When the Grass Is Greener

Three-Shot Search on Exponentiated Gaussian Landscapes

Peter Cotton

First version: April 6, 2022. This version: September 8, 2026.

Abstract

How should a firm spend its last pivot? A searcher knows an incumbent strategy's performance, may run one more trial, and must then commit, on a landscape given by the exponential of a stationary Ornstein–Uhlenbeck path. We solve the final placement exactly and give a closed-form phase diagram for the case of equal observations: abandon a below-median incumbent; after a disappointing trial revert to the better position's own two-shot rule; commit between two good positions; and stay put once a position is strong enough. The commitment between positions rests on a constant variance-ratio identity: every interior position has the same conditional mean log-performance as an explicit exterior alternative and $(1+\rho)/(1-\rho)$ times its conditional variance. Terminal placement is a projection of the observed values onto a feasible set of correlations, so the geometry solves in every Euclidean dimension. The small-incumbent opening constant rises from 0.540 on the line to a Lambert-W value in the plane, while the large-incumbent behaviour holds across non-Markov landscapes whose log-correlation is a metric.

1 Introduction

How many radical pivots does a firm make? Across four randomized trials covering 759 firms, 59% never pivoted radically during the study, 24% pivoted exactly once, and 17% pivoted more than once (Camuffo et al. 2024; the shares in the smaller 2020 trial are 74%, 19%, and 7%). The data were collected over windows of up to sixteen months, nine for the largest trial, so these are pivot rates over the study rather than a lifetime total.[1] Training founders to search more scientifically raised the frequency of pivoting once or twice and lowered the frequency of pivoting more than twice. Radical pivots are rare either way, which is what motivates a small-budget model, and the interesting version of the problem is farsighted rather than reactive.

Callander (2011) modeled this kind of search as trial and error on an unknown function drawn from a Brownian motion. A sequence of short-lived agents each implements one trial and observes its outcome, so search is myopic by demographic structure rather than by assumption. Most later work keeps this feature. Bardhi and Callander (2026, §3.1.3) survey this line: two-period foresight is understood, while farsighted search by an infinitely lived agent with a free choice of location is their largely open problem. The smallest instance of it is a single searcher

[1] A widely repeated figure holds that a successful startup pivots about 2.5 times. We have not been able to trace it to a primary source, so we lean on the randomized-trial counts above rather than on the round number.

who observes an incumbent, places one informative trial while already anticipating where she will commit, and then commits.

That same horizon is where the optimization literature stands on firm ground. With one informative trial before a terminal recommendation, the knowledge-gradient policy is optimal by construction (Frazier, Powell and Dayanik 2009); on general priors it is evaluated numerically, and on our landscape we solve it — the terminal placement in closed form, the opening trial by one-dimensional quadrature and optimization. The landscape is the exponential of a stationary Ornstein–Uhlenbeck path: mean reversion means performance far from known territory regresses to the industry average, and the exponential means small differences in fit produce large differences in sales, so an untried strategy carries extra value on average. The exponential is a production or revenue mapping from log-performance to money, not a convex utility over a fixed prize, and the searcher is risk neutral in dollars.

The searcher knows the performance of an incumbent strategy. She may try one more. Then she must commit. With two evaluations the optimal rule is a threshold rule: abandon a below-median incumbent, keep an exceptional one, and otherwise step a distance that we give in closed form.

With three evaluations the question is where to commit after the trial: between the two known strategies, beyond the better one, or at one of them. We compute the answer exactly. Commitment between the two is optimal when both values are good but not exceptional. The commitment point is the midpoint for moderate values and moves toward the better-known end as the values improve. Above an exact ceiling the searcher should stay where she is.

The interior option is worth about one percent of expected value at the optimum at this scale. Most of the value comes from the information the trial produces, not from the final placement. On landscapes that decorrelate quickly, it is better to probe far away, and the formulas say so: the interior premium is first order in the bracket correlation.

Gaussian conditioning alone turns the problem into a projection: terminal placement projects the observed values onto a feasible set of correlations, three-point covariance geometry and nothing more. The exponential Euclidean kernel then gives an exact spatial solution in every dimension, with a planar opening constant that differs from the line's. A weaker condition, that the negative log-correlations obey the triangle inequalities, gives high-incumbent universality across landscapes that need not be Markov. Small-incumbent placement is sensitive to the feasible geometry; large-incumbent behaviour is shared.

2 Related work

Callander (2008) introduced the Brownian policy landscape. In that paper, delegation to a biased expert survives only when the bias is below $\sigma^2/2\mu$. The 2011 papers use the landscape for search. An agent experiments when the incumbent outcome is worse than $\alpha = \sigma^2/2|\mu|$ and stops otherwise. Search ends almost surely, often at a mediocre policy. Simulated runs change policy between one and four times, which is consistent with the pivot counts above.

Later papers derive similar thresholds for precedent (Callander and Clark 2017), agenda control (Callander and McCarty 2024), and risk-averse learning (Callander and Matouschek 2019). In all of these the number of evaluations is endogenous and payoffs accrue every period. In our problem the number of evaluations is fixed at three, only the final placement is paid, and the landscape mean-reverts.

Callander and Hummel (2014) analyze a two-period game in which each of two policymakers can experiment once. The equilibrium has two exact cutoffs, and the first mover sometimes experiments although the status quo already delivers his ideal outcome. The players are strategic and payoffs accrue in both periods. Ganz (2020) studies an organization that allows a manager one offline trial before an online choice. That is a two-evaluation problem in which only the second evaluation is paid, and some of its boundaries are computed numerically. We solve the three-evaluation single-agent problem: the final placement exactly and, for equal observed values, with every phase boundary in closed form; the opening trial reduces to a one-dimensional numerical optimization.

Other papers obtain farsighted results under other restrictions. Urgun and Yariv (2025) require search to be contiguous; the searcher chooses speed and stops at a closed-form drawdown. Cetemen, Urgun and Yariv (2023) extend this to teams and Wong (2025) to flow payoffs. Bardhi (2024) chooses k sample locations at once to estimate a project. Garfagnini and Strulovici (2016), Carnehl and Schneider (2025), and Callander, Lambert and Matouschek (forthcoming) study infinite sequences of short-lived agents.

Aybas and Callander (2025) allow Ito diffusion landscapes, including the Ornstein–Uhlenbeck case, in a one-shot cheap-talk game. That is the only mean-reverting landscape we know of in this literature. Bardhi (2024) also proves that the OU kernel is the only powered-exponential covariance with the Markov property, which supports our choice of landscape.

In the optimization literature the same object appears with large budgets. Calvin (1997) derives average-case convergence rates under Wiener measure. Grill, Valko and Munos (2018) approximate the maximum of a Brownian path. Grosse, Zhang and Hennig (2023) extend this to Gaussian-process samples and prove regret rates. Alabert and Caballero (2018) compute the law of the minimum of a conditioned bridge in order to compare fixed sampling designs.

Malladi (2025) and Banchio and Malladi (2025) solve worst-case versions with no prior and a Lipschitz bound in place of volatility. Hodgson and Lewis (2025) estimate a Gaussian-process search model on clickstream data and solve the policy numerically. The probabilistic line search of Mahsereci and Hennig (2017) is the nearest algorithmic relative: a Gaussian-process line search for stochastic gradient descent, built on gradient observations and an integrated-Wiener model, without closed-form placement. We are not aware of a closed-form solution for a small fixed budget of adaptive, derivative-free evaluations under a Bayesian prior.

The closest methodological neighbor is non-myopic Bayesian optimization. Ginsbourger and Le Riche (2010) observed that the usual one-step expected-improvement rule is suboptimal under a genuine multi-evaluation budget, since it treats each evaluation as the last. Lam, Willcox and Wolpert (2016) cast finite-budget optimization as a dynamic program and solve it approximately by rollout, Wu and Frazier (2019) implement a practical two-step lookahead, and Jiang et al. (2020) optimize the full multi-step lookahead tree in one shot. These methods approximate the lookahead dynamic program on general Gaussian-process priors. Our horizon is smaller and cleaner: with one informative trial before a terminal recommendation, the knowledge-gradient policy of Frazier, Powell and Dayanik (2009) is optimal by construction, and we characterize it exactly on the OU landscape rather than approximately.

The rule is itself an acquisition function, and it can be located using two familiar distinctions in Bayesian optimization (Shahriari et al. 2016). One is myopic versus lookahead: expected improvement (Mockus et al. 1978; Jones et al. 1998), probability of improvement (Kushner 1964), the upper confidence bound (Srinivas et al. 2010), and the entropy-search family (Hennig and Schuler 2012; Wang and Jegelka 2017) are one-step rules, while multi-step-optimal

acquisitions exist only in approximate form. The other is cumulative-regret versus terminal-recommendation payoff: the upper confidence bound bounds cumulative regret, knowledge gradient scores the payoff of a finally reported point, and entropy search is information-directed toward the terminal optimizer. Our rule is lookahead and terminal-recommendation, the knowledge-gradient corner.

Within that corner our contribution is not a new objective but an exact characterization. Because the horizon carries a single informative trial, knowledge gradient is optimal here by construction, so the closed forms describe the optimal policy rather than an approximation to it: the final placement is algebraic, quartic in the unequal case, and the opening trial reduces to one-dimensional quadrature and optimization. The roughly one-percent figure in Section 9 is a separate quantity: it isolates the incremental value of permitting interior terminal recommendations relative to a restricted but still farsighted policy.

Glick and Myers (2015) tested the one-evaluation problem on lab subjects. Subjects respond to the Brownian structure but modify too much when the variance is high.

3 The model: three evaluations of an exponentiated OU path

Let X be a stationary Ornstein–Uhlenbeck process on $\mathbb{R}$ with $\mathbb{E}[X_t] = 0$, unit marginal variance, and covariance kernel

$$k(s,t) \;=\; e^{-\kappa|s-t|}. \tag{1}$$

Exponential decay of correlation with distance is a standard empirical model, not an assumption special to this paper: it has described shadow fading in radio engineering since Gudmundson (1991). On 77,040 WiFi measurements across a university floor (Feng, Nguyen and Luo 2024), we find the detrended log signal close to Gaussian with spatial correlation fitting an exponential at $R^2 = 0.98$, alongside a short-scale component of comparable size; the scripts are in the paper's repository.

The same shape is standard in several other fields. In Gaussian-process modeling the exponential kernel is the Matérn class at $\nu = 1/2$ and is the covariance of the OU process (Rasmussen and Williams 2006, §4.2.1). Practical Bayesian optimization prefers Matérn kernels because squared-exponential sample paths are too smooth for real objectives (Snoek, Larochelle and Adams 2012).

Weinberger (1990) measures the ruggedness of a fitness landscape by the exponential decay rate of a random walk's autocorrelation and shows that Kauffman's NK model is generically of this type. In evolutionary biology the OU process is the standard model of a trait held near an adaptive optimum (Hansen 1997; Butler and King 2004). The landscape is $f(t) = \exp(X_t)$: locally continuous, mean-reverting in its logarithm, with occasional high peaks where excursions are magnified by the exponential. A searcher samples f at a first location, which we take to be $t_0 = 0$ by translation invariance, discovering $X_0 = b$. Two further evaluations at locations t_1 and then t_2 follow, each chosen with knowledge of all values so far, and the payoff is the value at the final location, $u = \mathbb{E}\left[f(t_2)\right]$. Conditional on the observations gathered by the time t_2 is chosen, X_{t_2} is Gaussian with some mean μ and variance ν, and

$$\log \mathbb{E}\big[e^{X_{t_2}} \,\big|\, \mathcal{F}\big] \;=\; \mu + \tfrac{1}{2}\nu, \tag{2}$$

where $\mathcal{F}$ denotes the observed locations and values. (The unconditional distribution after adaptive placement need not be Gaussian; every use of (2) below is conditional.) The searcher

is thus paid for conditional mean and conditional variance in equal measure: a final location with residual uncertainty carries a lognormal bonus of half its variance.

The final evaluation is a recommendation, not a measurement. Write $R(\mathcal{D}) = \sup_t \mathbb{E}[e^{X_t} \mid \mathcal{D}]$ for the best expected payoff attainable from the information $\mathcal{D}$. The opening trial at t_1 maximizes $\mathbb{E}\big[R\big(\mathcal{D} \cup \{(t_1, X_{t_1})\}\big) \mid \mathcal{D}\big]$, and subtracting the action-independent $R(\mathcal{D})$ leaves the knowledge-gradient acquisition (Frazier, Powell and Dayanik 2009). Because exactly one informative trial precedes the terminal recommendation, this one-step rule is optimal by construction: knowledge gradient is the optimal policy whenever a single measurement remains. What we add is not a new acquisition but its exact solution on this landscape — the terminal-placement geometry and phase diagram in closed form, and an algebraic inner optimization that reduces the opening knowledge-gradient calculation to a one-dimensional numerical problem.

Figure 1 shows the decision this paper solves.

The problem has one essential scale parameter. A payoff exponent β alone is without loss of generality, since $e^{\beta X}$ on a unit-variance path equals e^{Y} on a path of standard deviation β; only the product $s = \beta\sigma$ of payoff convexity and path scale matters. We fix $s = 1$ as an illustrative scale; no empirical calibration is claimed.

For general s every comparison maximizes $\mu + s\nu/2$ with observations in path-standard-deviation units, the two-shot step becomes $\lambda^* = \mathrm{clip}(b/s, 0, 1)$, and the pitchfork and ceiling of Section 7 scale to $s \cdot 2\sqrt{\rho}/(1-\rho)$ and $s(1+\rho)/(1-\rho)$; equivalently, the phase diagram is universal in the coordinates $(\rho, b/s)$. The length scale κ, by contrast, is fully absorbed by working in correlations. The parameter plays the role Callander's complexity $\sigma^2/2|\mu|$ plays under drift, and the one-percent figure of Section 9 is specific to this illustrative $s = 1$.

Two conventions tidy the boundary cases. A searcher may take a location to infinity, which by mean reversion is equivalent to sampling a fresh $N(0,1)$ value independent of everything seen; and a searcher may finish at an observed location, receiving its known value. Neither changes the substance of any result.

4 Two evaluations: flee, explore, or stay

Given only $X_0 = b$, a candidate second location at distance t has $X_t \sim N\big(b\lambda,\, 1-\lambda^2\big)$ with $\lambda = e^{-\kappa t} \in (0, 1]$. From (2),

$$\log \mathbb{E}\big[e^{X_t} \,\big|\, X_0 = b\big] \;=\; b\lambda + \tfrac{1}{2}\big(1-\lambda^2\big) \;=\; -\tfrac{1}{2}\,(\lambda - b)^2 + \tfrac{1}{2}\,\big(b^2+1\big)\,, \tag{3}$$

maximized over $\lambda \in [0,1]$ at $\lambda^* = \min\{\max\{b, 0\}, 1\}$.

Proposition 1 (Two-shot policy). The optimal second location and its log-value $\zeta(b)$ are

$$t^* \;=\; \begin{cases} \infty & b < 0 \quad \text{(flee)}, \\ -\frac{1}{\kappa}\log b & 0 \le b \le 1 \quad \text{(explore)}, \\ 0 & b > 1 \quad \text{(stay)}, \end{cases} \qquad \zeta(b) \;=\; \begin{cases} \frac{1}{2} & b < 0, \\ \frac{1}{2}\,\big(b^2+1\big) & 0 \le b \le 1, \\ b & b > 1. \end{cases} \tag{4}$$

A below-median first value is abandoned, and an exceptional one is kept: already at two shots, an observed location can be the optimal final location. Practitioners know this fork as pivot or persevere (Ries 2011); the proposition supplies its thresholds and the size of the

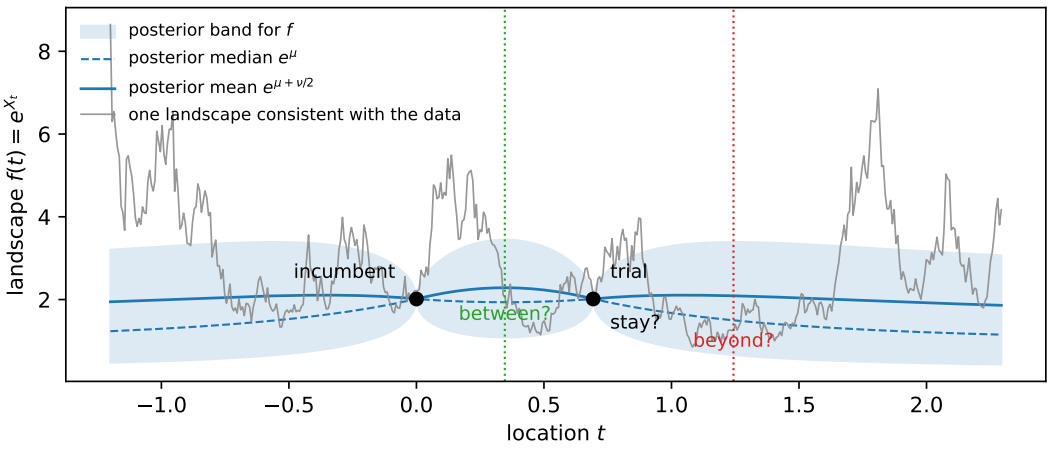

Figure 1: The problem. Two locations have been evaluated and returned equal values (dots); the shaded region is the posterior band, the dashed curve the posterior median e^{μ} and the solid curve the posterior mean $e^{\mu+\nu/2}$. The final evaluation is the consequential one: settle between the anchors, step out past an end, or stay at an observed point. The paper computes which, exactly, as a function of the anchor value and the bracket correlation.

step when pivoting wins. In between, the searcher steps to the distance at which the prior pull toward the mean exactly balances the anchor's height. The value ζ serves as the exterior benchmark throughout: by the Markov property, any placement beyond all existing observations faces a fresh copy of this problem anchored at the nearest observation.

5 Bridge moments, and what composes how

Suppose values $X_0 = b$ and $X_{t_1} = d$ are known, with bracket correlation $\rho = e^{-\kappa t_1}$. For an interior location let $\lambda_2 = e^{-\kappa t_2}$ and $\lambda = e^{-\kappa(t_1 - t_2)}$ be the correlations to the two ends, $\lambda_2 \lambda = \rho$, with rapidities $\theta_2 = \operatorname{artanh} \lambda_2$ and $\theta = \operatorname{artanh} \lambda$. Gaussian conditioning gives the bridge moments

$$\mu \;=\; \frac{\lambda_2\big(1-\lambda^2\big)\, b + \lambda\big(1-\lambda_2^2\big)\, d}{1-\rho^2}, \qquad \nu \;=\; \frac{\big(1-\lambda_2^2\big)\,\big(1-\lambda^2\big)}{1-\rho^2}. \tag{5}$$

For $b = d > 0$ the bridge sags: the mean is below b everywhere strictly inside, pulled toward the process mean, while the variance rises from zero at the ends to a maximum at the middle.

A caution on what is additive. Correlations across consecutive intervals multiply, $\rho(t+s) = \rho(t)\,\rho(s)$, so the additive coordinate for composing intervals is $\kappa t = -\log \rho$, not the rapidity. The rapidity enters through a different identity: for equal anchors $d = b$ the bridge mean is

$$\mu \;=\; b\,\frac{\lambda_2 + \lambda}{1 + \lambda_2 \lambda} \;=\; b \tanh(\theta_2 + \theta)\,, \tag{6}$$

which matches the mean of the exterior point at correlation $\tanh(\theta_2 + \theta)$ to its anchor. This is a mean-matching device we use next, not a composition law. The bridge variance also takes

a hyperbolic product form: substituting $1 - \tanh^2 = \operatorname{sech}^2$ into (5),

$$\nu \;=\; \operatorname{sech}(\theta_2 + \theta)\operatorname{sech}(\theta_2 - \theta)\,, \tag{7}$$

while an exterior point at correlation $\lambda' = \tanh\theta'$ to its anchor has variance $1 - \tanh^2\theta' = \operatorname{sech}^2\theta'$.

6 Matched-mean dominance at a constant variance ratio

Theorem 1 (Bridge dominance, equal anchors). Let $d = b$ and let $t_2 \in (0, t_1)$ be an interior location with rapidities (θ_2, θ). The exterior location at composed rapidity $\theta' = \theta_2 + \theta$ has the same conditional mean $b\tanh(\theta_2 + \theta)$, and the ratio of conditional variances is the same at every interior point:

$$\frac{\nu_{\text{inside}}}{\nu_{\text{outside}}} \;=\; \frac{\cosh(\theta_2 + \theta)}{\cosh(\theta_2 - \theta)} \;=\; \frac{1 + \tanh\theta_2\tanh\theta}{1 - \tanh\theta_2\tanh\theta} \;=\; \frac{1+\rho}{1-\rho} \;>\; 1. \tag{8}$$

Consequently the interior point's conditional distribution dominates its mean-matched exterior rival in the Gaussian convex order, and its expected payoff is at least as large for every convex payoff with finite expectation, strictly so for the exponential.

Proof. The mean statement is (6). The ratio of (7) to $\operatorname{sech}^2(\theta_2+\theta)$ is $\cosh(\theta_2+\theta)/\cosh(\theta_2-\theta)$, and expanding both via $\cosh(A\pm B) = \cosh A\cosh B \pm \sinh A\sinh B$ gives $(1+\tanh\theta_2\tanh\theta)/(1-\tanh\theta_2\tanh\theta)$, which is $(1+\rho)/(1-\rho)$ because $\tanh\theta_2\tanh\theta = \lambda_2\lambda = \rho$ on the bracket. Two Gaussians with equal means and ordered variances are ordered in the convex order (the wider equals the narrower plus independent noise), and (2) is the exponential case. □

The amplification is a property of the bracket, not the position: threefold at $\rho = \frac{1}{2}$, nineteenfold at $\rho = 0.9$, unbounded as $\rho \uparrow 1$. At the bracket's ends both variances vanish; their ratio still tends to $(1+\rho)/(1-\rho)$. Dominance over mean-matched exterior rivals, which the theorem explicitly constructs, decides nothing by itself: the exterior player may choose any correlation, and an observed anchor may simply be kept. The object that can fail to exist is different: an *unvisited* point whose conditional mean matches the best *observed* value. Mean reversion caps unvisited conditional means strictly below a high anchor, and whether the variance bonus makes up the shortfall is precisely the question the phase diagram answers. The next section makes this exact.

What the theorem uses. The bracket entered only through the chaining $\lambda_2\lambda = \rho$: the interior point's correlations to the two ends multiply to the end-to-end correlation. This is the Gauss–Markov screening property, and it does the work by itself. It guarantees that the mean-matched exterior twin exists — its correlation $(\lambda_2 + \lambda)/(1+\rho) = \tanh(\theta_2 + \theta)$ is below one because $(1-\lambda_2)(1-\lambda) \geq 0$ — and fixes the ratio at $(1+\rho)/(1-\rho)$ at every interior position. Without chaining the construction can fail outright: under a squared-exponential covariance the interior conditional mean can exceed every exterior mean, so no mean-matched twin exists and no constant ratio holds.

The bridge identities are therefore shared by any field whose three evaluated points chain in this way. Chaining alone does not transfer the whole problem, however: it fixes the conditional means and variances, but not which exterior locations exist or whether extra placements are

available. A Gaussian Markov field on a branching tree, for one, offers terminal locations off the observed line and can pose a different problem at the same marginal variance. The transfer condition is that the available three-point covariance configurations agree with those of the unrestricted OU line — the feasible set made explicit in Section 10.

The Ornstein–Uhlenbeck process is the stationary representative, the unique stationary Markov powered-exponential covariance (Bardhi 2024); its marginal $N(0, \sigma^2)$ sets the exterior benchmark ζ and the reading of mean reversion — the levels of the phase diagram, not its shape.

7 The constrained parabola and the complete symmetric solution

Parametrize the interior by $x = \lambda_2 \in (\rho, 1)$ and set

$$C \;=\; \frac{1+\rho}{1-\rho}, \qquad m \;=\; \frac{x+\rho/x}{1+\rho} \;\in\; \left[\frac{2\sqrt{\rho}}{1+\rho},\, 1\right), \tag{9}$$

under which mirror points $x \leftrightarrow \rho/x$ coincide, the lower endpoint of m is the bracket middle, and, using $\left(1-x^2\right)\left(1-\rho^2/x^2\right) = (1+\rho)^2\left(1-m^2\right)$, the equal-anchor interior log-utility becomes a constrained parabola:

$$L_{\text{inside}}(m) \;=\; b\,m + \frac{C}{2}\left(1-m^2\right), \qquad m^* \;=\; \operatorname{clip}\!\left(\frac{b}{C};\; \frac{2\sqrt{\rho}}{1+\rho},\, 1\right). \tag{10}$$

This is the organizing object of the paper: the entire symmetric analysis reads off the position of the unconstrained peak b/C relative to the interval.

Theorem 2 (Pitchfork). The bracket middle is the interior optimum iff $b/C \le 2\sqrt{\rho}/(1+\rho)$, i.e. $b \le b_{\text{p}}$, writing

$$b_{\text{p}} \;=\; \frac{2\sqrt{\rho}}{1-\rho} \tag{11}$$

for the pitchfork threshold. For $b_{\text{p}} < b < C$ the peak is interior and attained by the mirror pair

$$x_{\pm} \;=\; \frac{b\,(1-\rho) \pm \sqrt{b^2\,(1-\rho)^2 - 4\rho}}{2}, \qquad x_+ x_- = \rho, \tag{12}$$

with optimal value

$$U_{\text{off}} \;=\; \frac{C}{2} + \frac{b^2}{2C}. \tag{13}$$

For $b \ge C$ the peak $b/C \ge 1$ lies at or beyond the interval's open end: L_{inside} is increasing in m, no interior maximum is attained, and the interior supremum is the end limit b.

Proof. $dL/dm = b - Cm$ vanishes at $m = b/C$; the interval endpoints translate via (9) into the stated b-thresholds ($b/C = 2\sqrt{\rho}/(1+\rho)$ is $b = 2\sqrt{\rho}/(1-\rho)$, and $b/C = 1$ is $b = (1+\rho)/(1-\rho)$). Solving $x + \rho/x = (1+\rho)\,b/C = b\,(1-\rho)$ gives (12), and substituting $m = b/C$ into (10) gives (13). □

Theorem 3 (Symmetric phase diagram). For $d = b \geq 0$ the best interior point strictly beats every alternative — exterior placements and observed anchors alike — precisely for

$$b_-(\rho) \;<\; b \;<\; C, \qquad b_-(\rho) = \frac{\sqrt{\rho}\left(2 - \sqrt{2\,(1-\rho)}\right)}{1+\rho}, \tag{14}$$

the lower edge being the smaller root of $b^2\,(1+\rho) - 4b\sqrt{\rho} + 2\rho = 0$. For $b \geq C$ the optimal final location is an *observed* one: staying at an anchor beats every unvisited point, interior or exterior.

Proof. If $b \geq C$, Theorem 2 gives interior supremum b, unattained, while every exterior point loses to the anchor pointwise: for $b > 1$ and any $\lambda \in [0, 1)$,

$$b - \left[b\lambda + \tfrac{1}{2}\left(1 - \lambda^2\right)\right] \;=\; (1-\lambda)\left[b - \tfrac{1+\lambda}{2}\right] \;>\; 0 \tag{15}$$

(the suprema coincide at b; the dominance is strict at every unvisited point): staying at an anchor wins. If $b_{\mathrm{p}} < b < C$ the interior optimum is U_{off}, and both comparisons are one line: for $b > 1$,

$$U_{\mathrm{off}} - b \;=\; \tfrac{1}{2}\,C^{-1}(b-C)^2 \;>\; 0, \tag{16}$$

the arithmetic–geometric mean inequality applied to the two terms of (13), with equality only at the tangency $b = C$ defining the upper edge; for $b \leq 1$,

$$U_{\mathrm{off}} - \tfrac{1}{2}\left(b^2+1\right) \;=\; \tfrac{1}{2}(C-1)\left(1 - b^2 C^{-1}\right) \;>\; 0. \tag{17}$$

If $b \leq b_{\mathrm{p}}$ the optimum is the middle, with value $U_{\mathrm{mid}} = b\,\frac{2\sqrt{\rho}}{1+\rho} + \frac{1-\rho}{2(1+\rho)}$; for $b \leq 1$, $U_{\mathrm{mid}} > \zeta(b)$ is the stated quadratic, and $b_- \leq 2\sqrt{\rho}/\,(1+\rho) < b_{\mathrm{p}}$ places the lower edge inside this regime. No gap opens above the middle regime: when $b_{\mathrm{p}} \leq 1$ (i.e. $\rho \leq 3 - 2\sqrt{2}$), the larger quadratic root satisfies $b_+ \geq b_{\mathrm{p}}$, equivalent to $(1-\rho)^3 \geq 8\rho^2$, which holds on that range (left side decreasing, right increasing, valid at the endpoint); when $b_{\mathrm{p}} > 1$, $b_+ \geq 1$ reduces, with $u = \sqrt{\rho}$, to $2u^2\,(1+u) \geq (1-u)^3$, true at $u = \sqrt{2} - 1$ (it reads $20\sqrt{2} \geq 28$) and increasingly so; and on $1 < b \leq b_{\mathrm{p}}$, $U_{\mathrm{mid}} \geq b$ reduces to $b \leq (1+u)\,/\,(2\,(1-u))$, which covers the stretch because $b_{\mathrm{p}} \leq (1+u)\,/\,(2\,(1-u))$ is $(1-u)^2 \geq 0$. □

Along the two-shot-inherited bracket $\rho = b$ the region opens at $b^* = 0.4233\ldots$, the root of $U_{\mathrm{mid}}(b, b) = \zeta(b)$. The geometry of the upper edge is easy to misread: for $b > 1$, interior dominance requires $\rho > (b-1)\,/\,(b+1)$, so higher anchors demand a larger bracket correlation, which is a *shorter* physical bracket. Very good news is forgiven only by narrow brackets; on a wide bracket, very good news means stay.

Corollary 1 (Below-median anchors). For equal anchors $d = b < 0$ every interior point is dominated by flight, whose log-value is $\frac{1}{2}$.

Proof. An interior point's conditional mean is a nonnegative combination of the two negative anchors with positive weight on at least one, so $\mu < 0$, and its conditional variance is $\nu < 1$; hence $\mu + \nu/2 < \frac{1}{2} = \zeta(b)$. □

With Theorem 3 this settles the symmetric final-placement game for every real anchor. Throughout, $\rho \in (0, 1)$. As $\rho \downarrow 0$ the anchors separate infinitely: a candidate can remain correlated with at most one anchor, while points far from both converge to fresh draws. As $\rho \uparrow 1$ the bracket collapses to a point. Both are degenerate limits rather than cases of the phase diagram.

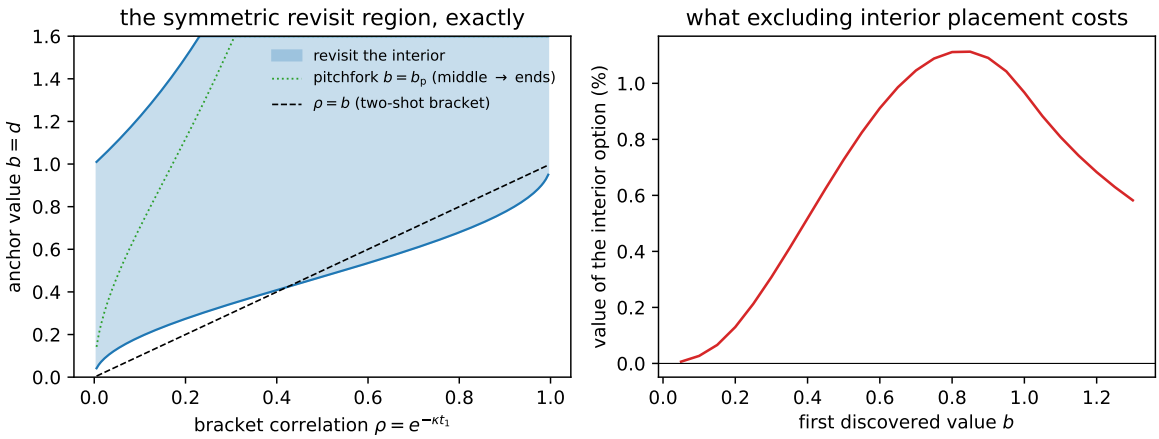

Figure 2: Left: the exact symmetric revisit region (14) in the plane of bracket correlation ρ and anchor value $b = d$, bounded below by the quadratic root b_- and above by C. The dotted curve is the pitchfork $b = b_p$: below it the optimal interior point is the bracket middle, above it the mirror pair (12) hugging the ends. The dashed line is the two-shot bracket $\rho = b$, entering the region at $b^* = 0.4233$; the upper boundary continues beyond the plotted range. Right: the value of the interior option — the percentage increase in expected payoff of the full three-shot optimum over the best policy denied interior placements (it may still place on either exterior ray) — as a function of the first discovered value b, with the bracket optimized in both cases.

8 Beyond equal anchors

For general (b, d) the interior stationary points solve the quartic

$$x^4 - (b - \rho d)\, x^3 + \rho\,(d - \rho b)\, x - \rho^2 \;=\; 0, \tag{18}$$

obtained by differentiating (5); its real roots in $(\rho, 1)$, compared against $\zeta(b)$ and $\zeta(d)$, decide the move exactly and replace any grid. The symmetric window does not extend indefinitely in d. At $b = \frac{1}{2}$, $\rho = \frac{1}{2}$, the interior wins precisely for $d \in (\approx 0.4715,\, 1)$: below, the ray past the better end wins; at $d = 1$ every strictly interior point loses, by the identity

$$L(x) - 1 \;=\; -\frac{(2x-1)^2\,(x^2 + x + 1)}{6x^2} \;<\; 0, \qquad \tfrac{1}{2} < x < 1, \tag{19}$$

and for $d \geq 1$ staying at the second anchor is optimal, since d's coefficient in the bridge mean is below one and raising d only widens the loss. Symmetric dominance is a window, not a half-line.

The stopping side of the problem has a closed form for unequal anchors as well.

Proposition 2 (Unequal-anchor stopping). Let $B = \max(b, d)$ and $D = \min(b, d)$. For the unrestricted final placement, staying at the better observed anchor B is optimal if and only if

$$B \geq 1 \qquad \text{and} \qquad \left(1 + \rho^2\right) B - 2\rho D \;\geq\; 1 - \rho^2. \tag{20}$$

For general scale s the conditions read $B \geq s$ and $\left(1 + \rho^2\right) B - 2\rho D \geq s\left(1 - \rho^2\right)$.

Proof. Orient the bracket so B is at the near end, write $C = (1+\rho)/(1-\rho)$ and $m_0 = 2\sqrt{\rho}/(1+\rho)$, and set $a = (B+D)/2$, $\delta = (B-D)/2$. Parametrize an interior point by $x \in [\sqrt{\rho}, 1]$, its correlation to the near end B, and write $m = (x+\rho/x)/(1+\rho) \in [m_0, 1]$ as in Section 7 (at the midpoint $x = \sqrt{\rho}$ gives $m = m_0$). The half nearer B carries at least the mean of its mirror image and the same variance, with log-value

$$F(m) = am + C\delta\sqrt{m^2 - m_0^2} + \tfrac{C}{2}\left(1 - m^2\right), \qquad m_0 \le m \le 1. \tag{21}$$

Since $F''(m) = -C - C\delta\, m_0^2\left(m^2 - m_0^2\right)^{-3/2} < 0$, F is concave and is maximized at $m = 1$ — collapsing onto B — exactly when $F'(1) = a + C^2\delta - C \ge 0$; with $\sqrt{1 - m_0^2} = 1/C$ this is the second inequality of (20). The exterior rays are beaten by the anchor precisely when $B \ge 1$, since $\zeta(B) = B$ there while $\zeta(B) > B$ for $B < 1$. The criterion agrees with the quartic (18) on 8,000 random states. □

For equal anchors (20) reduces to $B \ge (1+\rho)/(1-\rho)$, the ceiling of Theorem 3; at $\rho = \frac{1}{2}$, $B = 1$ it gives $D \le \frac{1}{2}$. It supplies the explicit stopping region for unequal observations that the quartic (18) describes only implicitly.

9 The three-shot policy, computed

Writing $V(b, d, \rho)$ for the best of $\zeta(b)$, $\zeta(d)$ and the interior optimum (via (18)), the three-shot value of a bracket ρ is $\log \mathbb{E}_d\left[e^{V(b,d,\rho)}\right]$ with $d \sim N\left(b\rho, 1-\rho^2\right)$, computed by adaptive quadrature and optimized over ρ. The comparison policy denies interior placement but keeps both exterior rays. It is not a ban on reversing direction.

Proposition 3 (Below-median start). For every $b \le 0$ the three-shot value equals $W_3 = \frac{1}{2} + \log\left(1 + 1/\sqrt{2\pi}\right) = 0.83572\ldots$, attained at $\rho = 0$. It is a constant on $b \le 0$, not a limit.

Proof. When the incumbent is nonpositive the abandonment argument (Appendix A) removes every interior recommendation, so the continuation from a revealed d is $\zeta(d)$ and the value is $\log \mathbb{E}_d\left[e^{\zeta(d)}\right]$ with $d \sim N\left(b\rho, 1-\rho^2\right)$. Since e^ζ is convex and nondecreasing, and $b \le 0$ makes the mean $b\rho$ nonincreasing in ρ while the variance $1-\rho^2$ decreases, the fresh draw $\rho = 0$, that is $d \sim N(0,1)$, dominates. Splitting $\mathbb{E}\left[e^{\zeta(Z)}\right]$ over $Z < 0$, $0 \le Z \le 1$, and $Z > 1$ gives $\mathbb{E}\left[e^{\zeta(Z)}\right] = e^{1/2}\left(1 + 1/\sqrt{2\pi}\right) = 2.30647\ldots$, so $W_3 = \log \mathbb{E}\left[e^{\zeta(Z)}\right] = \frac{1}{2} + \log\left(1 + 1/\sqrt{2\pi}\right)$. □

Three observations, all in Table 1 and Figure 2. First, a below-median start still means flight, and the interior option is then worthless; for every $b \le 0$ the three-shot value is exactly $\frac{1}{2} + \log(1 + 1/\sqrt{2\pi}) = 0.8357\ldots$ (Proposition 3). Second, the three-shot bracket is wider than the two-shot step, but most of the widening is *not* bought by the interior option: at $b = 0.5$ the two-shot policy steps to $\rho = 0.5$, excluding interior placement already widens the optimum to $\rho = 0.323$, and admitting it widens further to $\rho = 0.289$. The extra evaluation and the choice it enables account for most of the stretch; the interior option adds a further, smaller pull.

Third, the interior option's worth peaks near 1.1% of expected value around $b \approx 0.83$ and vanishes at both extremes. Excluding interior placement is thus cheap by value; it is the prescription that errs, and for equal anchors it errs exactly on the revisit window (14) — high, but below the equal-anchor ceiling C above which both policies stay.

b	optimal ρ	value	value, interior excluded	interior worth
0.1	0.053	0.8396	0.8393	0.03%
0.3	0.167	0.8723	0.8692	0.31%
0.5	0.289	0.9386	0.9314	0.73%
0.7	0.402	1.0375	1.0271	1.05%
0.83	0.469	1.1183	1.1072	1.12%
0.9	0.502	1.1670	1.1562	1.09%
1.2	0.622	1.4059	1.3991	0.68%

Table 1: The three-shot game: optimal bracket correlation, log-value, the log-value with interior placement excluded (exterior rays in both directions permitted), and the percentage the interior option adds to expected payoff. Adaptive quadrature over the second value d; interior optima exact via the stationarity quartic (18). For $b \le 0$ the optimal bracket correlation is zero and the value equals exactly $\frac{1}{2} + \log\bigl(1 + 1/\sqrt{2\pi}\bigr) = 0.8357\ldots$ (Proposition 3).

For unequal observations the erring region is a joint condition on (b, d, ρ), decided by the quartic (18), not membership of each value separately: at $\rho = \frac{1}{2}$, $b = \frac{1}{2}$, $d = 2$, both values lie in the symmetric window, yet staying at the second anchor is optimal and the exclusion costs nothing.

Proposition 4 (Opening asymptotics). For $s = 1$ the optimal opening correlation obeys $\rho^*(b) = \kappa_0\, b + O(b^3)$ as $b \downarrow 0$, with $\kappa_0 = 0.540062\ldots = 2a_0/(1 + a_0^2)$, where $a_0 \in (0, 1)$ is the unique root of

$$\bigl(1 + a^2\bigr) \log a + 2 - \tfrac{a^2}{2} + \tfrac{a^4}{6} - \sqrt{\tfrac{\pi}{2}}\,(1 - a)^2 = 0. \tag{22}$$

For every $b > 0$ the optimum is interior, $0 < \rho^*(b) < 1$, and $1 - \rho^*(b) \sim c_\infty / b^2$ as $b \to \infty$ with $c_\infty = 0.749096\ldots$

The coefficient is characterized rather than elementary. The opening is thus closed at $b = 0$ (Proposition 3) and asymptotically at both extremes, and numerical for intermediate b; the derivation and checks are in the repository. A costless informative trial is always worth running — even an exceptional incumbent is probed, only increasingly locally — so the stopping decision belongs entirely to the terminal placement, where for equal anchors it is the ceiling of Section 7. This separates acquiring information from acting on it, a distinction the myopic search of Callander (2011) does not draw.

10 Covariance geometry

The one-dimensional results are a projection in disguise. For a centered, unit-variance Gaussian field the joint law of three evaluated values is fixed by their three pairwise correlations, so both the trial reveal and the terminal decision see only this three-point covariance geometry. Write z for the observed values, Σ for their covariance, and r for a candidate's correlation vector with them; then

$$\log \mathbb{E}\bigl[e^{X_{\text{term}}} \mid z\bigr] = \tfrac{1}{2} + \tfrac{1}{2}\, z^\top \Sigma^{-1} z - \tfrac{1}{2}\,(r - z)^\top \Sigma^{-1}\,(r - z)\,. \tag{23}$$

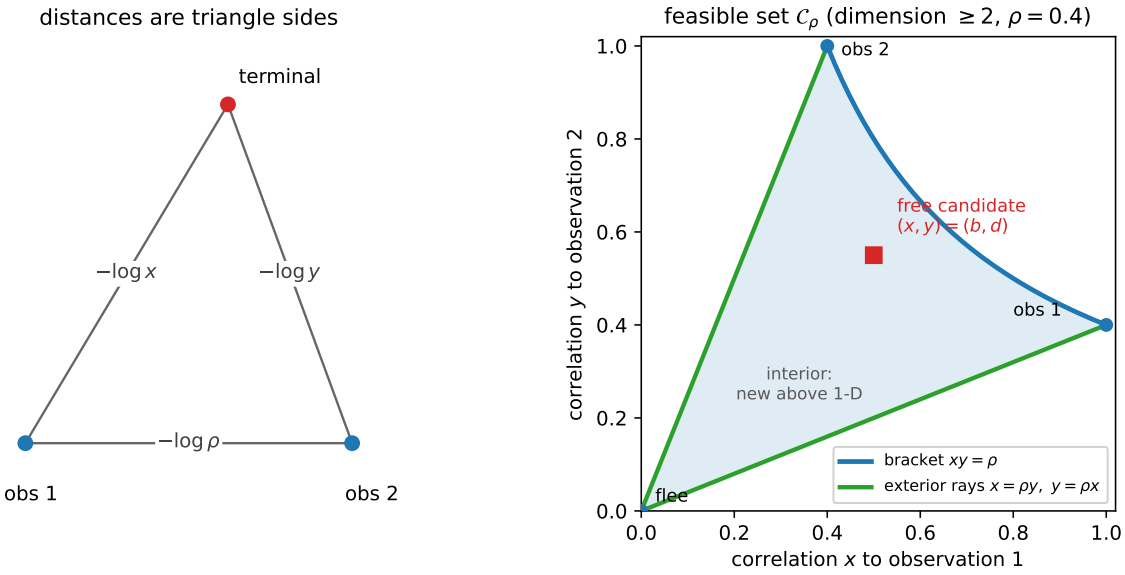

Figure 3: Left: three evaluated points form a Euclidean triangle whose sides are the distances $-\log$ of the pairwise correlations. Right: the feasible terminal correlations $\mathcal{C}_\rho$ in dimension $m \geq 2$ (here $\rho = 0.4$); the one-dimensional placements are the boundary — the bracket arc $xy = \rho$ and the two exterior rays — while the interior, containing the free candidate $(x, y) = (b, d)$, is available only above one dimension.

Terminal placement therefore projects the observed-value vector onto the feasible set of correlation vectors in the Σ^{-1} metric. Two fields with the same feasible correlations pose the same three-shot problem, whether or not one is a reparametrization of the other; the exponentiated OU line is the instance whose feasible correlations are the bracket $\lambda_2\lambda = \rho$ together with the two exterior rays of Sections 5–8.

Fix a radial covariance and allow unrestricted locations. Any three points form a triangle realizable in a plane, and a radial kernel turns edge lengths into pairwise correlations, so the three-shot problem is identical in every Euclidean dimension $m \geq 2$; extra dimensions cannot change a three-point problem. Take the exponential covariance $\mathrm{Cov}(X_s, X_t) = e^{-\|s-t\|}$ on $\mathbb{R}^m$, whose restriction to a line is the OU kernel (Rasmussen and Williams 2006). A terminal point's correlations (x, y) to the two observations have distances $-\log x, -\log y$, so the triangle inequalities against $-\log \rho$ cut out the feasible set exactly,

$$\mathcal{C}_\rho = \{(x, y) : xy \leq \rho,\ x \geq \rho y,\ y \geq \rho x\}\,. \tag{24}$$

Its boundary is precisely the Section-7 bracket ($xy = \rho$) and the two exterior rays ($x = \rho y$, $y = \rho x$), so the one-dimensional placements are the boundary of a two-dimensional region (Figure 3).

By (23) the log-payoff is a strictly concave quadratic in (x, y), so the only candidate off the boundary is its unconstrained maximizer $(x, y) = (b, d)$, with value

$$L_{\text{free}} = \tfrac{1}{2} + \frac{b^2 - 2\rho bd + d^2}{2\,(1 - \rho^2)}. \tag{25}$$

The spatial terminal solution adds exactly this one candidate to the line solution, $V_m(b,d,\rho) = \max\{V(b,d,\rho), L_{\text{free}}\}$, with L_{free} admitted only when $(b,d) \in \mathcal{C}_\rho$. When either observation is at least the unit scale the maximizer cannot lie strictly in the interior and $V_m = V$: for $b \geq 1$ the problem is exactly that of the line, in every dimension.

The opening coefficient reads off this geometry. Write $J_m(b,\rho) = \mathbb{E}_D\, e^{V_m(b,D,\rho)}$ with $D \sim N(b\rho, 1-\rho^2)$ for the opening value and $\rho = kb$. In every dimension $m \geq 2$ its leading expansion is $e^{-1/2} J_m(b,kb) = 1 + p + b^2 Q_m(k) + O(b^4)$ with $p = 1/\sqrt{2\pi}$ and

$$Q_m(k) = \tfrac{1}{4} + \tfrac{k}{2} - \tfrac{k^2}{4} - p\,k\log k, \qquad 0 < k \leq 1, \tag{26}$$

and Q_m strictly decreasing for $k \geq 1$, so the maximizer lies in $(0,1)$ (Appendix B). It is an entropy-against-quadratic tradeoff, and it closes in the principal Lambert W function.

Proposition 5 (Planar opening). For $\mathrm{Cov}(X_s, X_t) = e^{-\|s-t\|}$ on $\mathbb{R}^m$ with $m \geq 2$, the optimal opening obeys $\rho^*_m(b) = \kappa_{\geq 2}\, b + O(b^3)$, where

$$\kappa_{\geq 2} = \frac{W_0\big(c\, e^{\,c-1}\big)}{c} = 0.604170\ldots, \qquad c = \sqrt{\tfrac{\pi}{2}}. \tag{27}$$

The line and the plane give different constants because they give different feasible geometries: the plane admits the free candidate (25), which the line does not. Why the planar coefficient closes is then immediate. With $c = 1/(2p) = \sqrt{\pi/2}$, the stationary condition of Q_m is

$$Q_m'(k) = \frac{1-k}{2} - p(\log k + 1) = 0 \iff \log k = c(1-k) - 1 \iff (ck)\, e^{ck} = c\, e^{c-1}, \tag{28}$$

so $ck = W_0(ce^{c-1})$. The coefficient is the same for every $m \geq 2$; the three-shot policy is insensitive to further dimension.

Above the unit scale even the Gaussian–Markov structure is unnecessary. Say a field has *metric log-correlations* if its positive correlations satisfy $R_{ij} \geq R_{ik}R_{kj}$ for every triple, equivalently that $-\log R$ obeys the triangle inequalities; these are the path-product inequalities of Gaussian graphical models (Johnson and Smith 1999; Lauritzen, Uhler and Zwiernik 2019, §§4.1–4.2).

Theorem 4 (High-incumbent universality). Let X be a centered, unit-variance Gaussian field with metric log-correlations about an initial location o, and suppose every correlation in some interval $(\rho_0, 1)$ is attainable from o. For $b \geq 1$ and any trial location t, writing $\rho = R(o,t)$,

$$\mathbb{E}\, e^{\max(b,D)} \leq J_X(b,t) \leq J_{\mathrm{OU}}(b,\rho), \qquad D \sim N\big(b\rho, 1-\rho^2\big)\,, \tag{29}$$

and $0 \leq J_{\mathrm{OU}}(b,\rho) - \mathbb{E}\, e^{\max(b,D)} \leq K\, e^b/b^3$ for a constant K, uniformly in ρ. Consequently every global maximizer of the opening satisfies, as $b \to \infty$, $1 - \rho^*(b) \sim c_\infty/b^2$ and $\log J^*(b) = b + A_\infty/b + o(b^{-1})$.

The constants are analytic: $a_\infty = 0.612003\ldots$ is the unique root of $\phi(a) = 2a\,\Phi(-a)$, and $c_\infty = 2a_\infty^2 = 0.749096\ldots$, $A_\infty = c_\infty \Phi(-a_\infty) = 0.202456\ldots$ The upper bound holds because for $b \geq 1$ the maximizer of (23) cannot lie strictly inside $\mathcal{C}_\rho$, so the value is controlled by its OU-boundary maximum, which every metric field's attainable correlations underlie. The attainability interval is needed: correlations that merely accumulate at one, with gaps, can

keep the optimizer from the b^{-2} scale. The class covers powered-exponential fields $e^{-\lambda\|s-t\|^\alpha}$ with $0 < \alpha \leq 1$ and mixtures of exponentials, which are generally not Markov.

The small-incumbent constant thus detects the feasible terminal geometry — 0.540062 on the line, the Lambert-W value 0.604170 in every higher dimension — while the high-incumbent constants are universal across landscapes whose log-correlation is a metric. Full derivations and numerical checks are in the repository (`grass-covariance-geometry.md`).

11 The rule as a line search

The same policy is a derivative-free line search. Used as the inner loop of an iterated random-direction optimizer, with mean, scale, and correlation length estimated online, it is the best inner loop on several rough and high-dimensional objectives of a public physics benchmark, the HumpDay suite (Cotton 2021), and mid-field as a standalone optimizer against a panel of eleven standard methods. It helps when the evaluation budget is comparable to the number of local minima a slice presents and loses when the minima greatly outnumber the budget, consistent with a policy built for terminal placement rather than many-call optimization.

The benchmark table, the heuristic implementation and its caveats, an exponentiated-OU control, an ablation that isolates the third placement, and a pre-registered basin holdout are reported in the extended version and the repository.

12 The rules in business terms

The correlation length sets the regime. A searcher who can place trials many correlation lengths apart should probe far and take the best; the interior premium is first order in the bracket correlation while the value of a fresh draw is not. A venture is usually tethered. Capabilities, customers, and identity keep the next position within roughly one correlation length of the incumbent, and that is the regime in which the results above carry the decision.

The rules are short. Abandon a below-median incumbent for fresh ground. If the trial lands at or below the median, drop it and fall back to the better position's own two-shot rule (Appendix A): move past that position if it is modest, but stay if it is strong, since returning to a strong anchor beats any exterior move; a trial only somewhat below the anchors can still be worth keeping. If both positions are good and neither dominates, commit between them: near the midpoint for balanced moderate values, closer to the better-known position for high or uneven pairs, since the midpoint is exact only when the two are equal. Stay at the better position $B = \max(b, d)$ when $B \geq 1$ and, jointly, $(1 + \rho^2)B - 2\rho D \geq 1 - \rho^2$ (Proposition 2); a weaker second position makes staying *more* likely, not less, and the ceiling C is the threshold only when the two are equal. The interior option is worth about one percent at the optimum, measured against a policy already free to leave in either direction, so most of the value of the last pivot is the information it produces, and the case for making the trial at all is much stronger than the case for any particular refinement afterward.

13 Discussion

Two distinct comparisons should not be merged. Against a *visited* point of matching conditional mean, any unvisited location wins by Jensen alone: it carries variance, the known

value carries none. Between two *unvisited* rivals of matching mean, an interior point and its constructed exterior counterpart, Theorem 1 gives the constant variance ratio C. Neither comparison guarantees that a matched-mean unvisited point exists when the observed value is high.

Mean reversion caps the conditional mean of every unvisited point. Above $b = C$ the cap binds and the optimal final location is one already seen. Below it, the optimal location is past the better end when the news is bad, and between the anchors when the news is good.

A related mid-game rule is proved in Appendix A: a trial that lands at or below the process median is abandoned rather than repaired, and more generally the stay-or-leave decision after an interior trial is a threshold rule in the revealed value, with the measured threshold below the anchors.

With one evaluation remaining there is nothing to explore, since no discovery at the final point can inform a later choice. The last move is pure placement. The preference for uncertain locations comes from Jensen's inequality, not from any value of information.

Exploration proper enters only with more shots remaining. One might guess a two-phase policy, stepping out while values climb and then refining between the two best, but a direct implementation of that heuristic on rough one-dimensional landscapes showed no systematic gain over applying the two-shot rule repeatedly, so we advance no conjecture about its form. The k-shot dynamic program is well defined and each policy evaluation is cheap through the bridge formulas; the adaptive optimization it requires remains to be analyzed.

The same placement problem arises as the line-search step inside derivative-free optimizers, where one-dimensional slices of real objectives can be rough; we pursue that use separately in the humpday benchmark package.

The problem solved here is terminal placement: the payoff is the value at the final location, not the running maximum of the path. Section 2 locates it in the literature; the tradition's exact objects are thresholds in $\sigma^2/|\mu|$ on drifting Brownian paths with endogenous evaluations and flow payoffs, whereas the fixed-budget, terminal-payoff, mean-reverting problem solved here is new. Underneath, terminal placement projects the observed values onto a feasible correlation set, and the problem is exact in every Euclidean dimension. The two regimes read that geometry differently: small-incumbent placement is sensitive to it, so the opening coefficient jumps from 0.540062 on the line to the Lambert-W value 0.604170 in the plane, while large-incumbent behaviour is universal across landscapes whose log-correlation is a metric.

Reproducibility

The results are reproduced by scripts in the browniansearch repository. The three-shot table and phase-diagram figure come from `papers/grass/numerics2.py`; the problem figure from `make_problem_figure.py`; the identities (the constant variance ratio, the $d = 1$ factorization, the stationarity quartic, the flee limit) from `verify/check_review.py`, with earlier counterexample checks in `verify/check_inside.py` and `check_middle.py`; the WiFi diagnostic from `casestudy/wifi/`; the line-search benchmark from `linesearch/`; and the basin holdout, with its pre-registration committed before the results, from `linesearch/exp3_basins/` (`PREREGISTRATION.md` at commit `98f7997`). The opening and spatial results, with their verifiers, are in `grass-opening-analysis.md`, `grass-covariance-geometry.md`, `grass_opening_verify.py`, and `grass_spatial_verify.py`; Figure 3 is from `make_geometry_figure.py`.

A The abandonment threshold

Suppose a bracket has known end log-values b and d and an interior trial reveals log-value v. The searcher may continue inside one of the two sub-brackets or leave past an end.

Proposition 6 (A bad draw is abandoned). Let $v \le 0$, the process median or below. Every placement interior to the two sub-brackets, and the revisit of v, is weakly dominated by exiting to $\zeta(\max(b,d))$; strictly if $v < 0$.

The benchmark $\zeta(\max(b,d))$ is the better anchor's two-shot value: an exterior move when $\max(b,d) < 1$, but a return to that anchor when $\max(b,d) \ge 1$, where $\zeta(\max(b,d)) = \max(b,d)$ is attained by staying. A bad trial is dropped, not necessarily fled past an endpoint.

Proof. Consider the sub-bracket with near end b and far end v, and an interior point with correlations λ_2 and λ to its ends, $\rho_1 = \lambda_2\lambda$. The conditional mean is $\mu = Ab + Bv$ with $A = \lambda_2(1-\lambda^2)/(1-\rho_1^2)$ and $B = \lambda(1-\lambda_2^2)/(1-\rho_1^2)$, both nonnegative, and $A \le \lambda_2$ because $\lambda \ge \rho_1$ implies $1-\lambda^2 \le 1-\rho_1^2$. The conditional variance obeys $\nu = (1-\lambda_2^2)(1-\lambda^2)/(1-\rho_1^2) \le 1-\lambda_2^2 \le 1$ by the same inequality. If $b \ge 0$, then $Ab \le \lambda_2 b$ and $Bv \le 0$ give $\mu \le \lambda_2 b$, so $\mu + \nu/2 \le \lambda_2 b + (1-\lambda_2^2)/2 \le \sup_\ell[\ell b + (1-\ell^2)/2] = \zeta(b)$. If $b < 0$, then $b, v \le 0$ with $A, B \ge 0$ give $\mu \le 0$, so $\mu + \nu/2 \le \frac{1}{2} = \zeta(b)$. Either way the placement loses to $\zeta(b) \le \zeta(\max(b,d))$.

For $v < 0$ a strictly interior point has $B > 0$, so $Bv < 0$; since $Ab \le \lambda_2 b$ for $b \ge 0$ and $Ab \le 0$ for $b < 0$, the mean bound is strict and $\mu + \nu/2 < \zeta(b)$. The revisit of v pays $v \le 0 < \frac{1}{2} \le \zeta(\max(b,d))$. The same bounds apply to the other sub-bracket with d in place of b. □

Proposition 7 (Monotone value). The value of the game is nondecreasing in every observed value.

Proof. Every candidate placement's conditional mean is a nonnegative linear combination of observed values, and every conditional variance depends on locations only. The final payoff is a maximum of terms that are nondecreasing in each observed value. For earlier stages, raise one observed value and couple the reveal at any placement by shifting it with its conditional mean. By the Markov property of the OU landscape each continuation is a fresh game anchored at the current observations, so induction over stages preserves monotonicity through the expectation and the maximum. □

Exits are anchored at the outer values and do not depend on the interior reveal v, while continuations inside are nondecreasing in v by the monotonicity above. The stay-or-leave decision after an interior trial is therefore a threshold rule in v. The threshold can lie below both anchors: at $b = d = 0.7$, bracket correlation $\rho = 0.5$, and the trial at the midpoint, staying is strictly preferable when $v > v_*$ with $v_* \approx 0.63010$, with indifference at equality; a quarter-bracket trial moves the threshold to 0.55061. A trial somewhat worse than both anchors is still worth staying for, because the sub-bracket keeps one good end and the variance. This comparison assumes one final placement remains: over a longer horizon an exterior trial followed by a possible return can still carry continuation value depending on v, which Proposition 7 does not by itself exclude. Verification scripts accompany the paper's repository.

B Proofs of the opening asymptotics

Throughout $p = \phi(0) = 1/\sqrt{2\pi}$, $Z \sim N(0,1)$, and $h(d) = e^{\zeta(d)}$, with $h(d)\phi(d) = e^{1/2}p$ on $0 < d < 1$. Full detail is in the repository (`grass-opening-analysis.md`, `grass-covariance-geometry.md`).

The line coefficient (Proposition 4)

Write $J(b,\rho)$ for the opening value, the terminal value averaged over the reveal $D \sim N(b\rho, 1-\rho^2)$. At $b = 0$ strict convex ordering gives $J(0,\rho) < J(0,0)$ for every $\rho > 0$, and coupling the reveal at incumbent b to that at 0 raises each terminal conditional mean by at most b, so $J(b,\rho) \le e^b J(0,\rho)$ and $\rho^*(b) \to 0$. For the sharper bound, a bridge point with correlations x, y to the two observations and $t = (x-\rho y)/(1-\rho^2) \ge 0$ has $L = dy + \frac{1}{2}(1-y^2) + (b-\rho d)t - \frac{1}{2}(1-\rho^2)t^2$, whose last two terms are at most $b^2/[2(1-\rho^2)]$ for $d \ge 0$; abandonment gives the same bound for $d < 0$. Hence $J(b,\rho) \le J_0 + Cb^2 + Cb\rho - c\rho^2$ locally, and comparison with $J(b,0) \ge J_0$ forces $\rho^* = O(b)$.

Put $\rho = kb$. For fixed $0 < d < 1$ the limiting optimum attached to d is $y = d$ (at finite small b, $y = d + O(b^2)$), and the envelope theorem gives $\sup_y L = \zeta(d) + b^2 f(d,k) + O(b^4)$ with $f(d,k) = k(d^{-1} - d) - \frac{k^2}{2}(d^{-2} - d^2)$; the exterior option keeps the uplift at $[f]_+$, and $f > 0$ iff $d > a(k) = k/(1+\sqrt{1-k^2})$. Using $h\phi = e^{1/2}p$ and expanding the reveal density, $e^{-1/2}J(b,kb) = 1 + p + b^2 Q(k) + O(b^3)$ with

$$Q(k) = \tfrac{1}{4} + \tfrac{1+p}{2}k - \left(\tfrac{1}{4} + \tfrac{2p}{3}\right)k^2 + p\int_{a(k)}^{1} f(d,k)\,dd, \tag{30}$$

the integral elementary. Then $Q'(k) = \frac{1-k}{2} + p[-\log a + \frac{a^2}{2} - \frac{k}{a} - \frac{ka^3}{3}]$, and substituting $k = 2a/(1+a^2)$ reduces $Q' = 0$ to the equation of Proposition 4, root $a_0 = 0.293253$, $\kappa_0 = 2a_0/(1+a_0^2) = 0.540062$. Uniqueness holds because $Q''(k) = -\frac{1}{2} + p(-\frac{1}{2a} + a + \frac{a^3}{6}) < 0$ (the parenthesis rises in a to at most $2/3$, and $2p/3 < \frac{1}{2}$), with $Q'(0^+) = +\infty$, $Q'(1^-) = -5p/6 < 0$, and the $k \ge 1$ branch decreasing. The active region is bounded away from $d = 0$ near κ_0 and the k-independent b^3 term (from $0 < d < b$) does not move the maximizer, so $0 = \partial_k J = e^{1/2}b^2\{Q'(k^*) + O(b^2)\}$ gives $k^*(b) = \kappa_0 + O(b^2)$.

The endpoints are excluded for every $b > 0$. The fresh-draw derivative $\partial_\rho J(b, 0^+) > 0$ is a strictly positive Gaussian integral, so $\rho = 0$ is never optimal; and near $\rho = 1$ the value carries a strictly positive $\sqrt{1-\rho}$ term from the selected positive reveals, so $\rho = 1$ is never optimal. Hence $0 < \rho^*(b) < 1$.

The planar coefficient (Proposition 5)

In dimension $m \ge 2$, writing the terminal correlation to the incumbent as $x = bu$, the free candidate $(x,y) = (b,d)$ is admissible when $(b,d) \in \mathcal{C}_\rho$, i.e. $kb^2 \le d \le \min(k, 1/k)$ with $\rho = kb$, and it improves on the exterior by $\frac{b^2}{2}[(1-kd)^2 - (u-1)^2] + O(b^4)$ for fixed parameters, maximized at $u = 1$. Integrating gives $Q_m(k) = \frac{1}{4} + \frac{k}{2} - \frac{k^2}{4} - pk\log k$ on $0 < k \le 1$. For $k \ge 1$ the admissible reveal region does not vanish but shrinks to $kb^2 \le d \le 1/k$, contributing the $p/(6k)$ term, so $Q_m(k) = \frac{1}{4} + \frac{1+p}{2}k - (\frac{1}{4} + \frac{2p}{3})k^2 + \frac{p}{6k}$ is strictly decreasing, and the maximizer lies in $(0,1)$. On $(0,1)$, $Q_m'(k) = \frac{1-k}{2} - p(\log k + 1)$ and $Q_m''(k) = -\frac{1}{2} - \frac{p}{k} < 0$, so Q_m is strictly concave with a unique interior maximizer. With $c = 1/(2p) = \sqrt{\pi/2}$,

$Q'_m(k) = 0 \iff \log k = c(1-k) - 1 \iff (ck)e^{ck} = ce^{c-1}$, hence $ck = W_0(ce^{c-1})$ and $\kappa_{\geq 2} = W_0(ce^{c-1})/c = 0.604170$, independent of m. The localization and remainder arguments are those of the line, with the small-reveal crossover now at $d = kb^2$, so no cubic-in-b value term arises.

High-incumbent universality (Theorem 4)

For $b \geq 1$ the exterior-only reward equals $e^{\max(b,D)}$, the lower bound; by (23) a maximizer that cannot lie strictly inside $\mathcal{C}_\rho$ is bounded above by the OU-boundary optimum, which is (29). The gap is uniform. Set $\varepsilon = 1 - \rho$ and $\Delta = D - b$: a bridge improvement is impossible once $b \geq (1+\rho)/(1-\rho)$, and otherwise $\varepsilon < 2/(b+1)$ confines the improving reveals to $|\Delta| \leq 2\varepsilon$ for large b. There the excess over $e^{\max(b,D)}$, divided by e^b, is at most $C\varepsilon$ (bridge mean at most the larger anchor, variance at most $\varepsilon/(2-\varepsilon)$). Since $\Pr\{|\Delta| \leq 2\varepsilon\} \leq C\sqrt{\varepsilon}\, e^{-b^2\varepsilon/16}$, the expected excess is at most $C\varepsilon^{3/2}e^{-b^2\varepsilon/16} \leq Kb^{-3}$, the uniform bound.

Set $\varepsilon = c/b^2$. Uniformly on compact $c > 0$, $b\, e^{-b}[J - e^b] \to A(c) = \mathbb{E}(\sqrt{2c}\, Z - c)_+ = \sqrt{2c}\,\phi(\sqrt{c/2}) - c\,\Phi(-\sqrt{c/2})$. No maximizing sequence leaves this scale: with $x = b^2\varepsilon$, the tilting bound on the anchor reward, plus the bridge remainder, gives $b\, e^{-b}[J - e^b] \leq \sqrt{2x}\,\phi((1 - 2/b)\sqrt{x/2}) + K/b^2$; the remainder vanishes as $b \to \infty$ and the first term vanishes as $x \to 0$ or $x \to \infty$, while $A(c) > 0$ at fixed c. Then $A'(c) = 0$ reads $\phi(a_\infty) = 2a_\infty\Phi(-a_\infty)$ with $a_\infty = \sqrt{c_\infty/2}$, $c_\infty = 2a_\infty^2 = 0.749096$, and $A_\infty = A(c_\infty) = c_\infty\Phi(-a_\infty) = 0.202456$; uniqueness holds because $a\,\Phi(-a)/\phi(a)$ increases strictly from 0 to 1. Compactness and uniform convergence give $b^2(1-\rho^*) \to c_\infty$ and $\log J^* = b + A_\infty/b + o(b^{-1})$ for every global maximizer.

References

[1] S. Callander. Searching for good policies. *American Political Science Review*, 105(4):643–662, 2011.

[2] S. Callander. Searching and learning by trial and error. *American Economic Review*, 101(6):2277–2308, 2011.

[3] A. Bardhi and S. Callander. Learning in a correlated world. *Annual Review of Economics*, forthcoming.

[4] A. Bardhi. Attributes: selective learning and influence. *Econometrica*, 92(2):311–353, 2024.

[5] C. Urgun and L. Yariv. Contiguous search: exploration and ambition on uncharted terrain. *Journal of Political Economy*, 133, 2025.

[6] Y. F. Wong. Forward-looking experimentation of correlated alternatives. *Theoretical Economics*, 20, 2025.

[7] S. Callander. A theory of policy expertise. *Quarterly Journal of Political Science*, 3(2):123–140, 2008.

[8] S. Callander and P. Hummel. Preemptive policy experimentation. *Econometrica*, 82(4):1509–1528, 2014.

[9] S. Callander and T. S. Clark. Precedent and doctrine in a complicated world. *American Political Science Review*, 111(1):184–203, 2017.

[10] S. Callander and N. Matouschek. The risk of failure: trial and error learning and long-run performance. *American Economic Journal: Microeconomics*, 11(1):44–78, 2019.

[11] S. Callander and N. McCarty. Agenda control under policy uncertainty. *American Journal of Political Science*, 68(1):210–226, 2024.

[12] S. Callander, N. S. Lambert, and N. Matouschek. Innovation and competition on a rugged technological landscape. *American Economic Journal: Microeconomics*, forthcoming.

[13] Y. C. Aybas and S. Callander. Cheap talk in complex environments. Working paper, 2025.

[14] S. C. Ganz. Hyperopic search: organizations learning about managers learning about strategies. *Organization Science*, 31(4):821–838, 2020.

[15] U. Garfagnini and B. Strulovici. Social experimentation with interdependent and expanding technologies. *Review of Economic Studies*, 83(4):1579–1613, 2016.

[16] D. Cetemen, C. Urgun, and L. Yariv. Collective progress: dynamics of exit waves. *Journal of Political Economy*, 131(9):2402–2450, 2023.

[17] C. Carnehl and J. Schneider. A quest for knowledge. *Econometrica*, 93(2):623–659, 2025.

[18] C. Hodgson and G. Lewis. You can lead a horse to water: spatial learning and path dependence in consumer search. *Econometrica*, 93(4):1299–1332, 2025.

[19] D. Glick and C. D. Myers. Learning from others: an experimental test of Brownian motion uncertainty models. *Journal of Theoretical Politics*, 27(4):588–612, 2015.

[20] P. Cotton. HumpDay: pure Python derivative-free optimization and its application suite. https://github.com/microprediction/humpday, 2021.

[21] H. J. Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering*, 86(1):97–106, 1964.

[22] J. Mockus, V. Tiesis, and A. Zilinskas. The application of Bayesian methods for seeking the extremum. In *Towards Global Optimization*, volume 2, pages 117–129. North-Holland, 1978.

[23] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.

[24] N. Srinivas, A. Krause, S. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning*, 2010.

[25] P. Hennig and C. J. Schuler. Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13:1809–1837, 2012.

[26] Z. Wang and S. Jegelka. Max-value entropy search for efficient Bayesian optimization. In *Proceedings of the 34th International Conference on Machine Learning*, 2017.

[27] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: a review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

[28] D. Ginsbourger and R. Le Riche. Towards Gaussian process-based optimization with finite time horizon. In *mODa 9 — Advances in Model-Oriented Design and Analysis*, pages 89–96. Physica-Verlag, 2010.

[29] R. Lam, K. Willcox, and D. H. Wolpert. Bayesian optimization with a finite budget: an approximate dynamic programming approach. In *Advances in Neural Information Processing Systems 29*, 2016.

[30] S. Jiang, D. R. Jiang, M. Balandat, B. Karrer, J. R. Gardner, and R. Garnett. Efficient nonmyopic Bayesian optimization via one-shot multi-step trees. In *Advances in Neural Information Processing Systems 33*, 2020.

[31] P. I. Frazier, W. B. Powell, and S. Dayanik. The knowledge-gradient policy for correlated normal beliefs. *INFORMS Journal on Computing*, 21(4):599–613, 2009.

[32] J. Wu and P. I. Frazier. Practical two-step lookahead Bayesian optimization. In *Advances in Neural Information Processing Systems 32*, 2019.

[33] M. Mahsereci and P. Hennig. Probabilistic line searches for stochastic optimization. *Journal of Machine Learning Research*, 18:1–59, 2017.

[34] S. Malladi. Searching in the dark and learning where to look. Working paper, 2025.

[35] M. Banchio and S. Malladi. Rediscovery. Working paper, arXiv:2504.19761, 2025.

[36] J. Grosse, C. Zhang, and P. Hennig. Optimistic optimization of Gaussian process samples. *Transactions on Machine Learning Research*, 2023.

[37] M. A. Butler and A. A. King. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist*, 164(6):683–695, 2004.

[38] T. F. Hansen. Stabilizing selection and the comparative analysis of adaptation. *Evolution*, 51(5):1341–1351, 1997.

[39] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[40] J. Snoek, H. Larochelle, and R. P. Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems 25*, 2012.

[41] E. Weinberger. Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biological Cybernetics*, 63(5):325–336, 1990.

[42] X. Feng, K. A. Nguyen, and Z. Luo. WiFi RTT RSS dataset for indoor positioning. Zenodo, 2024. doi:10.5281/zenodo.11558192.

[43] M. Gudmundson. Correlation model for shadow fading in mobile radio systems. *Electronics Letters*, 27(23):2145–2146, 1991.

[44] E. Ries. *The Lean Startup*. Crown Business, New York, 2011.

[45] A. Camuffo, A. Cordova, A. Gambardella, and C. Spina. A scientific approach to entrepreneurial decision making: evidence from a randomized control trial. *Management Science*, 66(2):564–586, 2020.

[46] A. Camuffo, A. Gambardella, D. Messinese, E. Novelli, E. Paolucci, and C. Spina. A scientific approach to entrepreneurial decision making: large scale replication and extension. *Strategic Management Journal*, 45(6):1209–1237, 2024.

[47] J.-B. Grill, M. Valko, and R. Munos. Optimistic optimization of a Brownian. In *Advances in Neural Information Processing Systems 31*, 2018.

[48] J. M. Calvin. Average performance of a class of adaptive algorithms for global optimization. *Annals of Applied Probability*, 7(3):711–730, 1997.

[49] A. Alabert and R. Caballero. On the minimum of a conditioned Brownian bridge. *Stochastic Models*, 34(3):269–291, 2018.

[50] C. R. Johnson and R. L. Smith. Path product matrices. *Linear and Multilinear Algebra*, 46(3):177–191, 1999.

[51] S. Lauritzen, C. Uhler, and P. Zwiernik. Maximum likelihood estimation in Gaussian models under total positivity. *The Annals of Statistics*, 47(4):1835–1863, 2019.