^{-q} \delta_{ij} \) and the \( q \)-deformed replicator dynamic

\[
\dot{x}_i = x_i^q \left( f_i(x) - \sum_j x_j^q f_j(x) \right) = x_i^q (f_i(x) - \mathbb{E}_x [f(x)]).
\]

The limiting case \( q \to 1 \) yields the replicator equation, as is usual for \( q \)-deformations. The \( q \)-deformed replicator dynamic is formally similar, but not identical, to the dynamic generated by the geometry of escort parameters, viewing \( q \) as parameter, derived in \[1\]. For detail regarding the escort logarithms and Tsallis entropy see \[13\]. For a detailed exposition of the information geometry in this case see \[23\]. The case \( q = 0 \) is the orthogonal projection dynamic \[17\].
5.4.1. Poincare dynamic. Consider the case $\phi(x) = x^2$. This generates the Poincare metric [16], and an entropy similar to Simpson diversity. Call the associated dynamic the Poincare dynamic.

Let us consider the following explicit example, adapted from [15]. For the replicator equation with a fitness landscape given by a matrix $A$, so that $f(x) = Ax$, the game matrix $A$ is called zero sum if $a_{ij} = -a_{ji}$. This yields a mean fitness $\bar{f}(x) = x \cdot Ax = 0$. For instance, consider the rock-scissors-paper matrix

$$A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$  

This gives a zero-sum game with fitness landscape $f(x) = (x_2 - x_3, x_3 - x_1, x_1 - x_2)$ and internal rest point $\hat{x} = (1/3, 1/3, 1/3)$. The dynamic has an integral of motion given by $\sum_i \hat{x}_i \log x_i$.

For the Poincare dynamic, it is not the case that $E_{\phi}^0 [f(x)] = 0$, but the quadratic fitness landscape $f(x) = A\phi(x) = (x_2^2 - x_3^2, x_3^2 - x_1^2, x_1^2 - x_2^2)$ does have zero escort mean. This property is an analog of zero-sum, having a similar effect on the Poincare dynamic which now takes the form $\dot{x}_i = x_i^2 f_i(x)$. The state $\hat{x} = (1/3, 1/3, 1/3)$ is still an interior equilibrium, and using the fact that $\log \phi(x) = 1 - \frac{1}{x}$, we have an integral of motion for the Poincare dynamic given by $\sum_i \hat{x}_i \log \phi x_i = \frac{1}{3} (3 - \frac{1}{x_1} - \frac{1}{x_2} - \frac{1}{x_3})$, or equivalently, $\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3}$. Observe that

$$\frac{d}{dt} \left( \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} \right) = - \left( \frac{x_1}{x_1^2} + \frac{x_2}{x_2^2} + \frac{x_3}{x_3^2} \right) = -(f_1(x) + f_2(x) + f_3(x)) = 0.$$  

The example easily generalizes to longer cycles and arbitrary escort functions, yielding a constant of motion $\sum_i \hat{x}_i \log \phi x_i$ in general.

5.5. Orthogonal Projection Dynamic. Let $\phi(x) = 1$, which generates the Euclidean metric. The resulting dynamic is the orthogonal projection dynamic,

$$\dot{x}_i = f_i(x) - \frac{1}{n} \sum_{i=0}^{n} f_i(x),$$  

where the expectation is the average. The Lyapunov function takes the form $\frac{1}{2} ||x - \hat{x}||^2$, which captures the known result. [25]
5.6. Exponential Escort Dynamic. Another example of a dynamic that is not forward-invariant is the dynamic generated by $\phi(x) = e^x$:

$$\dot{x}_i = e^{x_i} \left( f_i(x) - \frac{\sum_j e^{x_j} f_j(x)}{\sum_j e^{x_j}} \right).$$

For an explicit example, consider the fitness landscape $f(x) = e^{-x} = (e^{-x_1}, \ldots, e^{-x_n})$. After some simplification, the dynamic takes the form

$$\dot{x}_i = 1 - \frac{ne^{x_i}}{\sum_j e^{x_j}},$$

the right-hand side of which is clearly nonzero for $x_i = 0$, unless $x_j = 0$ for all $j$, which is impossible because of the constraint $\sum_j x_j = 1$. The barycenter $(\frac{1}{n}, \ldots, \frac{1}{n})$ is an interior rest point. Note that the escort geometry in this case does not diverge on the boundary.

5.7. Remarks. Notice that the escort replicator equation is forward-invariant on the simplex, in the case of a non-constant fitness landscape, if and only if $\phi(0) = 0$, which is the case for the $q$-deformed replicator dynamic if $q > 0$, but not the case for the orthogonal projection dynamic, though the geometry is trivial. The exponential escort dynamic yields an example that is not forward-invariant with a non-trivial geometry.

The escort dynamic satisfies a gauge invariance similar to the replicator equation. The orthogonal complement to the tangent space of the simplex is the line generated by the vector $1 = (1, 1, \ldots, 1)$. Any function $g(x)\mathbf{1}$ mapping into the complement can be added to the fitness landscape without affecting the trajectories of the escort dynamic because the quantity $f_i(x) - \mathbb{E}_x^\phi [f(x)]$ is preserved: $f_i(x) + g(x) - \mathbb{E}_x^\phi [f(x) + g(x)\mathbf{1}] = f_i(x) + g(x) - (\mathbb{E}_x^\phi [f(x)] + g(x)) = f_i(x) - \mathbb{E}_x^\phi [f(x)]$.

In some cases, the escort replicator equation can be transformed into the replicator equation with an altered fitness landscape. If $\phi$ is invertible and differentiable on $[0, 1]$, with nonzero derivative, the transformation $y = \phi(x)$ translates the escort replicator equation to

$$\dot{y}_i = \phi'(x_i) \phi(x_i) \left( f_i(\phi(x)) - \mathbb{E}_x^\phi [f(\phi(x))] \right) = \phi'(x_i) y_i \left( f_i(y) - \mathbb{E}_y [f(y)] \right),$$

so that for the fitness landscape $g(y) = f(\phi(x))$, we have that

$$\dot{y}_i = \phi' (\phi^{-1}(y_i)) y_i \left( g(y) - \bar{g}(y) \right).$$

If $\phi' (\phi^{-1}(y_i))$ is strictly positive and does not depend on $i$ then the last equation can be transformed into the replicator equation

$$\dot{y}_i = y_i \left( g(y) - \bar{g}(y) \right),$$
by a strictly monotonic change in time scale. An example is \( \phi(x) = \beta x \), which was shown above to produce a dynamic equivalent to the replicator equation.

To determine if the dynamics generated by \( \phi(x) \) and \( \psi(y) \) are distinct with respect to change of velocity, consider the equations

\[
\dot{x}_i = \phi(x_i) \left( f_i(x) - \mathbb{E}_x^\phi[f(x)] \right) \quad \text{and} \quad \dot{y}_i = \phi(y_i) \left( f_i(y) - \mathbb{E}_y^\psi[f(y)] \right).
\]

Assuming that \( \psi \) is differentiable with differentiable invertible and that \( \phi \) is differentiable, the transformation \( \phi(x_i) = \psi(y_i) \) yields the equation

\[
\dot{y}_i = (\psi^{-1})'(\phi(x_i)) \phi'(x_i) \dot{x}_i
\]

\[
\dot{y}_i = (\psi^{-1})'(\phi(x_i)) \phi'(x_i) \psi(y_i) \left( f_i(\phi^{-1}(\psi(y))) - \mathbb{E}_y^\psi[f(\phi^{-1}(\psi(y))) \right).
\]

If \((\psi^{-1})'(\phi(x_i)) \phi'(x_i)\) does not depend on \(i\), the \( \phi \) dynamic is a \( \psi \) dynamic (for an altered landscape) after a change in velocity. In the case of \( \phi(x) = x^p \) and \( \psi(y) = y^q \), this quantity becomes

\[
\frac{p}{q} x_i^{\frac{p-q}{q}} \left( f_i(x) - \mathbb{E}_x^\phi[f(x)] \right)
\]

which is independent of \(i\) if and only if \(p = q\).

5.8. Dynamics from Other Common Information Divergences.

The escort dynamic is rather general in that it captures the localized form of many other commonly used information divergences. Bregman divergences are summations \( \sum \varphi(u,v) \) of the form \( \varphi(u,v) = f(u) - f(v) - f'(v)(u-v) \), which localize to \( \ddot{\varphi}(v,v) = f''(v) \), which corresponds to an escort given by the relation \( f''(v) = \frac{1}{\phi(v)} \). Similarly for Burbea-Rao divergences, up to a multiplicative constant \cite{8}. Straight-forward computations show that Jensen-Shannon divergences \cite{18} and capacitory discrimination \cite{12, 28} localize to (possibly a constant multiple of) the Shahshahani form as well.

6. Solutions of the Escort Replicator Equation

In terms of the geometry, exponential families are normal coordinates on the manifold. The generalized exponential furnishes a formal solution to the escort dynamic. Let \( v \) be a solution to \( \dot{v}_i = f_i(x) \); then \( x_i = \exp_\phi(v_i - G) \) is a solution to the escort dynamic where \( G = \mathbb{E}_x^\phi[f(x)] \), which follows from the fact that \( \dot{x}_1 + \cdots + \dot{x}_n = 0 \). The following fact regarding the derivative of the escort exponential is
needed for the proof of the solution (easily shown with implicit differentiation).

**Lemma 1.**

\[
\frac{d}{dx} \exp_\phi x = \phi(\exp_\phi x)
\]

Now observe that

\[
\dot{x}_i = \frac{d}{dt} \exp_\phi (v_i - G) = \phi(\exp_\phi (v_i - G))(v_i - \dot{G}) = \phi(x_i)(f_i(x) - E^\phi_x [f(x)]).
\]

The escort exponential solution is the maximal entropy distribution with respect to the generalized information divergence [22]. Geometrically, the exponentials correspond to the exponential map induced by the manifold structure.

6.1. **Escort Affine Landscapes.** In the case that the fitness landscape is escort affine, i.e. \( f(\exp_\phi(x)) = Ax + b \) for a matrix \( A \) and vector \( b \), an explicit solution can be given. We assume that \( b = 0 \) for simplicity, referring to the fitness landscape as escort log-linear. Adapting the method of [5], let \( f(x) = A \log_\phi(x) \) for some matrix \( A \), and perform a gauge transformation on \( A \) so that the columns are orthogonally projected on the tangent space of the simplex. Explicitly, the matrix \( A = (a_{ij}) \) becomes \( A^c = (a_{ij} - \frac{1}{n} \sum_i a_{ij}) \), which is a valid gauge transformation as it amounts to adding a particular constant value to each column to ensure that the column sums to zero. The equation for the variable \( v \) becomes \( \dot{v}_i = Av_i - AG = A^c v_i \), using \( x_i = \exp_\phi (v_i - G) \). This equation can be solved by eigenvalue methods and, if the equation for \( G \) can be solved, will yield an explicit closed form.

7. **Discrete Escort Replicator Equation**

An escort also gives an escorted discrete escort replicator dynamic. The escorted versions take the form

\[
x'_i = \frac{\phi(x_i) f_i(x)}{E^\phi_x [f(x)]}.
\]

A particularly simple case is the discrete analog of the orthogonal projection dynamic

\[
x'_i = \frac{f_i(x)}{\sum_j f_j(x)},
\]

which treats the fitness landscape as a non-normalized probability distribution estimating the population distribution.
As for the replicator equation, the escort dynamic can be obtained by a limiting procedure from the discrete analog:

\[ x_i' - \phi(x_i) = \frac{\phi(x_i)f_i(x)}{\mathbb{E}_x^\phi[f(x)]} - \phi(x_i) \]

\[ = \frac{\phi(x_i)}{\mathbb{E}_x^\phi[f(x)]} (f_i(x) - \mathbb{E}_x^\phi[f(x)]). \]

After a change in velocity to remove the strictly positive scalar function \( \mathbb{E}_x^\phi[f(x)] \) from the denominator, the escort dynamic is obtained. Because of this correspondence, the escort inference equation is expected to behave well with respect to the escort divergences and escort exponential families, in analogy of the relationship between Bayesian inference and the Kullback-Liebler divergences [27].

8. Vector-valued Escorts

As a final example, notice that we can use a more general escort function that is a vector-valued function of the population distribution. For instance, consider the escort \( \psi(x) = (e^{\beta_1 f_1(x)}, \ldots, e^{\beta_n f_n(x)}) \). Then the above definitions still go through with the appropriate changes, e.g. the metric is now of the form \( g_{ij}(x) = \frac{\delta_{ij}}{\psi_i(x)} \), which is positive definite, and the induced dynamic is

\[ \dot{x}_i = \psi_i(x) \left( f_i(x) - \mathbb{E}_x^\psi[f(x)] \right), \]

which differs because of the dependence on \( i \) and the entire distribution in the leading term \( \psi(x) \). The notational convention that \( \psi(x) \) means \( \psi \) is applied to each coordinate can be dropped for a vector-valued \( \psi \) to obtain the definitions for escort mean and related quantities. If \( \psi \) is identical in each coordinate, the leading factor can be eliminated with a change of velocity, since the dependence on \( i \) is removed. (This is subtly different than for the escort dynamic, where \( \phi(x_i) \) has dependence on the distribution argument but not the functional argument.) For a concrete example, the escort \( \psi(x) = (\beta_1 x_1, \ldots, \beta_n x_n) \) yields a replicator-like dynamic with differential selection pressures on the individual types.

With general \( \psi \), care must be taken with the induced logarithms, as they differ for each coordinate and the defining integrals are now possibly vector-valued, and with the form of the information divergences. Nevertheless, many of the computations above hold with only minor changes (e.g. introduction of indicies for the escort), such as the generalization of Fisher’s fundamental theorem.
9. Discussion

Incorporating generalized information entropies from statistical thermodynamics yields interesting new dynamics and defines a class of dynamics that includes the replicator equation and the orthogonal projection dynamic. Some of the dynamics have interesting combinations of features, such as non-forward-invariance, while retaining nice informative properties. By using constructions from information geometry, Lyapunov functions can be given for the class of dynamics arising from escorts that yield convex logarithms. Analogs of Fisher’s fundamental theorem and other results follow naturally, yielding new evolutionary models.

References

[1] Sumiyoshi Abe. Geometry of escort distributions. Physical Review E, 68, 2003.
[2] Ethan Akin. The geometry of population genetics. Lecture Notes in Biomathematics, 17(30-31), 1979.
[3] Shunichi Amari and A Fujiwara. Gradient systems in view of information geometry. Physica D: Nonlinear Phenomena, 1995.
[4] Shunichi Amari and Hiroshi Nagaoka. Methods of Information Geometry, volume 191 of Translations of Mathematical Monographs. Oxford University Press, 1993.
[5] Nihat Ay and Ionas Erb.
[6] Lisa Borland, Angel R. Plastino, and Constantino Tsallis. Information gain within nonextensive thermostatistics. Journal of Mathematical Physics, 39, 1998.
[7] N.N. Chentsov. Statistical decision rules and optimal inference. 1982.
[8] María del Carmen Pardo and Igor Vajda. On asymptotic properties of information-theoretic divergences. IEEE Transactions on Information Theory, 49(7), 2003.
[9] A. Dukkipati, M.N. Murty, and S. Bhatnagar. Properties of kullback-liebler cross-entropy minimization in nonextensive framework. International Symposium on Information Theory, pages 2374 – 2378, 2005.
[10] Ambedkar Dukkipati, M. Narasimha Murty, and Shalabh Bhatnagar. Generalized evolutionary algorithm based on tsallis statistics. Arxiv, 20044.
[11] Ambedkar Dukkipati, Narasimha M Murty, Shalabh, and Bhatnagar. Information theoretic justification of boltzmann selection and its generalization to tsallis case. IEEE Congress on Evolutionary Computation, pages 1667–1674, 2005.
[12] D. Endres and E. Schindelin. A new metric for probability distributions. IEEE Transactions on Information Theory, 49(7), 2003.
[13] Shigeru Furuichi. On the maximum entropy principle and the minimization of fisher information in tsallis statistics. Journal of Mathematical Physics, 50, 2009.
[14] Marc Harper. Information geometry and evolutionary game theory. 2009.
[15] Josef Hofbauer and Karl Sigmund. Evolutionary Games and Population Dynamics. Cambridge University Press, 1998.
[16] Ryote Horie. An optimization framework of biological dynamical systems. *Journal of Theoretical Biology*, (253):45–54, 2008.

[17] R. Joosten. Generalized projection dynamics in evolutionary game theory. *Papers on Economics and Evolution*, 0811, 2008.

[18] Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1), 1991.

[19] Alexandre Souto Martinez, Rodrigo Silva Gonzalez, and Csar Augusto Sangaletti Terariol. Generalized exponential function and discrete growth models. *Physica A: Statistical Mechanics and its Applications*, 387:5679–5687, 2008.

[20] Jan Naudts. Deformed exponentials and logarithms in generalized thermodynamics. *Physica A*, 316:323–334, 2002.

[21] Jan Naudts. Estimators, escort probabilities, and $\phi$-exponential families in statistical physics. *Journal of Inequalities in Pure and Applied Mathematics*, 5, 2004.

[22] Jan Naudts. Generalized exponential families and associated entropy functions. *Entropy*, 10:131–149, 2008.

[23] Atsumi Ohara. Geometry of distributions associated with tsallis statistics and properties of relative entropy minimization. *Physics Letters A*, 370, 2007.

[24] J.M. Pacheco, A. Traulsen, and M.A. Nowak. Active linking in evolutionary games. *Journal of theoretical Biology*, 243, 2006.

[25] William H. Sandholm, Emin Dokumaci, and Ratul Lahkar. The projection dynamic and the replicator dynamic. *Games and Economic Behavior*, 666-683, 64, 2008.

[26] S. Shahshahani. A new mathematical framework for the study of linkage and selection. *Memoirs of the AMS*, 17(221), 1979.

[27] C. R. Shalizi. Dynamics of bayesian updating with dependent data and misspecified models. *Electronic Journal of Statistics*, 3:1039–1074, 2009.

[28] Flemming Topsøe. Some inequalities for information divergence and related measures of discrimination. *IEEE Transactions on Information Theory*, 46(4), 2000.

[29] C. Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52:479–487, 1988.

[30] Stefan Turner, Fragiskos Kyriakopoulos, and Constantino Tsallis. Unified model for network dynamics exhibiting nonextensive statistics. *Physical Review E*, 76, 2007.

University of California Los Angeles

E-mail address: marcharper@ucla.edu