frontiers | Frontiers in Physics

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fphy.2022.861633

A Cost Metric for Team Efficiency

Ira S. Moskowitz *

Naval Research Laboratory, Information Management and Decision Architectures Branch, Washington, DC, United States

We use a Riemannian metric as a cost metric when it comes to the optimal decisions that should be made in a multi-agent/Team scenario. The two parameters of interest to us are Team skill and Team interdependence, which are modeled as Wiener process drift and the inverse of Wiener process diffusion, respectively. The underlying mathematics is presented, along with some approximating rules of thumb. It is noteworthy that the mathematics points to, what seems at first, counter-intuitive paradigms for Team performance. However, in reality the mathematics shows a subtle interplay between the factors affecting Team performance.

Keywords: agents, team, Wiener process, Brownian motion, multi-agent system

OPEN ACCESS

Edited by:
William Frere Lawless,
Paine College, United States

Reviewed by:
Sergio Da Silva,
Federal University of Santa Catarina, Brazil
Stephen Russell,
Jackson Memorial Hospital, United States

***Correspondence:**
Ira S. Moskowitz
ira.moskowitz@nrl.navy.mil

Specialty section:
This article was submitted to Interdisciplinary Physics, a section of the journal Frontiers in Physics

Received: *24 January 2022*
Accepted: *28 February 2022*
Published: *25 April 2022*

Citation:
Moskowitz IS (2022) A Cost Metric for Team Efficiency. Front. Phys. 10:861633. doi: 10.3389/fphy.2022.861633

1 INTRODUCTION

We are concerned here with how a multi-agent System (MAS) [2], or Team, reaches the successful conclusion of a task. In Team science, an important parameter for success is interdependence [1, 9, 10, 12]. The Team may be human, machine, or a hybrid. However, our mathematical assumptions implicitly assume that the Team is very machine-like in its behavior and discounts the vagaries of human psychology, e.g., [5]. We address this further at the conclusion of this article.

In [13] it was shown how to model Team behavior as (1-dimensional) Brownian motion [4] (starting at a point Z on the line). In particular, we proposed using Brownian motion $\mathscr{B}(t)$ with (high) drift μ, for (high) Team skill and (low) diffusion σ, for (high) interdependence, arbitrarily starting at a point Z, $0 \leq Z \leq A$. We consider that the Team has succeeded if it reaches point A before 0, and the Team has failed if it reaches 0 before A.

The drift, as mentioned, models Team skill. By way of motivation (using humans), imagine we have a restaurant kitchen crew (building on the restaurant Team given in [11]). We would like all the Team members to have the most skill possible; this would go into the calculation of the drift μ. High skill mapping to high μ. We would also like the Team to work together, hence we desire a high interdependence. Interdependence is the inverse of the diffusion, thus a Team with high interdependence has low diffusion σ, and a Team with each Team member acting independently of the other has high diffusion σ. This leads to the question of which is better—high drift μ, or low diffusion σ? Again, let us go back to our kitchen crew example. If everyone in the kitchen is skilled, but working independently of the others, the result will be a disaster. The dessert will be served before the main course, wine will be served after dessert, etc. Thus, skill alone does not lead to optimal success. On the other hand, consider a kitchen crew with no skill, but working together hand in glove. The results here are also less than optimal—very bad food served in an efficient manner. What is needed is a combination of both factors for optimal Team success, and that is what our idealized mathematics show.

Definition 1. *We say that a stochastic process $\mathscr{W}_t, t \geq 0$ is a Wiener process* [8] if

- $\mathscr{W}_0 = 0$.
- *With probability 1, the function $t \rightarrow \mathscr{W}_t$ is continuous in t.*
- *The stochastic process $\{\mathscr{W}_t\}, t \geq 0$, has stationary, independent increments.*

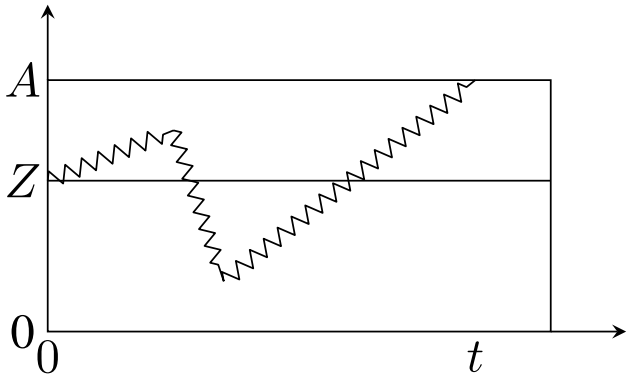

FIGURE 1 | Brownian motion, $\mathscr{B}(t)$, starting at Z and with absorbing boundaries at A (top) and 0 (bottom).

- *The increment, $\mathscr{W}_{t+s}-\mathscr{W}_s$, has the distribution of the standard normal random variable, N (0, t) (this latter part of the definition tells us that $\mathscr{W}_t$ has the distribution of N (0, t)).*

Definition 2. From [13, 15, 17, 18], we say that $\mathscr{B}(t)$ is Brownian motion with drift μ and diffusion σ, that starts at Z, $0 \le Z \le A$, if

$$\mathscr{B}(t) = \mu t + \sigma \mathscr{W}(t) + Z. \tag{1}$$

Let $P_Z(\eth = 0)$ be the probability that $\mathscr{B}(t)$ hits the bottom boundary first (Team failure), then $P_Z(\eth = A) = 1 - P_Z(\eth = 0)$ is the probability that it hits the top boundary first (Team success). **Figure 1** is an example of such a sample path. These probabilities are derived from stopping probabilities ([3]), and we use L'Hôpital's rule for $\mu = 0$.

$$P_Z\Big(\eth = 0\Big) = \begin{cases} \dfrac{e^{-\frac{2A\mu}{\sigma^2}} - e^{-\frac{2Z\mu}{\sigma^2}}}{e^{-\frac{2A\mu}{\sigma^2}} - 1}, & \text{if } \mu \neq 0 \\ 1 - \dfrac{Z}{A}, & \mu = 0 \end{cases} \tag{2}$$

and

$$P_Z\Big(\eth = A\Big) = \begin{cases} \dfrac{e^{-\frac{2Z\mu}{\sigma^2}} - 1}{e^{-\frac{2A\mu}{\sigma^2}} - 1}, & \text{if } \mu \neq 0 \\ \dfrac{Z}{A}, & \mu = 0. \end{cases} \tag{3}$$

For now, we will concentrate on **Equation 3**. Keep in mind that as $\sigma \to 0$, then $P_Z(\eth = A) \to 1$, if $\mu > 0$, and that $P_Z(\eth = A) \to 0$, if $\mu < 0$ (of course, if $Z > A$, this would not hold).

For $Z = 0$, we have that $P_Z(\eth = A) = 0$ since 0 is an absorbing boundary—that is, if we start at 0 we are done. Also, for $Z = A$, we have that $P_Z(\eth = A) = 1$, which also makes sense—if we start at A, we never leave A.

Now say that we need to make an assessment—is it better to modify the drift μ or modify the diffusion σ to increase $P_Z(\eth = A)$? Also, what is the cost of this modification, and how do we measure the cost?

2 FIRST STEPS

We start by defining our manifold **B** and its Riemannian structure.[1]

2.1 Our Manifold B

We would like to know the costs of changing $P_Z(\eth = A)$ as we vary μ and σ. We consider $\mathscr{B}(t)$ in terms of its two parameters, μ and σ.

With this in mind, we define a 2-dimensional Riemannian manifold **B**, homeomorphic to $\mathbb{R} \times \mathbb{R}^+$, and with a global μ, σ chart. We give **B** the Riemannian metric

$$ds^2 = d\mu \otimes d\mu + \frac{1}{\sigma^2} d\sigma \otimes d\sigma. \tag{4}$$

This metric captures the fact that for σ fixed, the difference in μ is simply the standard L_1 distance between them, and that it is independent of the diffusion value. However, the diffusion is also independent of the drift value, but as we attempt to make the diffusion (standard deviation) smaller, it costs more and more, until we approach ∞ at the Dirac distribution.

Note 1—We have chosen to give an infinitesimal distance between points (μ, σ) and $(\mu + d\mu, \sigma + d\sigma)$ and then extend it to a global distance. The Riemannian metric ds^2 captures the fact that changes in μ are Euclidean straight-line distance, whereas changes in σ are based on the inverse of the variance. This concept aligns with how normal distributions differ. We further note that this result is also similar to the Fisher information of the normal distribution (a normalized Poincaré upper-half-plane). What is important about our Riemannian metric is that only the $d\sigma^2$ is modified from the standard Euclidean metric. Again, this emphasizes the fact that changing the mean of the normal distribution is strictly Euclidean, whereas if we attempt to lower the variance, it requires much more "power," and in the limit approaches infinite power. This approach agrees with our thinking that total interdependence (exactly the opposite of independent behavior) has a diffusion of 0, where as totally uncorrelated behavior has infinite diffusion [13, 6.4.2].—

Thus, **B** has the first fundamental form

[1]In this article, we had to make a choice between readability for the non-expert in differential geometry and exact precision with respect to Riemannian geometry. We hope that we have achieved a happy middle ground, and we assure the interested reader that any of the missing fine points can be found in the literature (e.g. [20]).

$$[g_{ij}] = \begin{pmatrix} E & F \\ F & G \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix}, \text{ where } i, j \text{ are indexed independently over } \mu, \nu. \tag{5}$$

Assume we are at $p \in \mathbf{B}$, where $p = (p_\mu, p_\sigma)$, and with tangent vectors $\hat{U} = u_1\frac{\hat{\partial}}{\partial\mu} + u_2\frac{\hat{\partial}}{\partial\sigma}$, $\hat{W} = w_1\frac{\hat{\partial}}{\partial\mu} + w_2\frac{\hat{\partial}}{\partial\sigma}$ at p, where $\frac{\hat{\partial}}{\partial\mu}, \frac{\hat{\partial}}{\partial\sigma}$ are the canonical basis for the tangent space at p.

The inner product between them is

$$\langle \hat{U}, \hat{W} \rangle := u_1 w_1 + \frac{1}{(p_\sigma)^2} u_2 w_2.$$

The norm of a vector $\hat{W}$ is

$$\|\hat{W}\| := \sqrt{\langle \hat{W}, \hat{W} \rangle}.$$

Say $c(t)$ is a smooth curve in $\mathbf{B}$, c: $(a, b) \to \mathbf{B}$, then there is the velocity vector field (on the curve) denoted as $\dot{c}(t)$. This velocity vector field assigns to each point $c(t')$ on the curve $c(t)$ the velocity vector (which is also a tangent vector of M) of the curve at $c(t')$ expressed as $\dot{c}(t')$ (keep in mind that this is multi-dimensional).

That is, $c(t) = (c_\mu(t), c_\sigma(t))$, and $\dot{c}(t) = \dot{c}_\mu(t)\frac{\hat{\partial}}{\partial\mu} + \dot{c}_\sigma(t)\frac{\hat{\partial}}{\partial\sigma}$ (where the raised dot symbol is the usual differentiation with respect to t, and $\frac{\hat{\partial}}{\partial\mu}, \frac{\hat{\partial}}{\partial\sigma}$ are understood to be the canonical tangent space basis vectors at the point $c(t) \in M$). To simplify notation, we can express this as $\dot{c}(t) = \langle \dot{c}_\mu(t), \dot{c}_\sigma(t) \rangle$.

We define the **length** of $c(t)$, denoted as $L(c)$, as

$$L(c) := \int_a^b \|\dot{c}(\tau)\| d\tau = \int_a^b \sqrt{\left[\dot{c}_\mu(\tau)\right]^2 + \frac{\left[\dot{c}_\sigma(\tau)\right]^2}{\left[c_\sigma(\tau)\right]^2}}\, d\tau. \tag{6}$$

Given two points, $p, q \in \mathbf{B}$, and $c(t)$, any smooth curve between them (this can be relaxed to include piece-wise smooth, but not of class C^∞), we define the **distance** between them as

$$d(p, q) := \inf L(c). \tag{7}$$

2.2 Team Geometry

Our metric is modeled on the hyperbolic metric in the Poincaré half-plane model. An important difference is that E does not depend on the σ value. The change in drift is independent of the diffusion value which we feel is the correct way to model Team action. Furthermore, the μ distance is linear with respect to μ. This choice assumes that only the change of Team skill matters, not the values it ranges between. However, the change in diffusion, which is independent of drift, does depend on the different diffusion (interdependence values) that the Team is choosing between. This approach makes sense in terms of a normal distribution. Going from a normal distribution $N(\mu, 10)$ to $N(\mu, 9)$ requires much less change in the distribution itself than going from $N(\mu, 1)$ to $N(\mu, 0.9)$, and again going from $N(\mu, 0.5)$ to $N(\mu, 0.45)$.

We see in **Figure 2** that as $\sigma \to 0^+$, the difference in the normal plots is more severe. This behavior is in contrast to changing μ, which has the effect of shifting the graph to the left or right, but not changing its shape. We use a Riemannian manifold because it gives us the means to modify the metric for other models of Team behavior. This approach is accomplished by adjusting the g_{ij} in **Equation 5**.

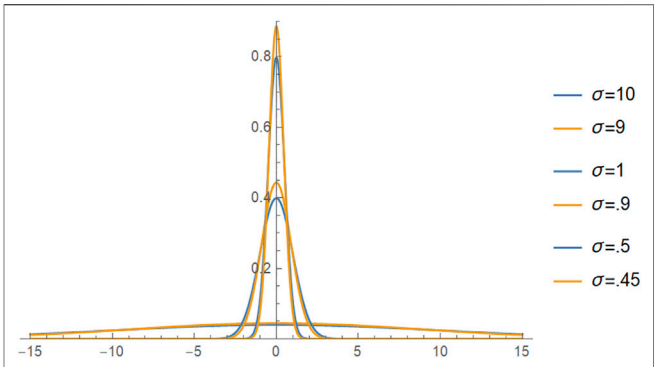

FIGURE 2 | $N(0, \sigma^2)$ for three groups. g The bottom pair is $\sigma = 10, 9$; the middle pair is $\sigma = 1, 0.9$; and the top two, which are the most different are $\sigma = 0.5, 0.45$.

2.3 Curvature and Geodesics

We start by considering the Gaussian (sectional) curvature K of $\mathbf{B}$ as a function of the first fundamental form.

First, using **Equation 5**, we consider the easily obtainable matrices (the sub-index indicates the partial differentiation with respect to that index) for $\mathbf{B}$.

$$\begin{pmatrix} E_\mu & F_\mu \\ F_\mu & G_\mu \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} E_\sigma & F_\sigma \\ F_\sigma & G_\sigma \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{-2}{\sigma^3} \end{pmatrix}. \tag{8}$$

From [14, Eq. (9.22), Eq. (9.33)] we use the Brioschi formula for a Riemannian 2-manifold in general with generic parameters μ, ν, which, for arbitrary E and G, and when $F = 0$ [14, Eq. 9.25], becomes

$$K = \frac{-1}{\sqrt{EG}}\left\{\frac{\partial}{\partial\mu}\left(\frac{1}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial\mu}\right) + \frac{\partial}{\partial\sigma}\left(\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial\mu}\right)\right\} \tag{9}$$

$$= \frac{-1}{2\sqrt{EG}}\left\{\frac{\partial}{\partial\mu}\left(\frac{G_\mu}{\sqrt{EG}}\right) + \frac{\partial}{\partial\sigma}\left(\frac{E_\sigma}{\sqrt{EG}}\right)\right\}. \tag{10}$$

For $\mathbf{B}$, $G_\mu = 0$ and $E_\sigma = 0$, we find that $K = 0$.

Now we move on to the geodesics of $\mathbf{B}$. First we have to find the Christoffel (tensor) symbols (symmetric in the lower indicies). We define these on a local patch of a Riemannian 2-manifold, M, in general with generic parameters μ, σ.

$$\Gamma^\mu_{\mu\mu} = \frac{GE_u + FE_\sigma - 2FF_u}{2(EG - F^2)} \tag{11}$$

$$\Gamma^\mu_{\mu\sigma} = \frac{GE_\sigma - FG_u}{2(EG - F^2)} \tag{12}$$

$$\Gamma^\mu_{\sigma\sigma} = \frac{-FG_\sigma - GG_u + 2GF_\sigma}{2(EG - F^2)} \tag{13}$$

$$\Gamma^{\sigma}_{\mu\mu} = \frac{-FE_u - EE_\sigma + 2EF_u}{2\left(EG - F^2\right)} \tag{14}$$

$$\Gamma^{\sigma}_{\mu\sigma} = \frac{EG_u - FE_\sigma}{2\left(EG - F^2\right)} \tag{15}$$

$$\Gamma^{\sigma}_{\sigma\sigma} = \frac{EG_\sigma + FG_u - 2FF_\sigma}{2\left(EG - F^2\right)}. \tag{16}$$

Thus, for our manifold **B**, we have that all of the Christoffel symbols are 0, except for $\Gamma^{\sigma}_{\sigma\sigma} = \frac{-1}{\sigma}$.

Definition 3. *For $t \in (0, 1)$, a smooth curve $c(t) = (c_1(t), c_2(t))$, $\dot{c}(t) = (\dot{c}_1(t), \dot{c}_2(t))$ in a Riemannian manifold with ∇ being the Levi-Civita* connection [20] is a **geo*desic*** *if*

$$\nabla_{\dot{c}(t)}\dot{c}(t) = 0. \tag{17}$$

In general, one does not need to restrict t to the unit interval, but we have done this as a convenience. In general, geodesics are unique up to an affine parametrization; without loss of generality, we have fixed this by setting the t interval to [0, 1]. Note that by the existence and uniqueness theorem for ordinary differential equations (ODEs), we can find a unique geodesic if we also include the vector values at $c(0)$ and $c'(0)$. (It turns out for **B** that this follows directly.)

We do not want to get into too many of the details of the ∇ operator above. It is covariant differentiation, which is the directional derivative of the vector field $\dot{c}(t)$ in the direction $\dot{c}(t)$ with adjustments for curvature K. Details can be readily found in the literature (e.g. [14]). Since the one local coordinate system (patch) we have given for **B** suffices, the geodesic equation reduces to

$$\ddot{c}_\mu(t) + \sum_{i,j\in\{\mu,\sigma\}} \Gamma^{\mu}_{ij}\dot{c}_i(t)\dot{c}_j(t) = 0, \text{ and} \tag{18}$$

$$\ddot{c}_\sigma(t) + \sum_{i,j\in\{\mu,\sigma\}} \Gamma^{\sigma}_{ij}\dot{c}_i(t)\dot{c}_j(t) = 0, \tag{19}$$

which, using the above values of the Christoffel symbols, simplifies to

$$\ddot{c}_\mu(t) = 0, \text{ and} \tag{20}$$

$$\ddot{c}_\sigma(t) - \frac{(\dot{c}_\sigma)^2}{c_\sigma(t)} = 0. \tag{21}$$

Trivially, we find that

$$c_\mu(t) = at + b.$$

To obtain $c_\sigma(t)$, we need to solve a non-linear second order ODE, so we simplify notation and use the auxiliary variable $w = \dot{c}_\sigma$, which gives $\ddot{c}_\sigma = \frac{dw}{dc_\sigma}\dot{c}_2 = \frac{dw}{dc_\sigma}w$. Now we perform the usual trickery, but check our answer at the end.

$$\begin{aligned}
\ddot{c}_\sigma &= \frac{(\dot{c}_\sigma)^2}{c_\sigma} \\
\frac{dw}{dc_\sigma}w &= \frac{w^2}{c_\sigma} \\
\frac{dw}{w} &= \frac{dc_\sigma}{c_\sigma} \\
w &= \beta c_\sigma \\
\dot{c}_\sigma &= \beta c_\sigma \\
c_\sigma(t) &= \alpha e^{\beta t}.
\end{aligned}$$

which when we check does solve **Eq. 21** for $c_\sigma(t)$ in its most general form. Thus,

$$c(t) = \left(at + b, \alpha e^{\beta t}\right). \tag{22}$$

Theorem 1. The constants a, b, α, β uniquely fix the geodesic.

Proof. Say there are two geodesics $c, \bar{c}$: $[0, 1] \to \mathbf{B}$ such that

$$c(t) = \left(at + b, \alpha e^{\beta t}\right), \text{ and}$$
$$\bar{c}(t) = \left(\bar{a}t + \bar{b}, \bar{\alpha} e^{\bar{\beta} t}\right).$$

Assume they are the same geodesic; then by evaluating the geodesic at $t = 0$, we have that

$$b = \bar{b}, \alpha = \bar{\alpha}.$$

Now using the above and evaluating the geodesics at $t = 1$, we have that

$$a = \bar{a}, \beta = \bar{\beta}.$$

□

So all we have to do now is to determine the four constants in the geodesic curve to uniquely specify it. As noted above, if we specify $c(0) = (\mu_0, \sigma_0)$ and $\dot{c}(0) = (\aleph, ב)$, then simple calculations show that we uniquely fix the geodesic as

$$c(t) = \left(\aleph t + \mu_0, v_0 e^{\frac{ב}{\sigma_0} t}\right). \tag{23}$$

However, we are interested in the boundary value problem to see if knowing $c(0)$, $c'(0)$ also gives us a unique solution. In general, for geodesics on an arbitrary Riemannian manifold, this result need not be true. By way of example, consider the geodesics (where the locus is a great circle) on S^2. Given $c(0)$, $c(1)$, there are infinitely many geodesics that satisfy the conditions (they just keep wrapping around). What is different in our situation, however, is that the geodesics never go back on themselves (this is seen by looking at the form of $c(t)$). If we have that $c(0) = (\mu_0, \sigma_0)$ and $c(1) = (\mu_1, \sigma_1)$, then simple calculations show that these boundary conditions uniquely fix the geodesic as

$$\begin{aligned}
c(t) &= \left(\mu_0 + (\mu_1 - \mu_0)t,\ \sigma_0 e^{t\ln\left(\frac{\sigma_1}{\sigma_0}\right)}\right) \\
&= \left(\mu_0 + (\mu_1 - \mu_0)t,\ \sigma_0\left(\frac{\sigma_1}{\sigma_0}\right)^t\right).
\end{aligned} \tag{24}$$

Thus, for the geodesics c: $[0, 1] \to \mathbf{B}$, we find that a solution exists and, given $c(0)$ and $c(1)$, that the geodesic is uniquely expressed as in **Equation 24**.

Equations 6 and **7** tell us how to obtain a topology based on the metric distance. This topology makes **B** homeomorphic to the upper half-plane with its standard topology. (Note though that **B** is not isometric to the upper half-plane with the standard Euclidean metric.) Since the latter space is complete, so is **B**. By the Hopf-Rinow theorem [20], given an initial point $p = (x_0, y_0)$, and a final point $q = (x_1, y_1)$, there exists a geodesic $c(t)$ between them such that $c(0) = p, c(1) = q$ and $L(c) = d(p, q)$. Given **Equation 24**, we have shown how to uniquely construct such a geodesic; therefore, the

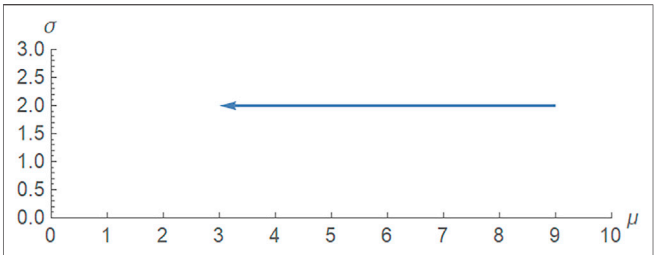

FIGURE 3 | Geodesic starting at p = (9, 2) and ending at q = (3, 2).

geodesic from **Equation 24** has the property that its length is the distance between the points.

The message from this result is that, given two points p, q on **B**, if we find the geodesic between them (remember we only use as a domain [0, 1]), then the length of that geodesic is the distance between them. This result is similar to what occurs in the Poincaré upper half-plane. We also note that this result would not work on S^2 because the boundary values do not uniquely determine the geodesic—as discussed above, the geodesics on S^2 can wrap around themselves, which does not happen on **B** or the Poincaré upper half-plane. Therefore, this leaves us with the following corollary to the above theorem.

Corollary 1.1. Given two points in $p = (\mu_0, \nu_0) \in \mathbf{B}$, and $q = (\mu_1, \sigma_1) \in \mathbf{B}$, there is a unique geodesic c: $[0, 1] \rightarrow \mathbf{B}$ between them such that $c(0) = p$, $c(1) = q$. Furthermore, $c(t)$ is length minimizing, that is, $L(c) = d(p, q)$. The geodesic is as given in **Equation 24**.

Now let us examine the length of our geodesic $c(t)$ between $p = (\mu_0, \sigma_0)$ and $q = (\mu_1, \sigma_1)$. By **Equation 6**, we have that. $L(c) = \int_0^1 \sqrt{(\mu_1 - \mu_0)^2 + [\ln(\frac{\sigma_1}{\sigma_0})]^2}\, d\tau$, thus,

Corollary 1.2. Given two points in $p = (\mu_0, \sigma_0) \in \mathbf{B}$, $q = (\mu_1, \sigma_1) \in \mathbf{B}$, the length of the unique geodesic $c(t)$ between them is

$$\sqrt{(\mu_1 - \mu_0)^2 + \left[\ln\left(\frac{\sigma_1}{\sigma_0}\right)\right]^2}. \tag{25}$$

Let us see what some of these geodesics look like (their traces).

Example 1: Let us examine the case where σ is held constant between two points. Let $p = (\mu_1, \sigma)$, $q = (\mu_2, \sigma)$. The geodesic between them is

$$c(t) = (\mu_0 + (\mu_1 - \mu_0)t, \sigma).$$

We illustrate this result with p = (9, 2), q = (3, 2) in **Figure 3**. When ν is fixed, we are in a standard Euclidean metric, the length of the geodesic is its distance, and it follows from **Eqs 6**, **7** that $d(p, q) = |\mu_1 - \mu_0|$,

The result is a horizontal straight line of length six. When ν is fixed, our geometry is standard Euclidean geometry. Example 2: Let us now look at the opposite situation, when we hold μ fixed and vary σ.

The trace of this geodesic, shown in **Figure 4**, is simply a vertical line, however, its length is not its Euclidean length of 2—0.5 = 1.5, rather its length is $|\ln(.5/2)| = 1.39$. The geometry here is far from Euclidean. But now let us consider the geodesic starting at p = (3, 0.2) and ending at q = (3, 0.05). Again, its length is not the Euclidean length of 0.15; rather, its length is $|\ln(.05/.2)| = 1.39$, the same as the first part of this example. Let us summarize these two examples. For **B**, the length of a geodesic connecting two points with fixed σ is simply their Euclidean distance. However, the length of a geodesic in **B** connecting two points with the same μ only depends on their ratio, the distance being the absolute value of the natural log of the ratio.

Let us look at the general geodesics a bit more. In **Figure 5**, we see paths of the geodesics that start at (μ_0, σ_0) and end at (μ_1, σ_1). These representative samples, along with **Figures 3**, **4**, show the general shape of the geodesics.

An interesting question becomes: Given a point (μ_0, σ_0), what is the locus of points (μ, σ) distance D from this point? From **Equation 25**, we easily have that

$$\sigma = \sigma_0 e^{\pm\sqrt{D^2 - (\mu - \mu_0)^2}}. \tag{26}$$

From this result, we see that $\mu \in [\mu_0 - D, \mu_0 + D]$ and $\sigma \in [\sigma_0 e^{-D}, \sigma_0 e^{D}]$.

We plot in **Figure 6** the locus of points for σ as a 2-valued "function" of μ, with distance 2 from the point (1.5, 3) and when Z=.6 (recall that we have normalized A to 1).

Let us go back to **Eq. 3** and see how $P_{.6}(\eth = 1)$ varies as we look at all of the points at a set distance from (μ_0, σ_1).

We start by summarizing some of the results from [16, Sec 6.2.2] that discuss how $P_Z(\eth = A)$ behaves.

- $P_Z(\eth = A)$ is an increasing function of μ.
- For $\mu > 0$, $P_Z(\eth = A)$ is a decreasing function of σ.
- For $\mu < 0$, $P_Z(\eth = A)$ is an increasing function of σ.

Since the center of our 2-ball is (1.5, 3), let us examine the two points with extreme μ values of -0.5 and 3.5. We find that

$$P_{.6}\Big(\eth = 1\Big)|_{(-.5,3)} = .587 < P_{.6}\Big(\eth = 1\Big)|_{(1.5,3)} = .639$$
$$< P_{.6}\Big(\eth = 1\Big)|_{(3.5,3)} = .690,$$

and

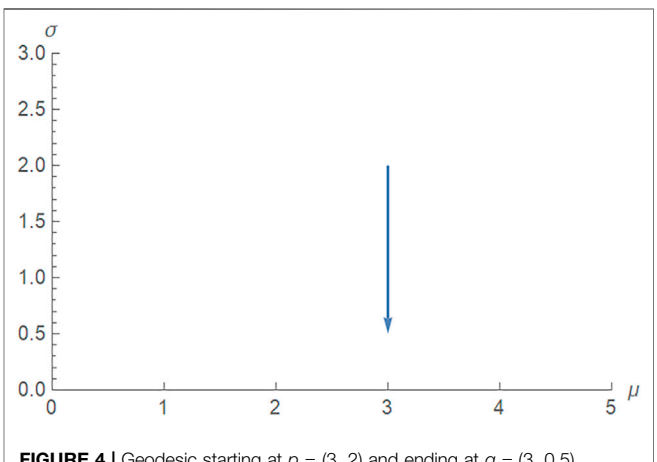

FIGURE 4 | Geodesic starting at p = (3, 2) and ending at q = (3, 0.5).

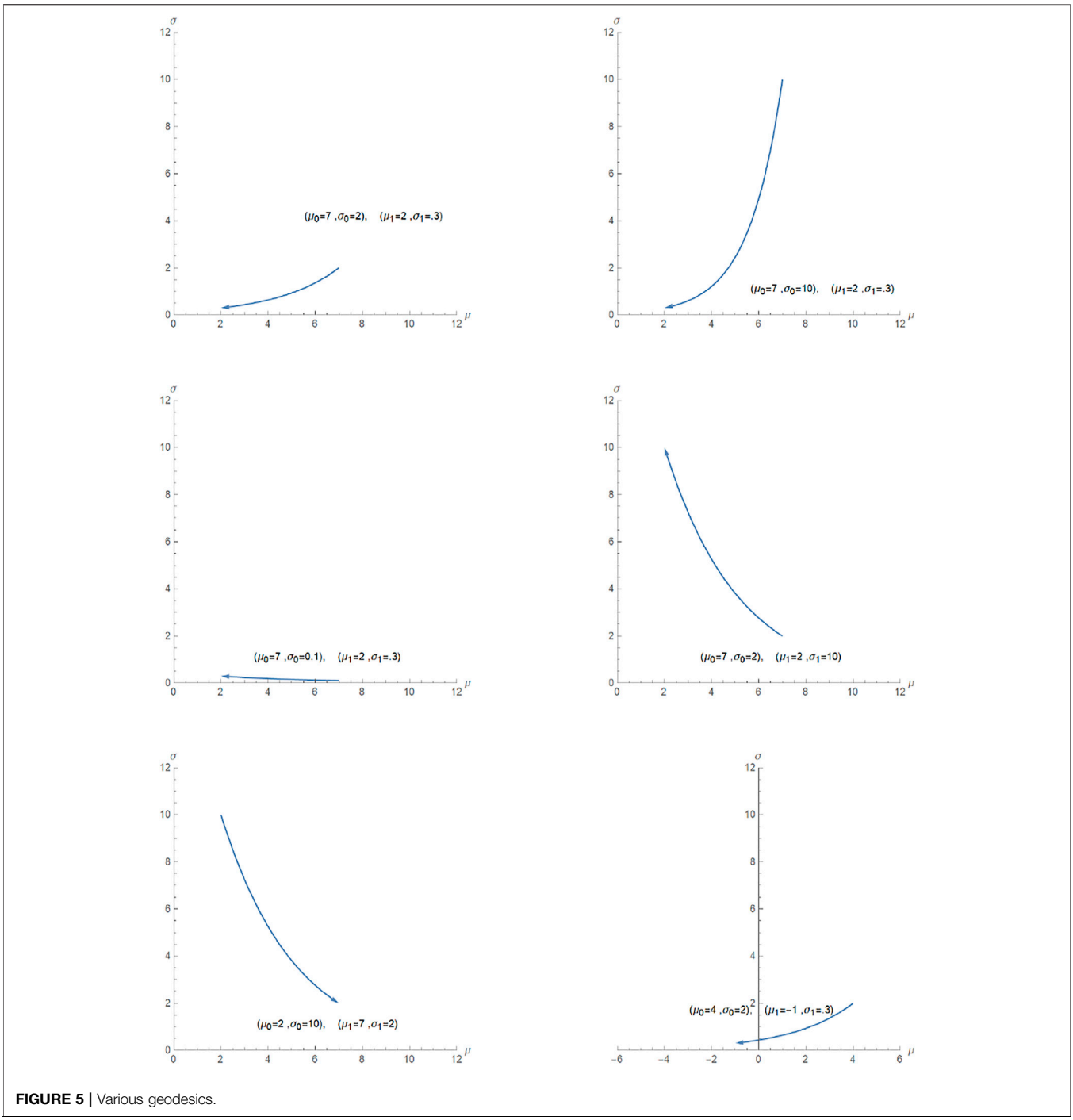

FIGURE 5 | Various geodesics.

$$P_{.6}\left(\eth = 1\right)|_{(3.5,.406)} = .999 > P_{.6}\left(\eth = 1\right)|_{(1.5,3)} = .639 > P_{.6}\left(\eth = 1\right)|_{(3.5,22.17)} = .601.$$

Keep in mind that for μ fixed at 3.5, as σ grows, the boundary probability approaches $0.6 = Z/A = 0.6/1$. This results in the values at the four "corners" of the metric-circle. One might think that the south pole is the highest probability. Let us plot $P_Z(\eth = A) = 1$ as a 2-valued function of μ. That is, we plot $P_Z(\eth = A) = 1$ as a function of μ with $\sigma = \sigma_0 e^{\sqrt{D^2-(\mu-\mu_0)^2}}$ (which corresponds to the top red semi-circle) and $\sigma = \sigma_0 e^{-\sqrt{D^2-(\mu-\mu_0)^2}}$ (which corresponds to the bottom blue semi-circle). The range of μ is $\mu \in [\mu_0 - D, \mu_0 + D]$. The point $(\mu, e^{\sqrt{D^2-(\mu-\mu_0)^2}})$ has a lesser probability than $(\mu, e^{-\sqrt{D^2-(\mu-\mu_0)^2}})$, since for μ fixed, the smaller σ becomes, the greater the probability when μ is positive.

Please note when comparing **Figures 6**, **7** that the blue and red regions have shifted.

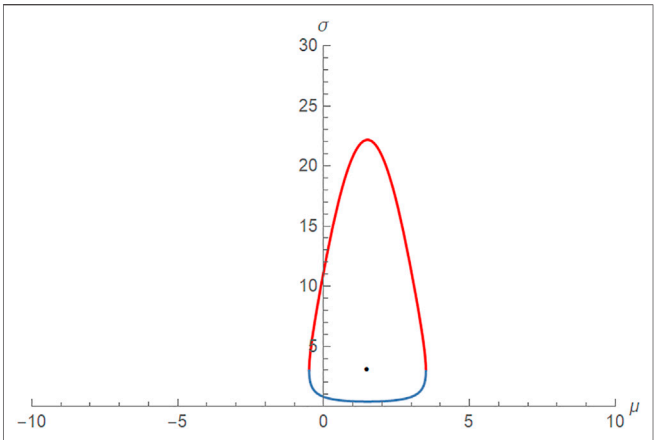

FIGURE 6 | Geodesic locus: Points at a distance 2 from (1.5, 3) with A = 1, Z =0.6. This is our metric "circle."

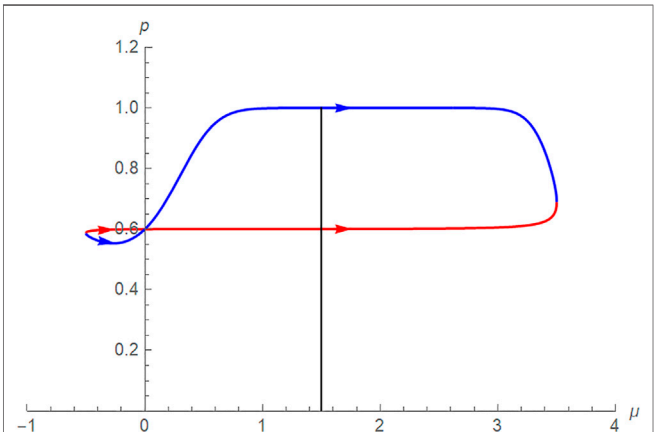

FIGURE 7 | $P_{.6}(\eth = 1)$ of point distance 2 from $(\mu, \sigma) = (1.5, 3)$ as μ increases (indicated by arrow direction) from −0.5 to 3.5. We see that for negative drift, the probabilities have an inverse behavior.

From **Figure 7**, one might infer that the maximum probability occurs at the bottom corner $(\mu, \sigma) = (1.5, 0.406)$. Further numerical analysis shows that this is not true; the actual maximum occurs closer to when $\mu = 2$ (with the corresponding $\sigma > 0.406$). Let us use a different example to show this result better, as in **Figure 8**. Here, it is much more obvious that the maximum probability does not occur at the south pole of the metric circle.

We also see that the minimum probability does not occur for the smallest μ; rather, it too is a combination of a small μ, but with a larger σ.

In **Figure 9**, we have combined the plots from **Figures 6**, **7**. That is, **Figure 6**, which is a plot of a 2-valued function of σ against μ, is sketched in $(\mu, \sigma, 0)$ space. In **Figure 7**, since σ is now a 2-valued function of μ, we see that the probability $P_{.6}(\eth = 1)$ (of points distance 2 from the center (1.5, 3)) is a 2-valued function of μ and lives naturally in (μ, σ, p) space. In other words, the points on the top of a point of distance 2 is

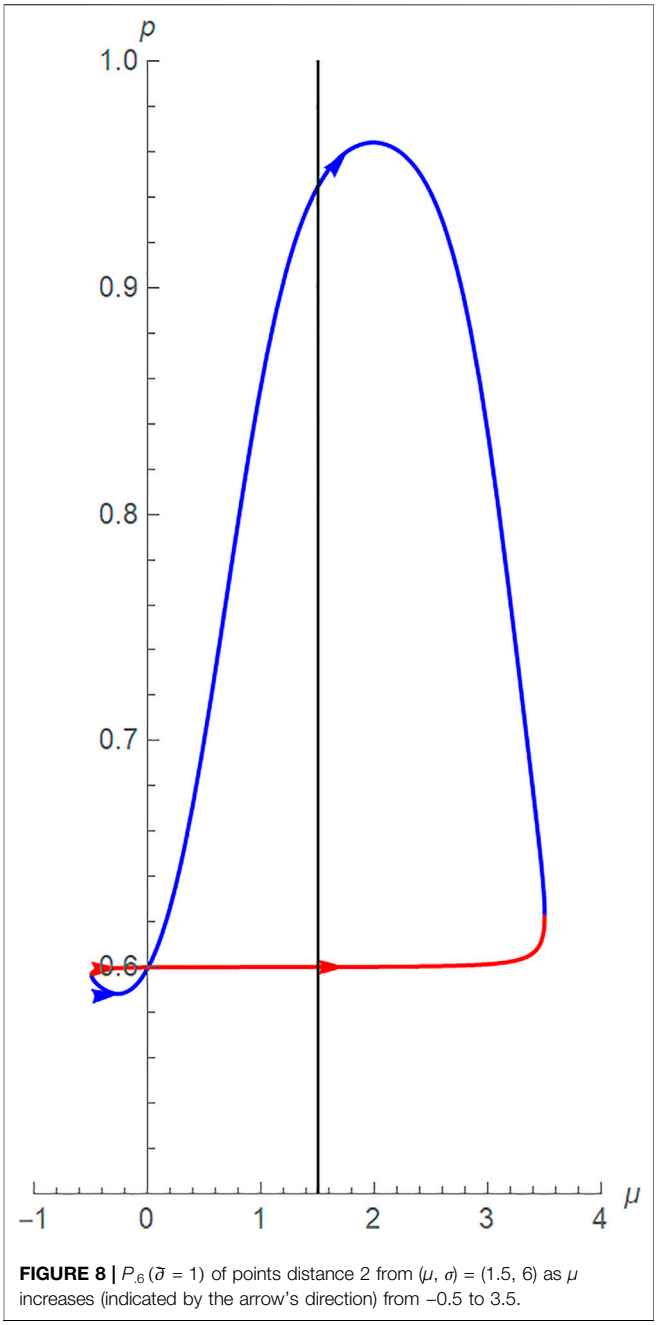

FIGURE 8 | $P_{.6}(\eth = 1)$ of points distance 2 from $(\mu, \sigma) = (1.5, 6)$ as μ increases (indicated by the arrow's direction) from −0.5 to 3.5.

on the top plot. The red curves correspond to $\sigma = 3e^{\sqrt{2^2-(\mu-1.5)^2}}$, and the blue curves correspond to $\sigma = 3e^{-\sqrt{2^2-(\mu-1.5)^2}}$. We see that the red probability hovers around 0.6, whereas the blue approaches, very closely in fact, to a probability of 1.

3 SURFACE GEOMETRY

Let us move away from points a certain Riemannian distance from a point and consider the surface and the level sets of $P_Z(\eth = A)$. As before, we will normalize A to be 1, and let

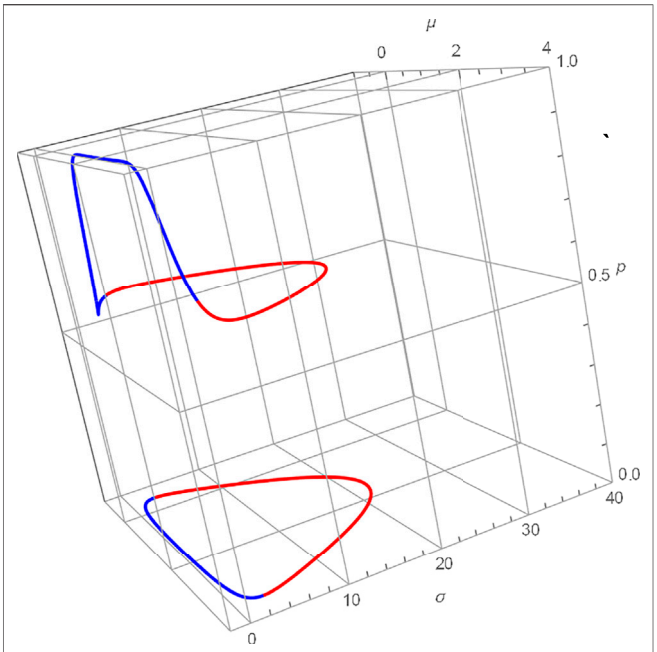

FIGURE 9 | Combo.

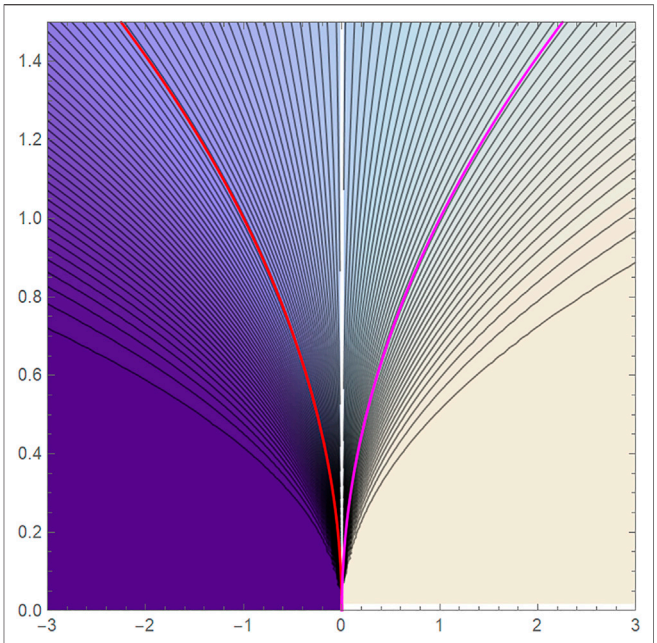

FIGURE 11 | Level sets of $P_{.6}(\eth = 1)$, horizontal axis is μ, vertical axis is σ.

$Z = 0.6$. The behavior of $P_Z(\eth = A)$ as functions of μ and σ has been analyzed in [13], so we will not repeat the results from there. The plot of $P_{.6}(\eth = 1)$ is given in **Figure 10**.

Let us consider the level sets of $P_{.6}(\eth = 1)$; in fact, this holds for the level sets of $P_Z(\eth = A)$ for $Z \in (0, 1)$.

Theorem 2. T*he level sets of* $P_Z(\eth = 1), 0 < Z < 1$ *correspond to constant values of* μ/σ^2.

Proof. ($\Leftarrow$) If $\mu/\sigma^2 = \mu'/\sigma'^2 = k$, then it is obvious that $P_Z(\eth = 1) = \frac{e^{-\frac{2Z\mu}{\sigma^2}}-1}{e^{-\frac{2\mu}{\sigma^2}}-1} = \frac{e^{-2Zk}-1}{e^{-2k}-1} = \frac{e^{-\frac{2Z\mu'}{\sigma'^2}}-1}{e^{-\frac{2\mu'}{\sigma'^2}}-1}$. ($\Rightarrow$) By (16, Cor 3.1), we have that for $C > D > 0$, that $\frac{e^{-Dk}-1}{e^{-Ck}-1}$ is an increasing function of x. Let $C = 2$, $D = 2Z$, we have that $f(k) = \frac{e^{-2Zk}-1}{e^{-2k}-1}$ is an increasing function of k and the result follows. □

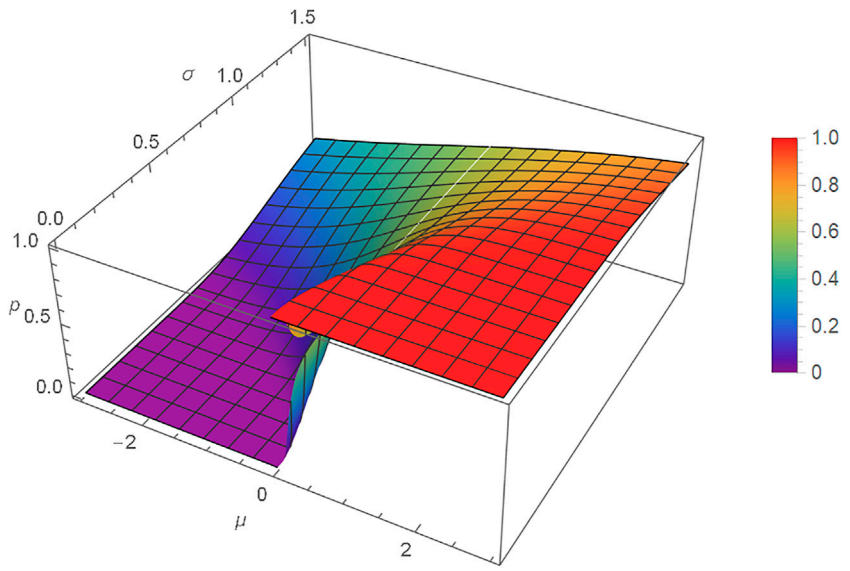

FIGURE 10 | Plot of $P_{.6}(\eth = 1)$ for $\mu \in (-3, 3)$, $\sigma \in (0, 1.5)$. The limited range is due to the fact that the plot is extremely close to 1 for large μ and small σ. The function is continuous and is equal to 0.6 when $\mu = 0$.

Let us look at the level sets of this surface in **Figure 11**. We see 100 level sets going from left to right in ascending order from 0.01 to 0.99 in approximate steps of 0.01. The middle level set is white and that corresponds to $P_{.6}(\eth = 1) = .6$ which occurs when $\mu = 0$.

Keep in mind that every level set of $P_{.6}(\eth = 1)$ corresponds to the curve given by $\mu/\sigma^2 = k$. For $k < 0$, we have the level set on the left hand side (LHS) of **Figure 11**. For $k = 0$, we have the vertical white line at $\mu = 0$, and for $k > 0$ we have the level sets on the right hand side (RHS) of **Figure 11**. Note that we specifically illustrated the level set corresponding to $\mu/\sigma^2 = -1$ which is equivalent to $\sigma = -\sqrt{\mu}$ and corresponds to the red level set on the LHS of the figure; and corresponds to $P_{.6}(\eth = 1) = .363$. We also illustrated the level set corresponding to $k = 1$, which is the purple curve on the RHS of the figure and corresponds to $P_{.6}(\eth = 1) = .808$. Of course, now we see why the maximum values that we discussed above are not at μ corresponding to the center of the metric circle, but to the right. This result occurs because the level curve that is tangent to the plot is where the maximum is found. This result can be analyzed with Lagrange multiplier theory, a direction we will pursue in future work. It suffices for this article to show that the trade-off between μ and σ is non-trivial.

4 IMPACT ON TEAMS AND MULTI-AGENT SYSTEMS

We have learned from the mathematics that the decision to attempt to increase skill or to increase interdependence is not trivial. The best answer is a complicated mathematical expression. We also could have looked at the time to obtain the correct answer, but this is even more complicated and will also be addressed in future work.

Presently, our problem boils down to the probability of reaching the correct answer by using the Riemannian distance described in this article—which gets a Team to the highest new probability of success by staying within the distance constraints on skill and interdependence.

Of course, general rules of thumb can be derived by studying the geometry of the question in hand, and near-optimal solutions may be good enough to satisfy a user.

Recall from [16], for $\mu > 0$ (the situations we have been looking into), the lower the diffusion, the greater the interdependence. We note that our mathematics shows that *to optimize Team performance, it takes a combination of increasing the drift/skill $\mu > 0$ and lowering the diffusion σ (increasing interdependence) to optimize the probability of the Team of multi-agents of reaching the correct conclusion to a problem that it confronts.*

5 CONCLUDING REMARKS

Presently, to avoid complicated Riemannian geometric discussions, it is best to use rules of thumb that can be derived from the various plots of the Teams in question. This present work continues the theme emphasized by Lawless [9, 10] of the importance of interdependence, but also has thrown the skill issue into the mathematical mix. This point is not to say that others have ignored skill, rather that their focus was in the interesting and not completely understood topic of interdependence. It was our desire in this article to present a framework incorporating more of the mathematics for decision making.

Future work needs to be done on this topic. We have presented an idealized mathematical model. If the Teams are not simply multi-agents systems, but rather human, or human-machine hybrid teams, our model must be tempered by human factors. Humans do not act as automatons. These ideas are discussed in detail in the beautiful books by Kahneman [6, 7], and also [21]. A good overview of Kahneman's Nobel prize work in behavioral economics can be found in [5]. In fact, from [7] we can take the concept of noise and view that in terms of diffusion. As we rely more and more on hybrid teams, we must factor in a behavioral economics type approach to Team decision making. How this relates to the mathematics holds promise as a new research area. Furthermore, Teams are often subject to the wisdom of crowds [21], or the stupidity of crowds [19] (this work involves ants, which might be better representative agents than humans when attempting to model a machine), and the mathematical model we have presented does not incorporate such human factors. Of course, future work could include looking at and measuring these factors for an actual Team/MAS.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

We thank Bill Lawless, Ruth Heilizer and the reviewers for their efforts. We are grateful to a reviewer and Ciara Sibley for pointing out, and discussing, the relevant research of Daniel Kahneman.

REFERENCES

1. National Research Council. Enhancing the Effectiveness of Team Science. In: Cooke NJ Hilton ML, editors. *Committee on the Science of Team Science; Board on Behavioral, Cognitive, and Sensory Sciences; Division of Behavioral and Social Sciences and Education; National Research Council*. Washington, DC: The National Academies Press (2015).
2. Springer. *Autonomous Agents and Multi-Agent Systems*. Berlin, Germany: Springer Journal (1998-2021).
3. Feller W. *An Introduction to Probability Theory and its Applications, Vols.1&2*. Hoboken, NJ, USA: Wiley (1950/1968).
4. Einstein A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann Phys* (1905) 322:549–60. doi:10.1002/andp.19053220806
5. Kahneman D. Maps of Bounded Rationality: Psychology for Behavioral Economics. *Am Econ Rev* (2003) 3(5):14491475. doi:10.1257/000282803322655392
6. Kahneman D. *Thinking, Fast and Slow*. New York, NY, USA: Farrar, Straus and Giroux (2013).
7. Kahneman D, Olivier S, Cass S. *Noise: A Flaw in Human Judgment*. New York City: Little, Brown Spark (2021).
8. Lalley S, Mykland P. *Lecture Note Statistics 313: Stochastic Processes II*. New York City: Spring (2013). Available from: https://galton.uchicago.edu/lalley/Courses/313/. (Accessed October 13, 2021).
9. Lawless WF. The Entangled Nature of Interdependence. Bistability, Irreproducibility and Uncertainty. *J Math Psychol* (2017) 78:51–64. doi:10.1016/j.jmp.2016.11.001
10. Lawless WF. The Physics of Teams: Interdependence, Measurable Entropy and Computational Emotion. *Front Phys* (2017) 2017. doi:10.3389/fphy.2017.00030
11. Lawless WF. Quantum-Like Interdependence Theory Advances Autonomous Human-Machine Teams (A-HMTs). *Entropy* (2020) 22(11):1227. doi:10.3390/e22111227
12. Moskowitz IS, Lawless W, Hyden P, Mittu R, Russell S. *A Network Science Approach to Entropy and Training*. Palo Alto, California, U.S.: AAAI Spring Symposia Series, AAAI Press, 2015.
13. Moskowitz IS. Agent Team Action, Brownian Motion and Gambler's Ruin. *Eng Artificially Intell Syst* (2021) 2021. doi:10.1007/978-3-030-89385-9_6
14. Moskowitz IS, Russell S, Lawless W. An Information Geometric Look at the Valuing of Information. In *Chapter 9. Human-Machine Shared Contexts*. Editors Lawless W. Amsterdam, Netherlands: Elsevier (2020). doi:10.1016/b978-0-12-820543-3.00009-2
15. Moskowitz IS, Brown NL, Goldstein Z. A Fractional Brownian Motion Approach to Psychological and Team Diffusion Problems. In: Lawless W, et al, editors. *Ch. 11, Systems Engineering and Artificial Intelligence*. Berlin, Germany: Springer (2021). doi:10.1007/978-3-030-77283-3_11
16. Moskowitz IS. Agent Team Action, Brownian Motion and Gambler's Ruin. Chaper 11. In: Lawless W, et al, editors. *Lecture Notes in Computer Science (LNCS)*. Berlin, Germany: Springer (2021).
17. Ratcliff R. A Theory of Memory Retrieval. *Psychol Rev* (1978) 85(2):59–108.
18. Ratcliff R, Smith PL, Brown SD, McKoon G. Diffusion Decision Model: Current Issues and History. *Trends Cogn Sci* (2016) 20(4):260–81. doi:10.1016/j.tics.2016.01.007
19. Sasaki T, Granovskiy B, Mann RP, Sumpter DJ, Pratt SC, Pratta SC. Ant Colonies Outperform Individuals when a Sensory Discrimination Task Is Difficult but Not when it Is Easy. *Proc Natl Acad Sci U S A* (2013) 110:13769–73. doi:10.1073/pnas.1304917110
20. Spivak M. *A Comprehensive Introduction to Differential Geometry. Volumes 1-5*. 3rd ed. Los Angeles: Publish or Perish (1999).
21. Surowiecki J. *The Wisdom of Crowds*. New York City: Anchor (2005).

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.