Exploratory Responsiveness and Adaptive Rigidity under AI-Assisted Optimization

Balaraju Battu[1,2]
[1]European University Institute, Florence, Italy
[2]New York University Abu Dhabi, UAE

May 20, 2026

Abstract

This paper develops a theory of exploratory adaptation under AI-assisted optimization. The central argument is that the long-run adaptive effects of AI systems depend critically on how predictive assistance interacts with exploratory responsiveness itself. We formalize this mechanism using a dynamical framework in which cognitive, institutional, and technological systems evolve over rugged epistemic landscapes characterized by multiple locally reinforced configurations. A central state variable in the model is adaptive responsiveness, which measures the capacity of a system to traverse unfamiliar conceptual and institutional trajectories under changing conditions. Under convergent predictive regimes, AI systems substitute for exploratory engagement, reducing adaptive responsiveness and generating metastable trapping, hysteresis, premature convergence, and exploration-collapse dynamics in which systems become locally efficient but globally rigid. The framework also identifies contrasting exploration-enhancing regimes in which AI systems amplify exploratory search, conceptual traversal, and adaptive mobility. The effective substitution parameter is therefore responsiveness-dependent: systems possessing weak exploratory routines are more vulnerable to exploratory substitution, whereas systems already possessing high adaptive responsiveness may use AI assistance to expand exploratory mobility across rugged landscapes. The long-run adaptive effects of AI consequently depend not only on AI capability itself, but also on institutional structure, developmental context, and the architecture of human-machine interaction.

1 Introduction

Recent advances in artificial intelligence have substantially increased the ability of computational systems to provide prediction, recommendation, and task-specific guidance across a wide range of economic and cognitive activities [1, 2, 3, 4]. In education, programming, scientific research, and professional decision-making, AI systems increasingly reduce the cost of acquiring information, navigating uncertainty, and generating acceptable solutions. These developments are widely viewed as a major source of productivity growth and economic transformation [5, 6, 7]. At the same time, a growing literature suggests that systems improving short-run performance may also weaken the exploratory processes through which adaptive flexibility develops. Research on cognitive offloading

shows that external systems reducing immediate cognitive demand can weaken independent memory formation, metacognitive engagement, and exploratory reasoning [8, 9, 10, 11]. Experimental and theoretical work on AI assistance similarly finds that systems improving local productivity may simultaneously reduce skill formation and adaptive learning [12, 13, 14]. In collaborative knowledge environments, generative AI adoption has also been associated with declining public knowledge sharing and reduced contribution activity [15, 16], while recursive dependence on AI-generated outputs may reduce informational diversity itself [17].

These findings point toward a broader economic problem. Many socially valuable forms of adaptation emerge through decentralized exploratory activity that individual agents do not fully internalize. Scientific discovery, institutional adaptation, apprenticeship, collective learning, and open-ended inquiry all depend on costly processes of experimentation, disagreement, revision, uncertainty navigation, and publicly observable problem-solving. Such exploratory activity generates positive externalities because it contributes not only to the production of knowledge, but also to the preservation of adaptive responsiveness itself.

The central argument of this paper is that the long-run adaptive effects of AI systems depend critically on how predictive assistance interacts with exploratory responsiveness. Systems possessing weakly developed exploratory routines may experience AI assistance primarily as a substitute for exploratory engagement, generating declining responsiveness and increasing rigidity over time. By contrast, systems already possessing high adaptive responsiveness may use AI assistance to expand exploratory variance, conceptual traversal, and adaptive mobility. The framework therefore distinguishes between convergent predictive systems that compress exploratory responsiveness and exploration-enhancing systems that amplify it. The central concern of the paper is consequently not AI assistance in general, but the conditions under which optimization compresses or preserves exploratory mobility across cognitive and institutional landscapes.

This perspective relates closely to recent work by Acemoglu, Kong, and Ozdaglar [18], who study how agentic AI systems may reduce incentives for costly human learning while weakening the accumulation of collective knowledge. In their framework, human learning generates informational externalities because individual effort contributes to a socially shared stock of useful knowledge. As AI systems increasingly substitute for human effort, equilibrium incentives for learning decline, potentially producing what the authors describe as knowledge collapse. The present paper builds on this insight but extends the analysis from knowledge accumulation to adaptive mobility itself. Knowledge accumulation alone does not guarantee continued adaptation. Societies may possess substantial informational resources while simultaneously losing the capacity to experiment, revise, and traverse unfamiliar conceptual or institutional trajectories under changing conditions. By introducing adaptive responsiveness as an endogenous state variable evolving over rugged epistemic landscapes, the present framework studies how optimization reshapes the exploratory mobility required for long-run adaptation itself.

To formalize this idea, we study collective cognition as a dynamical system evolving over a rugged epistemic landscape. Let $x_t \in \mathcal{L}$ denote the state of a cognitive, institutional, or behavioral system, where $\mathcal{L}$ is a high-dimensional space of possible configurations. The geometry of the

landscape is summarized by an effective potential function

$$V : \mathcal{L} \to \mathbb{R}.$$

The function $V(x)$ should not be interpreted as a literal utility function or consciously optimized objective. Instead, it summarizes the effective stabilization pressures acting on the system. These stabilization pressures emerge through institutional reinforcement, coordination effects, technological standardization, learning dynamics, social expectations, and repeated adaptive interaction. Lower values of $V(x)$ correspond to states that are relatively predictable, internally coherent, socially reinforced, institutionally stabilized, and low in coordination tension. Higher values correspond to states characterized by uncertainty, instability, novelty, unresolved contradiction, or cognitive and institutional strain.

Empirically, movement across the landscape can be understood as movement across problem formulations, organizational routines, technological standards, or scientific domains. The barriers separating local basins therefore represent the frictions constraining movement across alternative configurations. In technological environments, steep barriers may reflect the financial and organizational costs of abandoning established standards. In scientific environments, they may reflect the cognitive and professional costs associated with traversing distant epistemic regions. More generally, such barriers may correspond to coordination costs, retraining burdens, cognitive strain, institutional switching costs, legitimacy pressures, or accumulated infrastructure dependence. Recent evidence on scientific "pivot penalties" provides one observable analogue of such barriers: researchers moving too far from established trajectories face measurable traversal costs, consistent with the interpretation of rugged epistemic landscapes as locally reinforced knowledge structures [19].

This interpretation connects naturally to institutional theories emphasizing how stable rules, routines, shared belief systems, and social norms reduce uncertainty and structure adaptive behavior [20, 21, 22]. In this sense, locally stable regions of the landscape correspond not only to cognitive attractors, but also to institutionally reinforced configurations that economize on coordination and uncertainty. The resulting landscape is generally rugged and jagged rather than convex because adaptive systems often contain multiple locally reinforced configurations separated by barriers of varying steepness and uneven traversal difficulty [23]. Arthur [24] argues that self-reinforcing economic mechanisms can generate multiple locally stable outcomes through increasing returns, coordination effects, adaptive expectations, and learning dynamics. Early fluctuations may then push systems toward particular trajectories that gradually become locked in through positive feedback processes.

Attractors in this framework therefore represent locally stabilized cognitive, institutional, or technological configurations toward which adaptive systems repeatedly converge. Scientific paradigms, organizational routines, technological standards, and dominant explanatory systems may all function as attractors when reinforcement and coordination effects stabilize particular trajectories over time. Recent large-scale analyses of scientific and technological research document substantial "pivot penalties" when researchers shift too far from established trajectories, consistent with rugged epistemic landscapes and locally reinforced knowledge structures [19]. This interpretation also con-

nects to Bourdieu's conception of social fields as structured relational spaces in which established positions, routines, and forms of capital generate persistent trajectories and reproduce locally stable orders over time [25, 26]. More broadly, the framework connects naturally to work on complex adaptive systems emphasizing nonlinear feedbacks, path dependence, and emergent system-level organization [27]. Social and institutional systems may therefore become locally stable while remaining globally fragile because adaptation across rugged landscapes requires continued exploratory mobility.

The system evolves according to

$$dx_t = -z(t)\nabla V(x_t)\,dt + \sigma(t)dW_t. \tag{1}$$

The first term captures local stabilization dynamics. Systems tend to move toward nearby configurations that reduce uncertainty, coordination strain, or adaptive tension. This does not assume global optimization or fully rational behavior. Repeated local updating alone generates effective gradient-like stabilization dynamics. The second term captures exploratory perturbations generated through experimentation, disagreement, uncertainty navigation, and stochastic search. Environments characterized by rapid technological change, institutional instability, scientific uncertainty, or cultural fragmentation generally require greater exploratory traversal across the landscape, while stable and highly standardized environments permit systems to remain localized near existing attractors for longer periods.

The key state variable in the framework is adaptive responsiveness $z(t) \in [0, 1]$, where $z = 0$ denotes complete rigidity and $z = 1$ denotes maximal exploratory responsiveness and it measures the effective mobility of the system across epistemic space. This interpretation connects naturally to research in cognitive neuroscience emphasizing executive flexibility, exploratory engagement, and sensitive developmental periods in the maintenance of adaptive cognitive capacity [28, 29]. High responsiveness permits continued exploration and adaptation across regions of the landscape, while low responsiveness generates rigidity and local trapping. Intuitively, responsiveness captures whether a system remains capable of exploring beyond familiar trajectories even after predictive tools make local optimization easier. Responsiveness evolves according to

$$\dot{z}(t) = \eta z(t)^{\gamma} s_t(1 - z(t)) - \rho z(t)(1 - s_t). \tag{2}$$

The first term captures the regeneration of responsiveness through exploratory activity, while the factor $(1 - z)$ imposes saturation, reflecting the idea that responsiveness cannot increase indefinitely. The second term captures responsiveness erosion when exploratory activity declines and stabilization pressures dominate. The parameter $\gamma \in (0, 1]$ governs the responsiveness regeneration process. Lower values imply that responsiveness remains partially self-reinforcing even at low levels, allowing exploratory capacity to recover despite periods of stabilization or reduced exploration. Educational environments emphasizing open-ended reasoning, decentralized experimentation, and productive disagreement may exhibit this type of adaptive structure. Higher values imply that responsiveness regeneration depends more strongly on already-existing exploratory capacity. In such systems, exploratory decline reinforces future rigidity because adaptation increasingly depends on

accumulated expertise, institutional memory, and established exploratory routines.

Systems characterized by strong increasing returns, coordination effects, accumulated specialization, and path dependence may therefore exhibit high-γ dynamics [24, 20]. In such environments, exploratory responsiveness becomes increasingly dependent on previously accumulated routines, institutional memory, shared expectations, and established coordination structures, making adaptation progressively more difficult once exploratory flexibility deteriorates. Scientific paradigms, technological systems, and institutional environments may therefore become increasingly rigid over time as exploratory decline reinforces existing trajectories. The historical competition between VHS and Sony Betamax illustrates how self-reinforcing coordination effects can progressively lock systems into locally dominant trajectories once exploratory mobility narrows [24]. Recent evidence from scientific research similarly suggests that movement across distant epistemic regions becomes increasingly difficult as specialization deepens, consistent with highly path-dependent exploratory dynamics [19].

We define s_t as productive exploratory activity, including experimentation, uncertainty navigation, exploratory deliberation, self-guided problem solving, publicly observable reasoning, exploratory search, disagreement, and sustained engagement with unresolved problems. Exploratory activity is costly because it requires time, cognitive effort, uncertainty navigation, and sustained engagement with unfamiliar or unresolved conceptual regions. Empirically, declines in s_t may correspond to fewer discarded code branches, reduced exploratory drafting, weaker public problem-solving activity, narrower scientific search, or declining attempts to traverse unfamiliar conceptual regions.

We define α as the effective substitution parameter governing the extent to which AI systems substitute for or complement exploratory deliberation and search processes. In this sense, α does not measure AI use alone, but rather how AI systems alter the cost and frequency of exploratory traversal. Positive values of α correspond to environments in which predictive systems reduce costly exploratory engagement by replacing human uncertainty navigation with optimized prediction, recommendation, or automated solution generation. Negative values correspond to environments in which AI systems instead amplify exploratory activity by expanding conceptual traversal, adversarial reasoning, or exploratory variance.

A central case considered in the model is therefore that increasing AI assistance alters exploratory activity according to

$$s_t = s_0 - \alpha AI_t. \tag{3}$$

Here s_0 denotes baseline exploratory engagement, while AI_t measures the intensity of AI assistance.

The effective substitution parameter may be written more generally as

$$\alpha = \alpha(z, \mathcal{I}, \mathcal{A}),$$

where z denotes the existing stock of adaptive responsiveness, $\mathcal{I}$ captures the surrounding institutional environment, and $\mathcal{A}$ denotes the architecture of AI assistance itself. The sign of α governs whether AI systems compress or expand exploratory responsiveness. Positive values ($\alpha > 0$) cor-

respond to convergent predictive systems that substitute for exploratory engagement and reduce adaptive mobility over time, while negative values ($\alpha < 0$) correspond to exploration-enhancing systems that amplify exploratory search, conceptual traversal, and adaptive variation.

This interpretation implies that the adaptive effects of AI assistance are not technologically fixed, but depend on the interaction between existing responsiveness, institutional structure, and AI architecture. Systems already possessing high adaptive responsiveness may use AI assistance to expand exploratory variance, generate adversarial hypotheses, and traverse unfamiliar conceptual regions more effectively. In scientific environments, for example, generative systems capable of proposing anomalous hypotheses, simulating alternative mechanisms, or identifying unexpected cross-domain connections may increase exploratory search rather than compress it. By contrast, in environments where exploratory routines remain underdeveloped, predictive systems may strongly substitute for exploratory effort because foundational reasoning and uncertainty-navigation capacities have not yet stabilized. Early overreliance on predictive assistance may therefore weaken the adaptive routines required for later exploratory responsiveness [28, 29]. Under some institutional and developmental conditions, AI systems may consequently reduce or even reverse the effective substitution effect, potentially generating $\alpha < 0$ [30].

Substituting yields

$$\dot{z}(t) = \eta z(t)^{\gamma}(s_0 - \alpha A I_t)(1 - z(t)) - \rho z(t)\left[1 - (s_0 - \alpha A I_t)\right]. \tag{4}$$

The dynamics imply a simple but important mechanism. Under predictive systems for which $\alpha > 0$, increasing AI-assisted optimization reduces exploratory engagement, gradually lowering the adaptive responsiveness required for long-run mobility across epistemic and institutional landscapes. Systems may therefore become increasingly efficient within familiar trajectories while simultaneously losing the capacity to traverse unfamiliar conceptual regions when environments change. By contrast, exploration-enhancing systems characterized by sufficiently low or negative values of α may instead preserve or amplify exploratory mobility by lowering the cost of conceptual traversal, adversarial reasoning, and exploratory search. The long-run adaptive effects of AI therefore depend critically on the interaction between responsiveness, institutional structure, and the architecture of AI assistance itself.

Existing work on AI and economic performance largely focuses on improvements in prediction, information processing, productivity, and local task performance [5, 3]. By contrast, the present framework studies how predictive systems reshape the equilibrium geometry of collective exploration itself. The model naturally generates several phenomena increasingly relevant in AI-mediated environments, including underproduction of exploratory activity, metastable trapping, hysteresis, premature convergence, attractor rigidity, and exploration-collapse transitions. Greater landscape ruggedness amplifies these effects because systems confronting fragmented, uncertain, or rapidly changing environments require higher responsiveness to avoid becoming trapped near locally stable but globally suboptimal configurations.

The framework also connects several previously distinct literatures. Research in educational psychology shows that productive struggle and desirable difficulty improve long-run learning and transfer [31, 32]. Work in cognitive science emphasizes the importance of exploratory engage-

ment and uncertainty navigation for adaptive flexibility [33, 34]. Research on cognitive offloading demonstrates that systems reducing immediate cognitive demand can weaken independent memory formation and exploratory reasoning [8, 11, 35]. Recent empirical work further documents reductions in public knowledge generation following the diffusion of generative AI systems [15, 16]. The present paper proposes a unified economic framework connecting these phenomena through the dynamics of exploratory responsiveness.

More broadly, the paper contributes to a growing literature studying how AI systems reshape the production of knowledge, expertise, and human capital [36, 37, 38]. Existing approaches primarily focus on informational substitution and labor displacement. The present framework instead emphasizes a distinct channel: predictive systems may alter the equilibrium level of socially valuable exploration required for long-run adaptive capacity. The remainder of the paper develops the formal model, characterizes its dynamical structure, derives conditions for rigidity and exploration-collapse transitions, and studies how reductions in exploratory activity reshape the long-run adaptive capacity of collective intelligence systems.

2 A Model of Exploratory Adaptation

This section develops the formal framework linking exploratory activity, adaptive responsiveness, and AI-assisted optimization. The central premise is that long-run adaptation depends not only on the accumulation of information, but also on the preservation of exploratory mobility across cognitive and institutional landscapes. Systems confronting uncertain or rapidly changing environments must repeatedly traverse unfamiliar regions of epistemic and institutional space rather than remain confined near existing attractors. Exploratory activity therefore generates socially valuable adaptive capacity because experimentation, exploratory deliberation, uncertainty navigation, and publicly observable reasoning sustain the mobility required for long-run adaptation.

We represent local adaptive motion using gradient-like dynamics:

$$dx_t = -z(t)\nabla V(x_t)\,dt + \sigma(t)dW_t. \tag{5}$$

The gradient term captures stabilizing pressure toward nearby locally coherent configurations, while the stochastic term captures exploratory perturbations generated through experimentation, decentralized search, disagreement, uncertainty navigation, exploratory variation, and stochastic adaptation [39, 34, 40]. This specification does not assume globally rational optimization. Rather, it reflects the general tendency of adaptive systems to reinforce locally stable trajectories while exploration generates perturbations capable of escaping local attractors. The central state variable in the framework is adaptive responsiveness $z(t) \in [0, 1]$, which measures the effective mobility of the system across epistemic space. High responsiveness permits continued exploration and reconfiguration under changing conditions, while low responsiveness generates rigidity and confinement near existing attractors. The distinction is important because systems may possess substantial accumulated knowledge while simultaneously losing the capacity to adapt effectively to unfamiliar environments. A firm can become highly efficient within an existing production regime while losing the ability to respond to technological disruption, while scientific communities may accumulate

extensive knowledge while becoming increasingly conservative in exploratory behavior.

Exploration consequently plays a critical role because deterministic stabilization dynamics alone drive systems toward nearby local minima. Productive struggle, exploratory search, and publicly observable reasoning therefore sustain both exploratory variance and adaptive responsiveness itself. This interpretation connects naturally to research in educational psychology and cognitive science emphasizing the importance of desirable difficulty, productive failure, exploratory engagement, and uncertainty navigation for long-run adaptive flexibility [31, 32, 33]. Related evidence from cognitive offloading shows that systems reducing immediate cognitive demand can weaken independent memory formation and exploratory reasoning [8, 9, 11]. Recent work on AI assistance similarly finds that systems improving local productivity may simultaneously weaken long-run skill formation and adaptive learning [12, 13]. At the collective level, generative AI diffusion has also been associated with declining public knowledge sharing and reduced contribution activity on collaborative platforms [15, 16]. Taken together, this literature suggests that systems minimizing exploratory effort too aggressively may preserve local efficiency while weakening long-run adaptive responsiveness.

To formalize this mechanism, we model adaptive responsiveness as an endogenous stock evolving according to

$$\dot{z}(t) = \eta z(t)^{\gamma} s_t (1 - z(t)) - \rho z(t)(1 - s_t). \tag{6}$$

The first term captures responsiveness regeneration through exploratory activity, while the factor $(1 - z)$ imposes saturation, reflecting the idea that responsiveness cannot increase indefinitely. The second term captures responsiveness erosion when exploratory engagement declines and stabilization pressures dominate. The parameter $\gamma \in (0, 1]$ governs the responsiveness regeneration process itself. Lower values imply that responsiveness remains partially self-reinforcing even at low levels, allowing systems to recover adaptive capacity despite periods of stabilization or reduced exploration. Educational environments emphasizing open-ended reasoning, decentralized experimentation, and productive disagreement may exhibit this type of adaptive structure. Higher values imply that responsiveness regeneration depends more strongly on already-existing exploratory capacity. In such environments, exploratory decline reinforces future rigidity because adaptation increasingly depends on accumulated expertise, institutional memory, and established exploratory routines.

This distinction becomes especially important in environments characterized by strong increasing returns, coordination effects, and path dependence [24, 20]. The historical competition between VHS and Sony Betamax illustrates how self-reinforcing coordination dynamics can progressively lock systems into locally dominant trajectories once exploratory flexibility deteriorates [24]. Recent evidence documenting substantial "pivot penalties" in scientific research similarly suggests that movement across distant epistemic regions becomes increasingly difficult as specialization deepens [19]. Highly path-dependent systems may therefore exhibit strong hysteresis because exploratory decline erodes the very capacities required for future adaptation.

We define s_t as productive exploratory engagement, including experimentation, exploratory deliberation, uncertainty navigation, publicly observable reasoning, disagreement, and sustained engagement with unresolved problems. Exploratory activity is costly because it requires time, cognitive effort, uncertainty navigation, and sustained engagement with unfamiliar or unresolved

conceptual regions. Empirically, declines in s_t may correspond to fewer discarded code branches, reduced exploratory drafting, weaker public problem-solving activity, narrower scientific search, or declining attempts to traverse unfamiliar conceptual regions.

AI assistance alters exploratory activity according to

$$s_t = s_0 - \alpha A I_t,$$

where s_0 denotes baseline exploratory engagement and α captures the extent to which AI systems substitute for or complement exploratory effort. More generally, the effective substitution parameter may itself depend on the adaptive state of the system, the surrounding institutional environment, and the architecture of AI assistance:

$$\alpha = \alpha(z, \mathcal{I}, \mathcal{A}),$$

where z denotes adaptive responsiveness, $\mathcal{I}$ captures institutional structure, and $\mathcal{A}$ denotes the architecture of human-machine interaction. Systems with weakly developed exploratory routines and institutions optimized around frictionless prediction tend to increase α, making AI more substitutive. By contrast, systems already possessing high adaptive responsiveness, combined with architectures designed for adversarial reasoning, exploratory search, or conceptual translation, may reduce α, and under some conditions generate $\alpha < 0$.

The sign of α therefore governs whether AI systems compress or expand exploratory responsiveness. Positive values ($\alpha > 0$) correspond to convergent predictive systems that substitute for exploratory engagement and progressively reduce adaptive mobility over time, while negative values ($\alpha < 0$) correspond to exploration-enhancing systems that amplify exploratory search, conceptual traversal, adversarial reasoning, and adaptive variation. In scientific environments, for example, generative systems capable of proposing anomalous hypotheses, simulating alternative mechanisms, or identifying unexpected cross-domain connections may increase exploratory search rather than compress it. By contrast, in environments where exploratory routines remain underdeveloped, predictive systems may strongly substitute for exploratory effort because foundational reasoning and uncertainty-navigation capacities have not yet stabilized. Early overreliance on predictive assistance may therefore weaken the adaptive routines required for later exploratory responsiveness [28, 29].

Substituting into the responsiveness dynamics yields

$$\dot{z}(t) = \eta z(t)^{\gamma}(s_0 - \alpha A I_t)(1 - z(t)) - \rho z(t)\left[1 - (s_0 - \alpha A I_t)\right]. \tag{7}$$

The dynamics imply a simple but important mechanism. Under predictive systems for which $\alpha > 0$, increasing AI-assisted optimization reduces exploratory engagement, gradually lowering the adaptive responsiveness required for long-run mobility across epistemic and institutional landscapes. Systems may therefore become increasingly efficient within familiar trajectories while simultaneously losing the capacity to traverse unfamiliar conceptual regions when environments change. By contrast, sufficiently low or negative values of α may instead preserve or amplify exploratory mobility by lowering the cost of conceptual traversal, adversarial reasoning, and exploratory search. The

long-run adaptive effects of AI therefore depend critically on the interaction between responsiveness, institutional structure, and the architecture of AI assistance itself.

The model consequently generates several phenomena increasingly relevant in AI-mediated societies, including underproduction of exploratory activity, metastable trapping, hysteresis, premature convergence, attractor rigidity, path dependence, and exploration-collapse transitions. Unlike standard optimization and strategic experimentation frameworks with fixed exploration parameters [41, 42, 43], the present model allows adaptive responsiveness itself to evolve and potentially deteriorate under prolonged optimization pressure. This creates the possibility of persistent rigidity because restoring opportunities for exploration may not immediately restore exploratory mobility once responsiveness collapses. Systems may remain trapped because of accumulated stabilization, institutional lock-in, learned dependency, or diminished exploratory competence. The introduction of highly capable AI systems therefore alters not merely the efficiency of cognition, but the long-run dynamics of collective adaptation itself.

3 Dynamics of Exploratory Responsiveness

This section studies the dynamic implications of AI-assisted optimization for exploratory responsiveness and long-run adaptive capacity. The results develop progressively. We begin by characterizing the baseline effect of AI assistance on responsiveness. We then show how sufficiently large reductions in exploratory activity can generate responsiveness collapse, metastable trapping, and hysteresis. The analysis concludes by showing how decentralized optimization may underprovide the exploratory activity required to sustain long-run adaptive responsiveness in rugged epistemic environments.

Unless otherwise stated, the numerical illustrations use normalized responsiveness dynamics with parameter values chosen to illustrate the qualitative mechanisms of the framework rather than empirical calibration. Simulations use bounded responsiveness $z(t) \in [0, 1]$, baseline exploratory engagement $s_0 = 1$, responsiveness regeneration parameter $\eta = 2.5$, stabilization parameter $\rho = 1.0$, AI substitution intensity $\alpha = 0.65$, and nonlinear responsiveness parameter $\gamma = 0.75$. Rugged epistemic landscapes are generated using stylized effective potential functions of the form

$$V(x) = V_{\text{global}}(x) + V_{\text{local}}(x),$$

where $V_{\text{global}}(x)$ captures broad stabilizing structure and $V_{\text{local}}(x)$ generates irregular local basins, localized perturbations, and uneven traversal barriers. The resulting topology produces multiple locally reinforced configurations separated by barriers of varying steepness and curvature, allowing the simulations to illustrate metastable trapping, exploratory collapse, hysteresis, and path-dependent adaptive dynamics. Specific functional forms and numerical landscape specifications used in the simulations are reported in Appendix A. The qualitative results are robust to moderate parameter variation and alternative rugged landscape specifications.

The central mechanism linking these results is straightforward. AI systems improve local performance by reducing costly exploratory effort. However, exploratory activity also regenerates the adaptive responsiveness required for long-run mobility across epistemic and institutional land-

scapes. As exploratory engagement declines, systems may therefore become increasingly efficient locally while simultaneously becoming increasingly rigid globally.

3.1 Responsiveness Decline

We begin with the responsiveness dynamics

$$\dot{z}(t) = \eta z(t)^{\gamma} s_t (1 - z(t)) - \rho z(t)(1 - s_t), \tag{8}$$

where $z(t) \in [0, 1]$ denotes adaptive responsiveness, s_t denotes exploratory activity, $\eta > 0$ measures the regeneration of responsiveness through exploration, $\rho > 0$ measures erosion under stabilization pressure, and $\gamma \in (0, 1]$ governs nonlinear responsiveness regeneration. AI assistance reduces exploratory activity according to

$$s_t = s_0 - \alpha A I_t,$$

with $\alpha > 0$. Substituting yields

$$\dot{z}(t) = \eta z(t)^{\gamma} (s_0 - \alpha A I_t)(1 - z(t)) - \rho z(t) \left[1 - (s_0 - \alpha A I_t)\right]. \tag{9}$$

Proposition 1. *For any interior responsiveness level $z(t) \in (0, 1)$, an increase in AI assistance reduces the instantaneous growth rate of adaptive responsiveness whenever AI substitutes for exploratory activity.*

$$\frac{\partial \dot{z}(t)}{\partial A I_t} = -\alpha \left[\eta z(t)^{\gamma}(1 - z(t)) + \rho z(t)\right] < 0.$$

The result follows directly from the responsiveness dynamics. AI assistance reduces the exploratory activity that regenerates responsiveness while simultaneously increasing the relative force of stabilization. The implication is not that AI lowers performance in the short run. Predictive systems may substantially improve local task efficiency by reducing costly search, deliberation, and uncertainty navigation. However, the same exploratory practices that become privately unnecessary also contribute to the regeneration of adaptive responsiveness itself. Increasing optimization may therefore improve local performance while gradually reducing the exploratory mobility required for long-run adaptation across unfamiliar conceptual and institutional regions.

This distinction between local optimization and adaptive flexibility is central to the framework. Organizations automating routine analysis may improve short-run productivity because workers solve familiar problems more quickly and coordinate more efficiently within established workflows. At the same time, workers may engage less often in independent diagnosis, exploratory reasoning, adversarial deliberation, or unfamiliar problem solving. The organization consequently becomes increasingly efficient within existing routines while gradually losing the responsiveness required to adapt when environments shift unexpectedly. Similar dynamics may emerge in scientific, educational, and institutional systems. Predictive tools may make established tasks easier while simultaneously reducing engagement in the exploratory practices through which adaptive flexibility is maintained. Figure 1 illustrates this mechanism. Higher levels of AI assistance progressively lower the long-run responsiveness trajectory and can push the system toward the critical region in

which exploratory mobility deteriorates sharply.

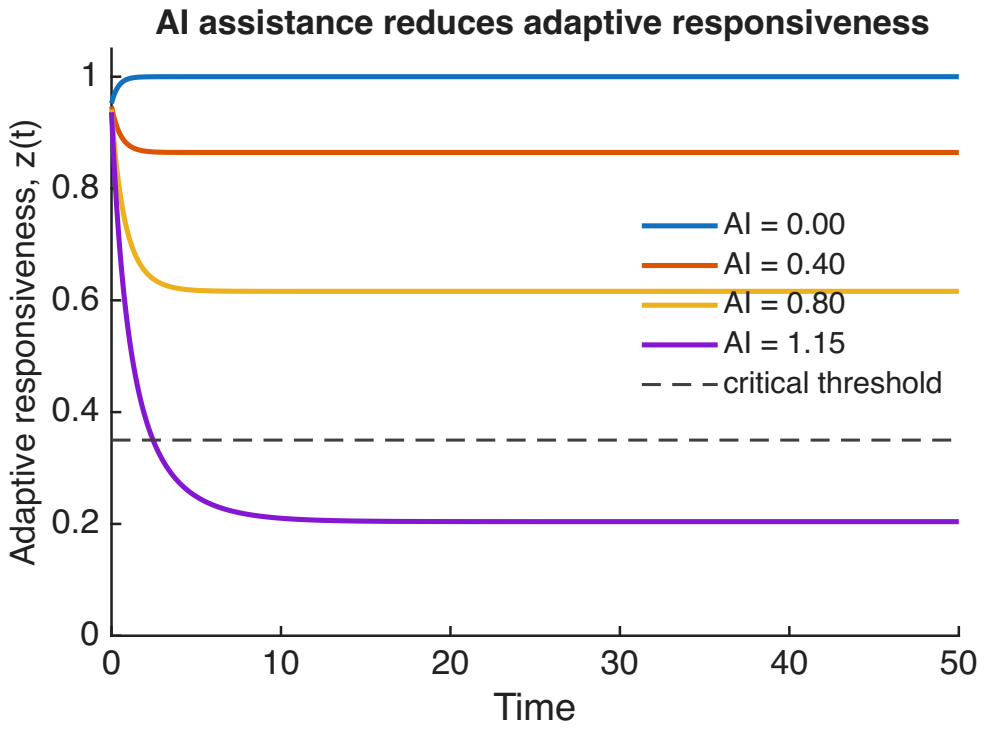


Figure 1: AI assistance reduces adaptive responsiveness by lowering exploratory engagement and weakening the regeneration of exploratory mobility. Higher levels of AI assistance progressively reduce the responsiveness stock $z(t)$, shifting the system toward increasingly rigid adaptive regimes. The dashed line denotes a critical responsiveness threshold below which exploratory mobility deteriorates sharply and systems become increasingly vulnerable to local trapping and premature convergence. Although locally optimized systems may continue performing efficiently under familiar conditions, declining responsiveness gradually reduces the capacity to traverse unfamiliar conceptual and institutional regions when environments change.

3.2 Responsiveness Collapse

The previous result establishes that AI assistance can gradually reduce adaptive responsiveness under convergent predictive regimes for which $\alpha > 0$. We now show that the resulting dynamics may become qualitatively discontinuous once exploratory engagement falls sufficiently. Define a critical responsiveness threshold

$$z_c \in (0, 1)$$

representing the minimum level of adaptive responsiveness required for sustained exploratory mobility across the landscape. Above this threshold, systems retain sufficient flexibility to traverse alternative regions of epistemic and institutional space and adapt to changing conditions. Below this threshold, exploratory mobility deteriorates sharply and systems become increasingly vulnerable to local trapping, premature convergence, and rigidity. Such thresholds become especially important in rugged environments containing multiple locally stable basins separated by steep barriers. Frontier scientific domains, rapidly changing technological sectors, and fragmented insti-

tutional environments generally require substantially greater responsiveness because systems must repeatedly escape established attractors and traverse unfamiliar regions of the landscape.

Proposition 2. *If AI-assisted optimization reduces exploratory activity sufficiently, the system converges toward a low-responsiveness regime characterized by reduced exploratory mobility and increasing rigidity.*

The proposition follows directly from the responsiveness dynamics. As AI assistance rises, exploratory activity declines according to

$$s_t = s_0 - \alpha A I_t.$$

Once exploratory engagement falls below the level required to sustain responsiveness regeneration, the system converges toward a low-mobility regime in which stabilization forces dominate exploratory renewal. The important feature of the result is that responsiveness decline is not merely gradual. Systems may continue exhibiting high local performance even while approaching collapse regions in which adaptive flexibility deteriorates rapidly. A student heavily dependent on automated problem-solving tools may continue performing well on familiar exercises while becoming increasingly unable to solve unfamiliar problems independently. Similarly, institutions optimized around predictive routines may remain highly efficient until environmental conditions shift sufficiently to require exploratory adaptation beyond existing workflows. Under sufficiently exploration-enhancing architectures ($\alpha < 0$), however, the same dynamics may instead preserve or strengthen responsiveness by increasing exploratory traversal rather than compressing it.

The distinction between optimization and adaptability is therefore central to the framework. Systems can remain locally competent even while progressively losing the capacity to respond effectively to novelty, structural change, or conceptual disruption. Figure 2 illustrates this mechanism under convergent predictive systems characterized by $\alpha > 0$. As AI assistance intensity increases, steady-state responsiveness declines. Once responsiveness crosses the critical threshold z_c, the system enters a collapse region characterized by low exploratory mobility, increasing rigidity, and declining capacity to traverse unfamiliar regions of the landscape. The effects become especially severe as landscape ruggedness rises. Systems operating over relatively smooth landscapes can remain adaptive even with modest responsiveness because nearby trajectories remain accessible. By contrast, highly rugged landscapes contain many locally stable but globally suboptimal configurations. In such environments, even moderate reductions in responsiveness can substantially increase the probability of metastable trapping and premature convergence.

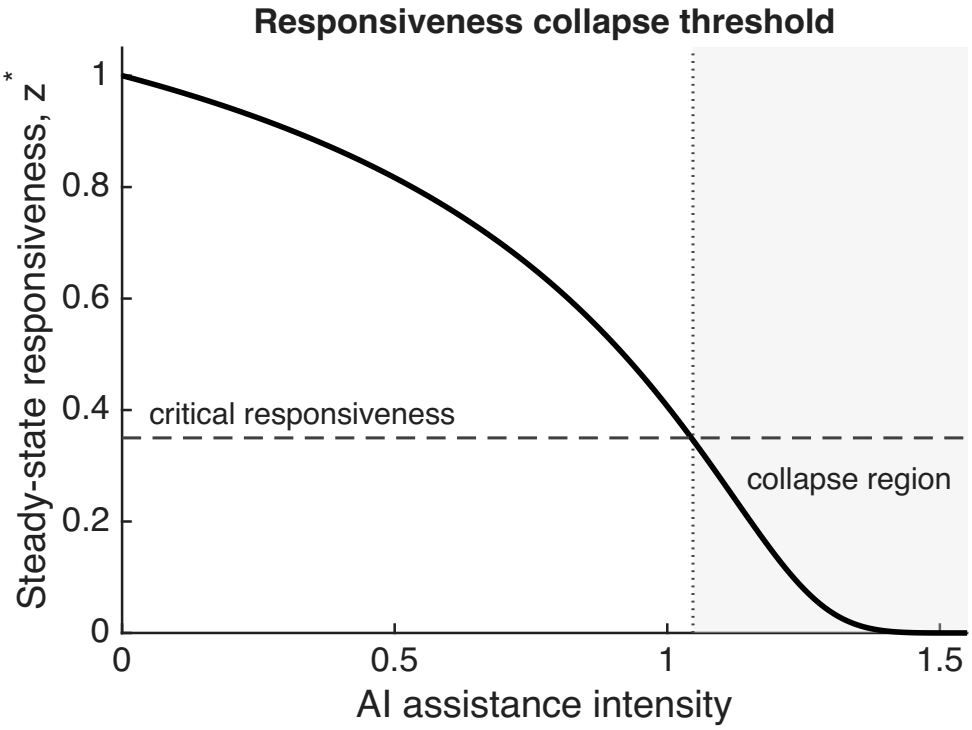


Figure 2: Responsiveness collapse threshold under convergent AI-assisted optimization ($\alpha > 0$). Steady-state adaptive responsiveness z^* declines as AI assistance intensity increases because predictive systems progressively substitute for exploratory engagement. Once responsiveness falls below the critical threshold z_c, the system enters a low-mobility regime characterized by reduced exploratory flexibility, increasing rigidity, and declining capacity to traverse unfamiliar regions of the landscape. The shaded collapse region denotes parameter configurations in which systems become increasingly vulnerable to metastable trapping, premature convergence, and path-dependent lock-in. The figure illustrates how locally efficient optimization trajectories may gradually erode the exploratory responsiveness required for long-run adaptive adaptation under changing conditions.

3.3 Metastable Trapping

The collapse of responsiveness has important implications for dynamics on rugged epistemic landscapes. The system evolves according to

$$dx_t = -z(t)\nabla V(x_t)\,dt + \sigma(t)dW_t. \tag{10}$$

When responsiveness remains high, exploratory perturbations permit trajectories to traverse multiple regions of the landscape. Systems can escape local attractors, experiment with alternative configurations, and adapt to changing conditions. Exploratory mobility therefore allows continued search across competing technological, institutional, and conceptual trajectories. Under convergent predictive regimes characterized by $\alpha > 0$, however, declining exploratory engagement progressively lowers responsiveness and constrains movement across the landscape. Transition probabilities between basins fall because exploratory mobility deteriorates, and the system becomes trapped near locally stable configurations even when superior alternatives exist elsewhere.

By contrast, exploration-enhancing systems characterized by sufficiently low or negative values of α may preserve or expand exploratory traversal by lowering the effective barriers separating alternative regions of the landscape.

Proposition 3. *For sufficiently low levels of adaptive responsiveness, the probability of escape from local attractors declines sharply, increasing metastable trapping and path dependence.*

The proposition follows from the interaction between declining responsiveness and rugged landscape geometry. Local minima that were previously traversable become effectively absorbing states once adaptive mobility deteriorates. Exploration no longer generates sufficiently large perturbations to overcome the coordination, institutional, or cognitive barriers separating competing regions of the landscape. Attractors in this context should therefore be interpreted as locally reinforced configurations toward which adaptive systems repeatedly converge. Such configurations may correspond to dominant scientific paradigms, established technological standards, organizational routines, institutional practices, or widely shared explanatory systems. These states persist because coordination effects, accumulated learning, institutional reinforcement, and adaptive expectations reduce the local costs of remaining within existing trajectories. Escaping a local basin consequently requires sufficiently high exploratory responsiveness to traverse unfamiliar regions of the landscape. When responsiveness declines, systems may continue functioning efficiently within established routines while becoming increasingly unable to adapt to conditions outside the existing basin.

Figure 3 illustrates this mechanism under convergent predictive dynamics for which declining responsiveness progressively compresses exploratory mobility. High responsiveness permits broader traversal across multiple basins and continued movement between alternative regions of the landscape, while low responsiveness confines trajectories near existing attractors, increasing local trapping and reducing adaptive flexibility.

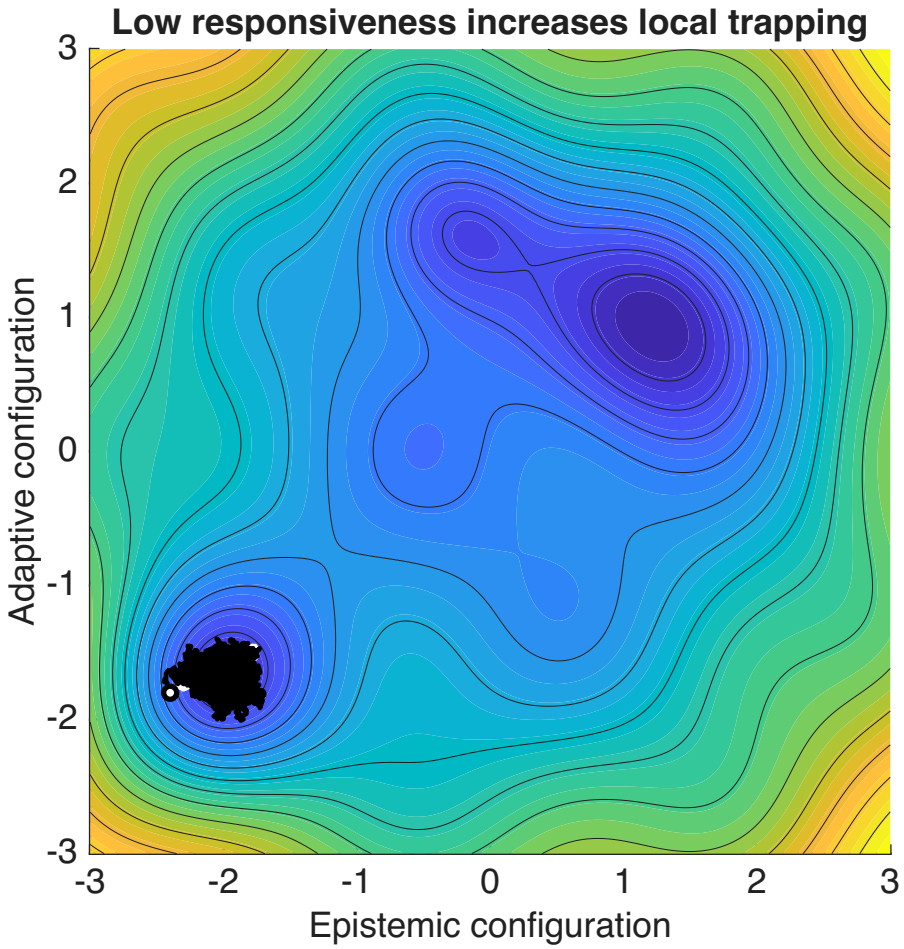


Figure 3: Low responsiveness increases local trapping on rugged epistemic landscapes. The figure illustrates a stylized rugged epistemic landscape containing multiple locally stable basins separated by steep barriers. The horizontal axes represent abstract epistemic or institutional configuration dimensions, where movement across the surface corresponds to traversal between alternative conceptual, technological, organizational, or scientific trajectories. Positive and negative values do not denote "good" or "bad" states, but different regions of the adaptive landscape. Surface height and color gradients jointly represent the effective stabilization potential $V(x)$, where darker blue valleys correspond to relatively stable and strongly reinforced configurations, while lighter elevated regions correspond to higher instability, uncertainty, coordination strain, or adaptive cost. The black trajectories illustrate system movement across the landscape under different levels of adaptive responsiveness. High responsiveness permits broader traversal across multiple basins and continued exploration of alternative configurations, while low responsiveness confines trajectories near existing attractors and reduces successful barrier crossing. As responsiveness declines, systems become increasingly trapped in locally stable but globally suboptimal regions of the landscape, generating metastable rigidity and path dependence.

This mechanism helps explain why highly optimized systems often become increasingly rigid over time. Firms that become highly optimized around existing technological trajectories can gradually lose exploratory responsiveness when environments shift. The historical difficulties experienced by Nokia during the smartphone transition illustrate how strong organizational reinforce-

ment, established routines, and accumulated coordination structures may generate locally efficient but increasingly rigid adaptive trajectories [44]. The firm remained highly adapted to an existing technological basin even as the surrounding landscape shifted toward smartphone-centered trajectories. Historical experience with automated flight-control systems illustrates a related mechanism. Modern autopilot technologies substantially improved routine operational efficiency while simultaneously reducing the frequency of manual exploratory engagement by human pilots. Under rare edge-case failures requiring rapid adaptive intervention, reduced independent responsiveness could become increasingly consequential [45, 46]. In both cases, systems remained highly efficient within familiar operational environments while becoming increasingly vulnerable to conditions requiring adaptive traversal beyond established routines.

Scientific systems may exhibit similar forms of local trapping. Mature paradigms can become resistant to conceptual revision as specialization deepens and exploratory mobility narrows. Recent evidence documenting substantial "pivot penalties" when researchers move too far from established areas of specialization suggests that scientific adaptation itself may become increasingly path dependent over time [19]. Bureaucratic systems may similarly continue reproducing inefficient routines despite mounting evidence of dysfunction, while technological ecosystems may converge toward highly standardized but increasingly brittle trajectories once coordination effects and institutional reinforcement become sufficiently strong. The central point is therefore not that optimization itself is harmful. Rather, prolonged optimization under strongly substitutive AI regimes may gradually reduce the exploratory responsiveness required to escape local attractors when adaptation becomes necessary. Recent work on the "jagged technological frontier" similarly suggests that AI systems may substantially improve performance within some regions of a task space while simultaneously degrading performance outside familiar operational domains [23]. Landscape ruggedness amplifies these effects substantially because adaptive systems operating in fragmented and rapidly changing environments must repeatedly traverse steep barriers separating competing local basins. Even moderate reductions in responsiveness can therefore generate large increases in trapping probability once landscapes become sufficiently rugged.

3.4 Hysteresis and Irreversible Rigidity

The nonlinear responsiveness dynamics naturally generate hysteresis because responsiveness regeneration depends on existing responsiveness itself. Collapse and recovery therefore need not follow identical trajectories. Systems experiencing prolonged periods of low exploratory engagement may remain trapped in low-responsiveness states even after exploratory opportunities partially recover.

Proposition 4. *Temporary periods of reduced exploratory activity can generate persistent rigidity even after exploratory opportunities are restored.*

The intuition follows directly from the responsiveness dynamics

$$\dot{z}(t) = \eta z(t)^{\gamma} s_t (1 - z(t)) - \rho z(t)(1 - s_t). \tag{11}$$

Responsiveness regeneration depends jointly on exploratory activity and existing responsiveness. Once responsiveness deteriorates sufficiently, future regeneration becomes increasingly diffi-

cult because the system loses part of the exploratory capacity required to rebuild responsiveness itself. Collapse and recovery therefore follow asymmetric trajectories, generating path dependence in adaptive mobility. Restoring opportunities for exploration may not immediately restore exploratory responsiveness because recovery depends not only on the availability of exploratory engagement, but also on whether the institutional, cognitive, and organizational structures required for exploratory regeneration remain intact.

Convergent predictive regimes ($\alpha > 0$)**.** Under convergent predictive regimes, AI assistance substitutes for exploratory engagement by reducing the need for independent search, deliberation, disagreement, and uncertainty navigation. In these environments, higher AI assistance lowers s_t, weakens responsiveness regeneration, and gradually pushes the system toward a low-mobility region. The resulting rigidity is not merely a temporary performance loss. Once exploratory routines, public reasoning practices, or institutional memory deteriorate, the system may lose part of the capacity required to regenerate responsiveness in the future. Collaborative knowledge systems provide a useful illustration. Recent empirical work documents declining public contribution activity following the diffusion of generative AI systems on collaborative platforms such as Wikipedia [16]. Once decentralized exploratory participation, mentorship pipelines, and publicly observable reasoning practices weaken, responsiveness may not rapidly recover even if reliance on AI systems later declines. Lost exploratory routines and reduced institutional memory can therefore generate delayed regeneration of adaptive capacity.

The parameter γ governs the severity of these persistence effects. Lower values imply that responsiveness remains partially self-reinforcing even at relatively low levels, allowing systems to recover despite periods of stabilization or reduced exploration. Higher values imply that responsiveness regeneration depends more strongly on already-existing exploratory capacity. In such environments, prolonged optimization can generate deeper rigidity because exploratory competence itself deteriorates over time. Scientific fields characterized by strong increasing returns, accumulated specialization, and institutional reinforcement may therefore exhibit high-γ dynamics [24]. The historical competition between VHS and Sony Betamax illustrates how early coordination advantages can generate increasingly self-reinforcing trajectories once adaptive flexibility narrows [24]. Recent evidence documenting substantial "pivot penalties" when researchers move too far from established areas of specialization similarly suggests that scientific adaptation itself may become increasingly path dependent over time [19]. More broadly, similar persistence effects may emerge across many adaptive domains. Scientific communities that lose traditions of exploratory inquiry may not rapidly recover them even after incentives for innovation improve. Workers heavily dependent on automated systems may experience persistent skill deterioration even after automation intensity declines, while institutions optimized around narrow predictive routines may remain rigid long after environmental conditions change.

Exploration-enhancing regimes ($\alpha < 0$)**.** The same formal structure also identifies a contrasting regime in which AI assistance complements rather than substitutes for exploration. When $\alpha < 0$, increases in AI assistance raise effective exploratory engagement rather than reducing it. In this regime, AI systems do not merely increase productivity within established trajectories. They

preserve or expand the exploratory mobility required to traverse rugged epistemic landscapes under changing conditions. Such systems may lower the cost of conceptual traversal, generate adversarial hypotheses, expose users to alternative explanations, simulate unfamiliar mechanisms, or make distant domains more navigable. The relevant distinction is therefore not between AI and non-AI systems, but between architectures that compress exploration and architectures that expand the system's capacity to explore.

This contrast reflects the responsiveness-dependent nature of the effective substitution parameter itself. More generally,

$$\alpha = \alpha(z, \mathcal{I}, \mathcal{A}),$$

where z denotes the existing stock of adaptive responsiveness, $\mathcal{I}$ denotes the surrounding institutional environment, and $\mathcal{A}$ denotes the architecture of AI assistance. Systems with weakly developed exploratory routines and institutions optimized around frictionless prediction tend to generate positive values of α, making AI assistance substitutive. By contrast, systems already possessing high adaptive responsiveness, embedded in institutions that preserve exploratory drafting, adversarial review, decentralized experimentation, or Socratic interaction, may reduce α and under some conditions generate $\alpha < 0$. In this sense, maintaining responsiveness is itself part of the solution: preserving exploratory capacity changes the way AI assistance enters the dynamics.

Figure 4 illustrates these contrasting adaptive regimes. The left panel shows hysteresis in the endogenous responsiveness stock under convergent predictive systems for which AI assistance substitutes for exploratory engagement ($\alpha > 0$). AI assistance is gradually increased, held high, and then reduced. Because responsiveness regeneration depends on existing responsiveness itself, the recovery trajectory differs from the collapse trajectory. Even after AI intensity declines, the system remains trapped in a lower-responsiveness region because responsiveness itself has deteriorated. The right panel illustrates an exploration-enhancing architecture in which AI systems complement rather than substitute for exploratory activity ($\alpha < 0$). In this regime, increasing AI assistance strengthens adaptive responsiveness by increasing experimentation, exploratory search, adversarial reasoning, and conceptual traversal rather than compressing it. The comparison highlights that persistent rigidity is not an inevitable consequence of AI assistance itself, but depends critically on whether AI systems reinforce convergent optimization or preserve exploratory mobility across rugged epistemic landscapes. Although the hysteresis trajectories are illustrated using the stylized responsiveness dynamics, the underlying persistence effects emerge from the interaction between responsiveness and traversal across rugged epistemic landscapes characterized by locally reinforced basins and uneven adaptive barriers.

The appendix provides an additional robustness illustration using direct stochastic traversal over a generic rugged epistemic landscape (Figure 6). Even without imposing a specialized responsiveness potential, declining responsiveness generates asymmetric recovery, reduced basin escape, and persistent reductions in exploratory mobility. The resulting trajectories are noisier than the bounded responsiveness dynamics shown here, but produce qualitatively similar hysteresis and path-dependent trapping effects. This suggests that the persistence dynamics arise not only from the specific responsiveness equation itself, but more generally from the interaction between declining exploratory mobility and rugged adaptive geometry.

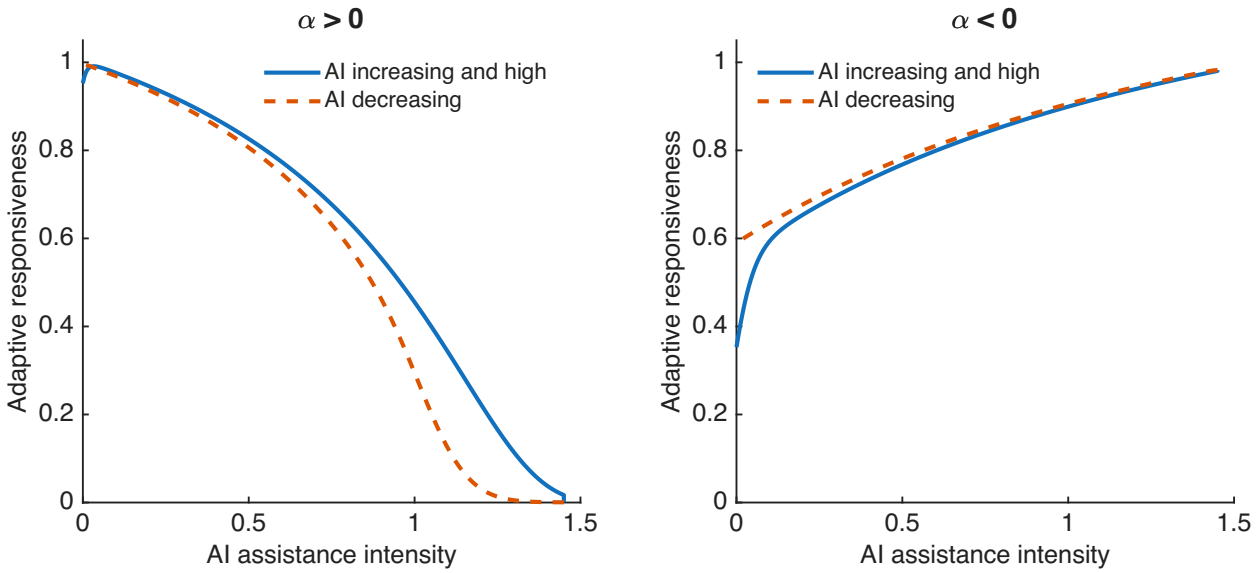


Figure 4: Hysteresis and the architecture of AI-assisted exploration. The left panel ($\alpha > 0$) illustrates hysteresis under convergent predictive systems for which AI assistance substitutes for exploratory engagement and progressively compresses adaptive responsiveness. AI assistance is gradually increased, held high, and then reduced. Because responsiveness regeneration depends on previously accumulated responsiveness itself, recovery follows a different trajectory from collapse. Even after AI intensity declines, the system remains trapped in a lower-responsiveness region, illustrating path dependence, delayed recovery, and persistent adaptive rigidity. The right panel ($\alpha < 0$) illustrates an alternative exploratory architecture in which AI assistance complements rather than substitutes for exploration. In this regime, AI systems expand exploratory engagement, adversarial reasoning, and conceptual traversal, strengthening adaptive responsiveness rather than compressing it. The comparison highlights that exploration collapse is not an inevitable consequence of AI assistance itself, but depends critically on whether AI systems reinforce convergent optimization or preserve exploratory mobility across rugged epistemic landscapes.

The framework therefore generates endogenous irreversibility in the convergent predictive regime, but also identifies the conditions under which such irreversibility may be mitigated. Adaptive rigidity can persist because responsiveness itself becomes increasingly difficult to regenerate once collapse occurs [47]. Restoring opportunities for exploration may not immediately restore exploratory mobility because systems remain constrained by accumulated stabilization, institutional lock-in, learned dependency, or diminished exploratory competence. Yet the same dynamics also imply that institutions can reduce rigidity by preserving the exploratory routines that keep α low or negative. This result distinguishes the present framework from standard optimization models with fixed exploration parameters. Here, exploratory responsiveness evolves endogenously and may deteriorate persistently under prolonged optimization pressure, but AI architecture and institutional design can alter whether optimization compresses or expands the exploratory engagement required for future adaptation.

3.5 Welfare and the Underproduction of Exploration

The preceding results imply a welfare problem arising from decentralized optimization under AI-assisted prediction. Individual agents adopt AI systems because predictive assistance improves local performance and reduces costly exploratory effort. However, exploratory activity also generates socially valuable adaptive responsiveness that decentralized agents do not fully internalize. The relevant distinction is therefore not merely between effort and convenience, but between obtaining the result of exploration and preserving the capacity to explore itself. Predictive systems may deliver high-quality answers, summaries, code, recommendations, or strategies without requiring agents to traverse the underlying problem space independently. Local output improves, but the behavioral, cognitive, and institutional routines through which exploratory responsiveness is regenerated may gradually deteriorate. A system may therefore accumulate useful answers while simultaneously losing the adaptive mobility required to generate new answers when environments change.

Let aggregate welfare be given by

$$W_t = P(AI_t) + \beta z(t)^{\nu}, \tag{12}$$

where $P(AI_t)$ captures local productivity gains from AI assistance, $\beta > 0$ measures the social value of adaptive responsiveness, and $\nu > 1$ captures the idea that highly responsive systems are disproportionately valuable for long-run adaptation. The nonlinear responsiveness term reflects the fact that adaptive losses may become increasingly severe once responsiveness deteriorates sufficiently. Small reductions in exploratory responsiveness may generate modest welfare losses, while deeper collapse can produce disproportionately large reductions in adaptive flexibility. Responsiveness consequently becomes especially valuable in environments characterized by uncertainty, structural change, and high landscape ruggedness, where long-run adaptation requires sustained traversal across competing local basins.

Proposition 5. *When exploratory activity generates positive responsiveness externalities, decentralized AI adoption leads to underinvestment in socially valuable exploration.*

The proposition follows because private agents internalize the local productivity gains from AI assistance while only partially internalizing the long-run value of exploratory responsiveness. A worker, student, firm, or platform user may rationally rely on predictive assistance because it reduces immediate effort and improves current output. Yet the foregone exploratory activity also reduces public reasoning, skill formation, institutional memory, conceptual diversity, and collective problem-solving capacity. These losses are partly external to the individual decision-maker. The resulting inefficiency therefore differs from a standard productivity tradeoff. The issue is not simply that agents choose too much convenience or too little effort. Rather, decentralized optimization may gradually erode the system-level capacity required to navigate future uncertainty.

In convergent predictive regimes characterized by strongly positive values of α, the long-run equilibrium may approach a low-responsiveness state in which systems retain high local throughput while progressively losing exploratory mobility itself. Such systems may continue generating predictions, recommendations, and locally efficient outputs, yet gradually lose the decentralized exploratory routines required for conceptual revision, institutional adaptation, and traversal across

unfamiliar regions of the landscape. In this sense, the relevant welfare problem concerns not only the quantity of information produced, but whether adaptive systems preserve the responsiveness required to escape local basins when environments shift. Highly optimized systems may therefore converge toward socially brittle equilibria characterized by strong local efficiency but weak adaptive resilience.

Figure 5 illustrates the resulting divergence between private and socially optimal AI adoption under convergent predictive dynamics. The private objective initially rises with AI assistance because predictive systems improve local productivity and reduce costly exploratory effort. Social welfare also initially increases because moderate AI assistance improves performance without substantially degrading responsiveness. Beyond a critical point, however, additional optimization reduces exploratory responsiveness sufficiently to lower long-run adaptive capacity. The socially optimal level of AI adoption therefore occurs before the private optimum, while excessive optimization pushes the system toward the collapse region.

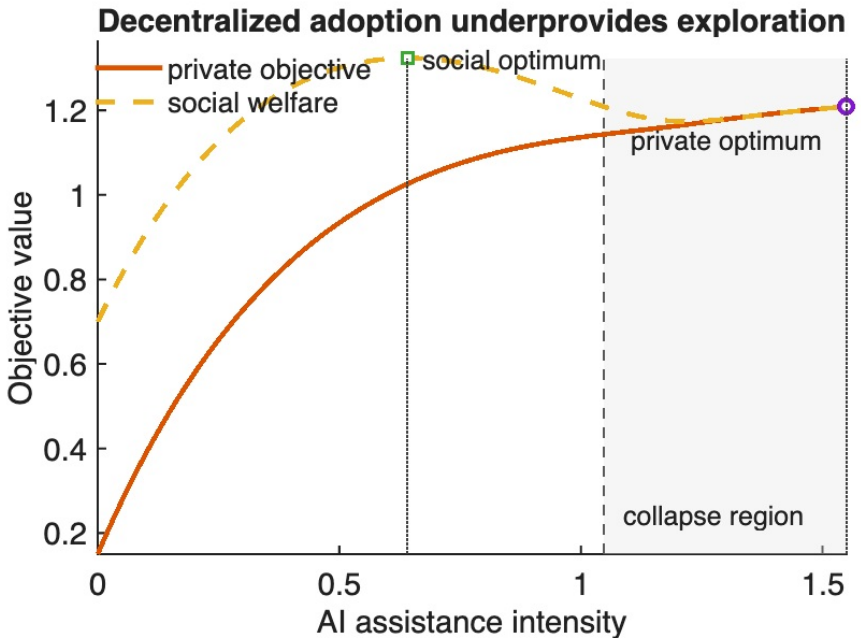


Figure 5: Decentralized AI adoption underprovides socially valuable exploration under convergent predictive regimes ($\alpha > 0$). The private objective initially rises with AI assistance because predictive systems improve local productivity and reduce costly exploratory effort. Social welfare also initially increases but peaks earlier because adaptive responsiveness generates positive externalities that decentralized agents do not fully internalize. Beyond the collapse threshold, excessive optimization reduces exploratory responsiveness and lowers long-run adaptive capacity, pushing the system toward increasingly rigid and locally trapped equilibria.

Collaborative knowledge systems illustrate the mechanism clearly. When users obtain private answers from generative systems rather than posting questions, explaining reasoning, correcting errors, and debating solutions in public forums, immediate problems may be solved more efficiently. Yet the public exploratory process through which collective knowledge is updated becomes weaker.

Recent evidence documenting reductions in public knowledge sharing and contribution activity following generative AI diffusion is consistent with this mechanism [15, 16]. The resulting inefficiency resembles a public-good problem, but with an important difference. The externality concerns not only knowledge production, but also the regeneration of adaptive mobility itself. Societies confronting uncertain and rapidly changing environments require more than accumulated information. They require systems capable of traversing unfamiliar conceptual and institutional regions under changing conditions.

The framework also identifies the conditions under which these dynamics may be mitigated or even reversed. The effective substitution parameter is itself responsiveness-dependent:

$$\alpha = \alpha(z, \mathcal{I}, \mathcal{A}).$$

Systems possessing weak exploratory routines and institutions optimized around frictionless prediction tend to generate strongly positive values of α, making AI assistance increasingly substitutive over time. By contrast, institutions preserving exploratory drafting, adversarial review, decentralized experimentation, public reasoning, and human-first diagnosis may sustain lower or even negative values of α. In such environments, AI systems may complement rather than substitute for exploratory engagement by lowering the cost of conceptual traversal, exposing agents to anomalous hypotheses, and expanding exploratory variance across distant epistemic regions. The relevant policy problem is therefore not whether AI systems should be adopted in general, but how institutional environments shape the responsiveness conditions under which AI systems become substitutive or exploration-enhancing.

The framework consequently suggests that sufficiently aggressive optimization may generate a paradoxical outcome. Systems become increasingly efficient under stable conditions while simultaneously becoming less capable of adapting when stability breaks down. Yet the same framework also implies that organizations, educational systems, and scientific institutions can preserve adaptive responsiveness by maintaining the exploratory structures that keep AI systems complementary to exploration rather than substitutive. Educational environments can require active reasoning before automated assistance. Organizations can preserve exploratory drafting, adversarial review, and decentralized experimentation before predictive optimization. Scientific systems can maintain public problem-solving spaces, conceptual disagreement, and exploratory search across distant domains. In the language of the model, such institutions preserve the responsiveness conditions under which α remains low or even becomes negative.

Taken together, the results describe a unified adaptive process. AI-assisted optimization reshapes exploratory engagement, which in turn alters adaptive responsiveness and the capacity of systems to traverse rugged epistemic landscapes. Under strongly convergent predictive regimes, declining responsiveness increases local trapping, rigidity, and hysteresis, potentially generating low-mobility equilibria that become increasingly difficult to escape. Under exploration-enhancing regimes, however, AI assistance may instead preserve or expand adaptive mobility by strengthening the exploratory processes through which responsiveness itself is regenerated. The long-run adaptive effects of AI therefore depend fundamentally on whether institutions and AI architectures preserve the exploratory responsiveness required for continued collective adaptation.

3.6 Exploration-Enhancing AI and Endogenous α

Most existing discussions of AI-assisted optimization implicitly focus on environments in which AI systems substitute for costly human exploration, corresponding to regimes where $\alpha > 0$. In such settings, predictive systems compress exploratory activity by providing optimized answers, recommendations, summaries, or strategic guidance that reduce the need for independent search, uncertainty navigation, exploratory drafting, and publicly observable reasoning. The baseline collapse dynamics studied in the previous sections emerge naturally under these conditions: reduced exploratory engagement lowers adaptive responsiveness, increasing rigidity, local trapping, and hysteresis over time.

The central implication of the framework is that the effective substitution parameter is responsiveness-dependent. The same AI system may therefore operate as a substitute in one environment and as a complement in another depending on the existing responsiveness stock, institutional structure, and architecture of human-machine interaction. Systems with low responsiveness are generally more vulnerable to exploratory substitution because foundational exploratory routines remain weakly developed, whereas systems already possessing strong exploratory capacity may use AI systems to expand exploratory variance, adversarial reasoning, and traversal across unfamiliar conceptual regions more effectively. This distinction becomes particularly important in environments characterized by rugged epistemic landscapes and high traversal costs. In scientific research, for example, moving across distant epistemic regions often requires overcoming substantial cognitive and institutional barriers. Recent evidence documenting "pivot penalties" when researchers move too far from established domains illustrates how specialization can progressively narrow exploratory mobility [19]. Under strongly convergent predictive systems, AI assistance may reinforce these barriers by accelerating movement toward familiar trajectories and reducing exploratory variance. Under exploration-enhancing architectures, however, AI systems may instead lower the effective cost of conceptual traversal by translating unfamiliar technical languages, generating anomalous hypotheses, exposing agents to distant conceptual connections, or facilitating adversarial reasoning across domains. In such environments, AI assistance may flatten portions of the rugged landscape itself, increasing the probability of successful traversal across previously inaccessible regions.

The responsiveness dependence of α also helps explain why the same AI system may generate qualitatively different outcomes across developmental environments. In educational settings where exploratory routines remain weakly developed, predictive systems may strongly substitute for exploratory effort because foundational reasoning and uncertainty-navigation capacities have not yet stabilized. Early overreliance on predictive assistance may therefore weaken the development of exploratory routines required for later adaptive flexibility [9, 14]. By contrast, systems already possessing high adaptive responsiveness may use AI assistance to generate adversarial hypotheses, explore alternative conceptual trajectories, and amplify exploratory search across distant epistemic regions. At the same time, recent evidence suggests that AI systems increasing short-run scientific productivity may also narrow conceptual diversity and contract exploratory focus under some institutional conditions [48]. Whether AI assistance ultimately amplifies or compresses exploration therefore depends critically on the responsiveness state of the surrounding system.

Institutional structure consequently becomes central to the long-run dynamics. Systems de-

signed primarily around frictionless optimization and immediate answer generation naturally increase α, making AI assistance increasingly substitutive over time. By contrast, institutions preserving exploratory drafting, adversarial review, decentralized experimentation, Socratic interaction, and human-first reasoning may sustain lower or even negative values of α. The relevant distinction is therefore not between AI and non-AI systems, but between architectures and institutions that compress exploratory mobility and those that preserve or expand it.

Figure 4 illustrates these contrasting adaptive regimes. Convergent predictive systems characterized by $\alpha > 0$ generate responsiveness decline and hysteresis under prolonged optimization pressure. Exploration-enhancing systems characterized by $\alpha < 0$, by contrast, strengthen adaptive responsiveness by increasing exploratory engagement, conceptual traversal, and adaptive variation. The comparison highlights that exploration collapse is not an inevitable consequence of AI assistance itself. Rather, the long-run adaptive effects of AI depend fundamentally on whether institutions and AI architectures preserve the exploratory responsiveness required for continued traversal across rugged epistemic landscapes.

3.7 Empirical Interpretation and Measurement

Although the framework is stylized, its central variables admit natural empirical interpretations. The model is not intended as a calibrated description of any single institutional or cognitive system. Rather, it provides a dynamical structure linking exploratory activity, adaptive responsiveness, institutional structure, and AI-assisted optimization. The framework consequently generates observable implications that may be examined empirically across organizational, scientific, educational, and collaborative knowledge environments.

The central quantities in the model correspond to different aspects of exploratory behavior and adaptive mobility. Exploratory activity s_t captures observable engagement in experimentation, uncertainty navigation, publicly visible reasoning, exploratory search, disagreement, and adaptive variation. Adaptive responsiveness $z(t)$ captures a more latent capability: the retained ability of a system to traverse unfamiliar conceptual or institutional regions under changing conditions. Landscape ruggedness reflects the complexity, fragmentation, and instability of the surrounding adaptive environment, while the effective substitution parameter

$$\alpha = \alpha(z, \mathcal{I}, \mathcal{A})$$

captures the extent to which AI systems substitute for or complement exploratory activity depending on the existing responsiveness stock z, institutional environment $\mathcal{I}$, and architecture of AI assistance $\mathcal{A}$.

One natural empirical direction concerns the measurement of exploratory activity itself. In the framework, exploratory systems generate high variance because trajectories repeatedly traverse unfamiliar regions of the landscape. Observable exploratory behavior may therefore appear as elevated experimentation rates, branching structures, reversibility, disagreement, or exploratory diversity. In software development environments, exploratory activity may be reflected in the variance and entropy of code modifications, branching behavior, exploratory commits, or the frequency of dis-

carded developmental trajectories before and after AI integration. Highly exploratory workflows generate irregular and decentralized search patterns, while strongly optimized workflows converge more rapidly toward standardized local solutions.

Entropy-based measures may provide especially useful proxies because exploratory systems tend to generate greater diversity and unpredictability in adaptive trajectories. In collaborative knowledge environments, exploratory engagement may therefore be approximated through the entropy of editing patterns, semantic diversity of contributions, branching variance in discussion networks, or the ratio of original drafting activity relative to post-hoc optimization and editing. The framework predicts that strongly substitutive AI systems ($\alpha > 0$) should gradually reduce such exploratory variance over time, while exploration-enhancing systems ($\alpha < 0$) may instead preserve or amplify exploratory diversity by increasing conceptual traversal and adversarial variation.

Adaptive responsiveness $z(t)$ represents a more latent capability because it concerns retained exploratory mobility rather than immediately observable output. Empirically, responsiveness may therefore be inferred indirectly through performance under domain shifts, unfamiliar environments, or unexpected perturbations. Systems with high responsiveness should retain greater capacity to adapt when landscapes change, while systems with low responsiveness should exhibit increasing rigidity, adaptation lag, narrowing exploratory search, and local trapping. One possible empirical proxy concerns cross-domain traversal costs. Recent work documenting substantial "pivot penalties" when researchers move across distant scientific domains suggests that exploratory mobility may become increasingly constrained as specialization deepens [19]. Similar dynamics may appear in organizational environments. Firms exhibiting strong local optimization may continue performing efficiently within established technological regimes while experiencing increasing difficulty adapting to rapidly changing environments. The historical difficulties experienced by Nokia during the smartphone transition illustrate how highly optimized organizational trajectories may remain locally functional while adaptive responsiveness to emerging technological landscapes deteriorates [44, 23].

The framework also generates measurable implications for responsiveness collapse and metastable trapping. Systems approaching low-responsiveness regimes should exhibit declining exploratory variance, increasing convergence toward standardized routines, reduced willingness to traverse unfamiliar conceptual regions, and growing sensitivity to environmental shocks. Organizations may display increasing adaptation delays under novel market conditions. Scientific systems may exhibit narrowing conceptual diversity and declining exploratory mobility across fields. Educational systems heavily optimized around predictive assistance may exhibit reduced independent problem-solving persistence under unfamiliar tasks. The hysteresis dynamics similarly imply that collapse and recovery should follow asymmetric trajectories. Systems experiencing prolonged periods of exploratory decline may therefore fail to recover rapidly even after optimization pressure weakens. Collaborative knowledge environments may again provide useful empirical settings. If publicly observable exploratory routines, mentorship pipelines, and decentralized participation deteriorate after generative AI diffusion, contribution recovery may remain persistently weak even if reliance on AI systems later declines.

The responsiveness dependence of α also generates a broader empirical research agenda centered

on institutional structure and AI architecture themselves. Systems optimized primarily around convergence, recommendation, and immediate answer generation correspond to environments with strongly positive values of α, where AI systems substitute for exploratory engagement. Such systems should reduce exploratory variance and increase convergence toward existing local attractors. By contrast, exploration-enhancing systems may generate low or negative values of α by lowering the cost of traversing unfamiliar conceptual regions. Open-ended simulation systems, adversarial research assistants, Socratic tutoring architectures, and generative scientific tools designed to produce anomalous hypotheses or unexpected cross-domain connections may therefore preserve or amplify exploratory mobility rather than compress it. The framework consequently predicts that systems optimized for convergent prediction and systems optimized for exploratory variance should generate fundamentally different long-run adaptive dynamics even when both improve short-run performance.

More broadly, the responsiveness dynamics evolve continuously over time, whereas many observable exploratory behaviors occur discretely and episodically. Exploratory activity may therefore be interpreted as a latent continuous process generated through the accumulation of discrete exploratory events such as anomalous code commits, exploratory edits, branching structures, public reasoning actions, unresolved discussion threads, or experimentally divergent trajectories [34, 39, 40]. Empirical implementation consequently requires a mapping between discrete observable events and the underlying continuous responsiveness dynamics represented by s_t and $z(t)$. One possible approach is to model exploratory activity as a rolling time integral or exponentially weighted moving average of exploratory event intensity. More generally, discrete exploratory actions may be interpreted as realizations of an underlying stochastic arrival process whose intensity reflects latent exploratory engagement [49]. Under this interpretation, observable exploratory events represent partial manifestations of a deeper adaptive process rather than direct one-to-one measurements of responsiveness itself.

The responsiveness parameters themselves also admit empirical interpretation. The regeneration parameter η may capture the effectiveness with which exploratory activity translates into future adaptive mobility, while the erosion parameter ρ captures the speed with which routinization, dependency, institutional lock-in, or excessive predictive optimization degrade exploratory responsiveness once exploratory engagement declines. The nonlinear parameter γ governs the persistence of exploratory regeneration itself and may therefore reflect the extent to which future adaptation depends on accumulated exploratory capacity, institutional memory, or established exploratory routines. More generally, the empirical challenge is not merely measuring information production, but measuring whether systems retain the exploratory responsiveness required to traverse unfamiliar conceptual and institutional landscapes under changing conditions.

4 Discussion and Conclusion

This paper develops a theory of exploratory adaptation under AI-assisted optimization. The central argument is not that AI systems inevitably reduce human adaptability, but that the long-run adaptive effects of AI depend critically on how predictive assistance interacts with exploratory

responsiveness itself. Exploratory activity generates socially valuable adaptive responsiveness because experimentation, uncertainty navigation, disagreement, publicly observable reasoning, and self-guided problem solving regenerate the mobility required to traverse rugged cognitive and institutional landscapes under changing conditions. Systems confronting uncertainty, fragmentation, and rapid structural change therefore require more than accumulated information. They require sufficient responsiveness to escape local attractors, revise existing routines, and traverse unfamiliar conceptual regions.

The framework studies these dynamics through the interaction between exploratory activity, adaptive responsiveness, institutional structure, and AI architecture. In convergent predictive regimes characterized by strongly positive values of α, AI assistance substitutes for exploratory engagement by reducing the need for independent search, deliberation, and uncertainty navigation. Under such conditions, increasing optimization lowers exploratory activity, gradually reducing adaptive responsiveness and increasing rigidity over time. The resulting dynamics generate metastable trapping, hysteresis, premature convergence, and exploration-collapse transitions in which systems remain locally efficient while becoming globally brittle. Rugged epistemic landscapes amplify these effects because fragmented and rapidly changing environments require sustained exploratory mobility to avoid becoming trapped near locally reinforced configurations.

The framework, however, also identifies a contrasting regime that is largely absent from existing automation-oriented literatures. The effective substitution parameter is itself endogenous:

$$\alpha = \alpha(z, \mathcal{I}, \mathcal{A}),$$

where z denotes the existing responsiveness stock, $\mathcal{I}$ denotes institutional structure, and $\mathcal{A}$ captures the architecture of AI assistance. The same AI system may therefore operate as a substitute in one environment and as a complement in another. Systems possessing weakly developed exploratory routines and institutions optimized around frictionless prediction tend to generate positive values of α, making AI assistance increasingly substitutive over time. By contrast, systems already possessing high adaptive responsiveness, combined with institutions preserving exploratory drafting, adversarial review, decentralized experimentation, and public reasoning, may reduce α and under some conditions generate $\alpha < 0$.

This distinction is central because it implies that exploration collapse is not an inevitable consequence of AI assistance itself. Exploration-enhancing systems may instead preserve or expand adaptive mobility by lowering the cost of conceptual traversal, exposing agents to anomalous hypotheses, facilitating adversarial reasoning, or enabling movement across previously inaccessible epistemic regions. Socratic tutoring systems, adversarial research assistants, open-ended simulation environments, and generative scientific tools capable of translating across disciplinary boundaries may therefore strengthen exploratory responsiveness rather than compress it. In highly rugged environments characterized by strong traversal barriers, such systems may increase the capacity of organizations, scientific communities, and educational systems to adapt under changing conditions.

The framework consequently differs from existing literatures emphasizing informational substitution, labor displacement, or knowledge collapse alone. Existing work on cognitive offloading demonstrates that external systems reducing immediate cognitive demand can weaken independent

memory formation and exploratory reasoning [8, 9, 11]. Research on AI assistance similarly shows that predictive systems improving local productivity may simultaneously reduce skill formation and adaptive learning [12, 13]. Collaborative knowledge systems further document reductions in public knowledge generation following generative AI diffusion [15, 16], while recursive dependence on AI-generated outputs may reduce informational diversity itself [17]. The present paper builds on these concerns but extends the analysis from knowledge production to exploratory mobility itself. The relevant question is not only whether societies continue producing information, but whether they preserve the responsiveness required to traverse unfamiliar conceptual and institutional landscapes when environments change.

The framework also differs from standard optimization models in economics and machine learning. Conventional approaches typically treat exploration as an exogenous parameter or temporary inefficiency. Here, exploratory responsiveness evolves endogenously and may either deteriorate or strengthen depending on the interaction between AI architecture, institutional structure, and existing responsiveness. This creates the possibility of multiple adaptive regimes. Under strongly convergent predictive dynamics, systems may gradually approach low-responsiveness equilibria characterized by high local throughput but weak adaptive resilience. Under exploration-enhancing regimes, by contrast, AI assistance may preserve or expand exploratory mobility by strengthening the exploratory processes through which responsiveness itself is regenerated.

These results have important implications for institutional design. If exploratory activity generates socially valuable responsiveness externalities, then environments optimized purely around immediate efficiency may become dynamically fragile over time. Educational systems emphasizing productive struggle, scientific systems preserving methodological pluralism, and organizations maintaining decentralized experimentation may therefore possess long-run adaptive advantages even when they appear locally less efficient in the short run. The relevant distinction is therefore not between AI and non-AI systems, but between institutions and architectures that compress exploratory mobility and those that preserve or expand it. Organizations may preserve adaptive responsiveness by maintaining structured forms of productive friction within highly optimized environments. Educational systems can require active reasoning before automated assistance. Scientific institutions can preserve decentralized experimentation, adversarial review, and publicly observable reasoning processes. Firms operating in rapidly changing environments can maintain exploratory conceptual development before predictive optimization.

Preserving exploratory responsiveness therefore becomes an institutional design problem rather than merely a technological one. Systems optimized exclusively around short-run throughput may gradually erode the exploratory routines, disagreement structures, and adaptive perturbations required for long-run mobility across rugged environments. The architecture of AI-assisted systems consequently matters fundamentally for long-run adaptation. Convergent predictive systems optimized for immediate answer generation may compress exploratory variance and accelerate local trapping, while exploration-enhancing systems may expand adaptive mobility by lowering the cost of traversing unfamiliar conceptual regions. The long-run adaptive effects of AI therefore depend fundamentally on whether institutions and AI architectures preserve the exploratory responsiveness required for continued collective adaptation.

The paper has several limitations. The framework intentionally abstracts from strategic interaction, heterogeneous agents, endogenous network formation, and political economy considerations. The responsiveness dynamics are stylized and designed to isolate the core mechanism linking optimization and exploratory adaptation. The model also does not distinguish formally between different AI architectures, although the analysis suggests that institutional structure and workflow design critically shape whether AI systems become substitutive or exploration-enhancing. These limitations point toward several important directions for future research, including heterogeneous responsiveness dynamics across agents and institutions, endogenous landscape formation, network structure and collective learning, strategic coordination, and political economy feedbacks under AI-assisted optimization.

More broadly, the framework opens a research frontier at the intersection of AI economics, adaptive systems, collective cognition, and social learning. The central question is no longer only how AI affects productivity, labor demand, or information acquisition. Increasingly, the relevant question concerns how predictive systems reshape the long-run geometry of exploration itself. By introducing adaptive responsiveness as an endogenous and socially underprovided stock generated through exploratory activity, and by linking AI-assisted optimization to rugged epistemic landscapes, metastable trapping, hysteresis, and exploration-enhancing regimes, the framework provides a unified account of how systems may either lose or preserve the adaptive mobility required for long-run collective adaptation.

References

[1] Henry Farrell, Alison Gopnik, Cosma Shalizi, and James Evans. Large ai models are cultural and social technologies. *Science*, 387(6739):1153–1156, 2025.

[2] OpenAI. Gpt-4. `https://openai.com/index/gpt-4-research/`, March 2023. Accessed January 2026.

[3] Erik Brynjolfsson, Danielle Li, and Lindsey Raymond. Generative ai at work. *The Quarterly Journal of Economics*, 140(2):889–942, 2025.

[4] Michael Chui, Eric Hazan, Roger Roberts, Alex Singla, and Kate Smaje. The economic potential of generative ai: The next productivity frontier. Report, McKinsey & Company, 2023.

[5] Daron Acemoglu and Pascual Restrepo. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542, 2018.

[6] Daron Acemoglu and Pascual Restrepo. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30, 2019.

[7] Ajay K. Agrawal, Joshua S. Gans, and Avi Goldfarb. The turing transformation: Artificial intelligence, intelligence augmentation, and skill premiums. Working Paper 31195, National Bureau of Economic Research, 2023.

[8] Evan F. Risko and Sam J. Gilbert. Cognitive offloading. *Trends in Cognitive Sciences*, 20(9):676–688, 2016.

[9] Timothy L. Dunn and Evan F. Risko. Toward a metacognitive account of cognitive offloading. *Cognitive Science*, 40(5):1080–1127, 2016.

[10] Xiao Hu, Liang Luo, and Stephen M. Fleming. A role for metamemory in cognitive offloading. *Cognition*, 193:104012, 2019.

[11] Sandra Grinschgl, Frank Papenmeier, and Hauke S. Meyerhoff. Consequences of cognitive offloading: Boosting performance but diminishing memory. *Quarterly Journal of Experimental Psychology*, 74(9):1477–1496, 2021.

[12] Michael Caosun and Sinan Aral. The augmentation trap: Ai productivity and the cost of cognitive offloading. *arXiv preprint arXiv:2604.03501*, 2026.

[13] Ali Aouad, Thodoris Lykouris, and Huiying Zhong. Human-ai productivity paradoxes: Modeling the interplay of skill, effort, and ai assistance. *arXiv preprint arXiv:2605.11350*, 2026.

[14] Hao-Ping Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. The impact of generative ai on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI conference on human factors in computing systems*, pages 1–22, 2025.

[15] R. Maria del Rio-Chanona, Nadzeya Laurentsyeva, and Johannes Wachs. Large language models reduce public knowledge sharing on online q&a platforms. *PNAS Nexus*, 3(9):pgae400, 2024.

[16] Liang Lyu, James Siderius, Hannah Li, Daron Acemoglu, Daniel Huttenlocher, and Asuman Ozdaglar. Wikipedia contributions in the wake of chatgpt. In *Companion Proceedings of the ACM on Web Conference 2025*, pages 1176–1179, 2025.

[17] Ilia Shumailov, Zakhar Shumaylov, Yiren Zhao, Nicolas Papernot, Ross Anderson, and Yarin Gal. Ai models collapse when trained on recursively generated data. *Nature*, 631(8022):755–759, 2024.

[18] Daron Acemoglu, Dingwen Kong, and Asuman Ozdaglar. Ai, human cognition and knowledge collapse. Technical report, National Bureau of Economic Research, 2026.

[19] Ryan Hill, Yian Yin, Carolyn Stein, Xizhao Wang, Dashun Wang, and Benjamin F Jones. The pivot penalty in research. *Nature*, 642(8069):999–1006, 2025.

[20] Douglass C. North. *Institutions, Institutional Change and Economic Performance*. Cambridge University Press, 1990.

[21] Douglass C. North. *Understanding the Process of Economic Change*. Princeton University Press, 2005.

[22] Cristina Bicchieri. *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press, 2005.

[23] Fabrizio Dell'Acqua, Edward McFowland III, Ethan Mollick, Hila Lifshitz, Katherine C Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R Lakhani. Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science*, 37(2):403–423, 2026.

[24] W. Brian Arthur. Self-reinforcing mechanisms in economics. pages 9–31, 1988.

[25] Pierre Bourdieu. *The Field of Cultural Production*. Columbia University Press, 1993.

[26] Neil Fligstein. Social skill and the theory of fields. *Sociological Theory*, 19(2):105–125, 2001.

[27] Rika Preiser, Reinette Biggs, Alta De Vos, and Carl Folke. Social-ecological systems as complex adaptive systems: Organizing principles for advancing research methods and approaches. *Ecology and Society*, 23(4):46, 2018.

[28] Eric I. Knudsen. Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, 16(8):1412–1425, 2004.

[29] Umberto León-Domínguez. Potential cognitive risks of generative transformer-based ai chatbots on higher order executive functions. *Neuropsychology*, 38(4):293–308, 2024.

[30] Shakked Noy and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192, 2023.

[31] Robert A Bjork and Elizabeth L Bjork. Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, pages 56–64. Worth Publishers, 2011.

[32] Manu Kapur. Productive failure. *Cognition and Instruction*, 26(3):379–424, 2008.

[33] Alison Gopnik, Sophie O'Grady, Christopher G Lucas, et al. Childhood as a solution to explore–exploit tradeoff. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711):20160050, 2017.

[34] Nathaniel D Daw, John P O'Doherty, Peter Dayan, Ben Seymour, and Raymond J Dolan. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–879, 2006.

[35] Nataliya Kosmyna, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes. Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task. *arXiv preprint arXiv:2506.08872*, 2025.

[36] Enrique Ide. Automation, ai, and the intergenerational transmission of knowledge. *arXiv preprint arXiv:2507.16078*, 2025.

[37] Enrique Ide and Eduard Talamàs. Artificial intelligence in the knowledge economy. *Journal of Political Economy*, 133(12):3762–3800, 2025.

[38] Zoë Cullen, Danielle Li, and Shengwu Li. Labor as capital: Ai and the ownership of expertise. Unpublished manuscript, 2025.

[39] Jacqueline Gottlieb, Pierre-Yves Oudeyer, Manuel Lopes, and Adrien Baranes. Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, 17(11):585–593, 2013.

[40] Kalina Christoff, Zachary C Irving, Kieran C. R. Fox, R Nathan Spreng, and Jessica R Andrews-Hanna. Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, 17(11):718–731, 2016.

[41] Patrick Bolton and Christopher Harris. Strategic experimentation. *Econometrica*, 67(2):349–374, 1999.

[42] Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

[43] Godfrey Keller and Sven Rady. Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010.

[44] Jens Wang, Jonas Hedman, and Virpi Kristiina Tuunainen. Path creation, path dependence and breaking away from the path: Re-examining the case of nokia. *Journal of Theoretical and Applied Electronic Commerce Research*, 11(2):16–27, 2016.

[45] Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile. Final report on the accident on 1st june 2009 to the airbus a330-203 registered f-gzcp operated by air france flight af 447 rio de janeiro–paris. Technical report, BEA, Le Bourget, France, 2012.

[46] Nick Oliver, Thomas Calvard, and Kristina Potočnik. The tragic crash of flight af447 shows the unlikely but catastrophic consequences of automation. *Harvard Business Review*, 2017.

[47] Nicholas Otis, Rowan Clarke, Solene Delecourt, David Holtz, and Rembrand Koning. The uneven impact of generative ai on entrepreneurial performance. 2024.

[48] Qianyue Hao, Fengli Xu, Yong Li, and James Evans. Artificial intelligence tools expand scientists' impact but contract science's focus. *Nature*, pages 1–7, 2026.

[49] Crispin Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, 2009.

Author Contributions

BB designed the research, developed the model, and wrote the manuscript.

Acknowledgements and Disclosures

The author received no financial support for the research, authorship, or publication of this article. The author declares no relevant financial relationships or other potential conflicts of interest.

A Appendix: Responsiveness Dynamics and Numerical Illustrations

This appendix provides additional details on the responsiveness dynamics and numerical illustrations used throughout the paper. The purpose is not to calibrate the framework empirically, but to clarify how the equations and landscape structure encode the central mechanism of the model. Exploratory activity regenerates adaptive responsiveness, while reduced exploration and stabilization pressures gradually erode the mobility required for long-run traversal across rugged epistemic and institutional landscapes.

A.1 Socioeconomic Interpretations of Rugged Landscapes

The baseline framework uses stylized rugged landscapes to represent the interaction between stabilization pressures and exploratory mobility. The purpose is not to model any single cognitive, organizational, or institutional environment directly, but to characterize the general adaptive geometry governing movement across multiple locally reinforced configurations. Nevertheless, the effective potential function $V(x)$ may be interpreted more concretely as a composite representation of socioeconomic and cognitive frictions constraining movement across alternative trajectories.

One possible decomposition is

$$V(x) = \lambda_C C(x) + \lambda_K K(x) + \lambda_I I(x), \tag{13}$$

where $C(x)$ captures coordination costs associated with deviating from established standards, routines, or conventions, $K(x)$ captures cognitive strain or learning effort associated with unfamiliar conceptual structures, and $I(x)$ captures institutional switching costs generated by accumulated infrastructure, organizational routines, legitimacy constraints, or path dependence. The weights $\lambda_C, \lambda_K, \lambda_I \geq 0$ determine the relative importance of each friction in shaping the effective landscape.

Under this interpretation, local minima correspond to configurations minimizing combined coordination, cognitive, and institutional strain. Steep barriers separating basins represent retraining burdens, organizational switching costs, legitimacy pressures, accumulated infrastructure dependence, or cognitive effort associated with traversing unfamiliar conceptual regions. Ruggedness therefore emerges when multiple stabilization mechanisms reinforce competing local configurations simultaneously.

In technological environments, ruggedness may arise through network externalities, installed infrastructure, retraining costs, and standardization pressures. The historical competition between VHS and Sony Betamax illustrates how early coordination advantages can generate self-reinforcing barriers that progressively constrain alternative trajectories [24]. In scientific environments, ruggedness may reflect cognitive specialization, disciplinary boundaries, and professional incentives that increase the difficulty of traversing distant epistemic regions. Recent evidence documenting substantial "pivot penalties" in scientific research is consistent with this interpretation [19]. In organizational environments, ruggedness may arise from legacy workflows, accumulated routines, managerial incentives, compliance structures, and the costs of retraining personnel around alternative

practices.

This interpretation is not intended to replace the general potential function used in the main text. Rather, it clarifies how the abstract geometry of $V(x)$ can be grounded in empirically meaningful frictions. Different empirical settings may place different weights on coordination costs, cognitive strain, and institutional switching barriers, while the general adaptive mechanism remains unchanged: stabilization pressures pull systems toward nearby low-strain configurations, while adaptive responsiveness determines whether systems retain sufficient mobility to traverse the barriers separating alternative regions of the landscape.

A.2 Numerical Landscape Specification

The numerical illustrations use stylized rugged landscapes designed to generate multiple locally stable basins and uneven traversal barriers consistent with the general framework developed in the main text. The simulations are intended to illustrate exploratory traversal, metastable trapping, responsiveness collapse, and hysteresis rather than to calibrate any particular institutional or cognitive environment.

The effective potential function used in the numerical simulations is given by

$$V(x, y) = 0.18 \sin(2.5x) \cos(2.2y) + 0.12 \sin(4.1x + 1.3y) + 0.08 \cos(3.7x - 2.4y) + 0.015(x^2 + y^2).$$

The sinusoidal components generate irregular local basins and uneven traversal barriers, while the quadratic term prevents trajectories from diverging without bound. The resulting topology produces a rugged landscape containing multiple locally reinforced configurations separated by barriers of varying height and curvature. Alternative specifications producing comparable ruggedness generate qualitatively similar trapping, collapse, and hysteresis dynamics.

A.3 Bounded Responsiveness Dynamics

The main text normalizes adaptive responsiveness so that

$$z(t) \in [0, 1].$$

The lower bound $z = 0$ represents complete rigidity. The upper bound $z = 1$ represents maximal exploratory responsiveness. This normalization is useful because responsiveness is interpreted as a stock of adaptive mobility rather than an unbounded quantity.

The responsiveness law is

$$\dot{z}(t) = \eta z(t)^{\gamma} s_t (1 - z(t)) - \rho z(t)(1 - s_t).$$

The first term captures responsiveness regeneration through exploratory activity. The factor s_t means that responsiveness increases only when the system engages in exploratory activity. The factor $(1 - z(t))$ imposes saturation. When responsiveness is already high, the marginal increase in responsiveness becomes small. This prevents the model from generating unbounded responsiveness.

The second term captures responsiveness erosion. When exploratory activity is low, stabilization, routinization, learned dependence, and institutional lock-in reduce responsiveness. The term $z(t)$ implies that erosion is proportional to the existing stock of responsiveness. A system with no remaining responsiveness cannot lose further responsiveness.

The parameter $\gamma \in (0, 1]$ governs nonlinear regeneration. It determines how existing responsiveness affects the capacity to regenerate future responsiveness. When $\gamma = 1$, responsiveness regeneration is proportional to the current responsiveness stock. When $\gamma < 1$, regeneration remains active even at relatively low responsiveness levels. This produces smooth but persistent recovery dynamics and allows collapse and recovery paths to differ.

A.4 AI Assistance and Exploratory Activity

Exploratory activity is modeled as

$$s_t = s_0 - \alpha AI_t,$$

where s_0 is baseline exploratory engagement, AI_t is AI assistance intensity, and $\alpha > 0$ measures the degree to which AI substitutes for costly exploratory activity.

Substituting into the responsiveness law yields

$$\dot{z}(t) = \eta z(t)^{\gamma}(s_0 - \alpha AI_t)(1 - z(t)) - \rho z(t)\left[1 - (s_0 - \alpha AI_t)\right].$$

This equation gives the central comparative statics. Differentiating with respect to AI_t gives

$$\frac{\partial \dot{z}(t)}{\partial AI_t} = -\alpha\eta z(t)^{\gamma}(1 - z(t)) - \alpha\rho z(t).$$

For $z(t) \in (0, 1)$ and $\alpha > 0$, this derivative is negative:

$$\frac{\partial \dot{z}(t)}{\partial AI_t} < 0.$$

Thus, increasing AI assistance lowers the growth rate of adaptive responsiveness whenever AI assistance substitutes for exploratory activity.

A.5 Steady States and Collapse Regions

For fixed AI intensity, steady states satisfy

$$0 = \eta z^{\gamma} s(1 - z) - \rho z(1 - s),$$

where

$$s = s_0 - \alpha AI.$$

The low-responsiveness region becomes relevant when s is small enough that erosion dominates regeneration. In that case, responsiveness converges toward low values and the system loses

exploratory mobility. The main text represents this by introducing a critical responsiveness level

$$z_c \in (0, 1).$$

When the long-run responsiveness level falls below z_c, the system enters a low-mobility regime.

The threshold z_c should not be interpreted as a universal constant. It is a reduced-form representation of the minimum responsiveness needed to traverse the relevant landscape. In smoother environments, the required threshold may be low. In rugged environments with many local basins and steep barriers, the required threshold is higher.

A.6 Rugged Landscapes and Metastable Trapping

The state dynamics are

$$dx_t = -z(t)\nabla V(x_t)\, dt + \sigma(t) dW_t.$$

The potential $V(x)$ represents effective stabilization pressures. Low values of $V(x)$ correspond to locally reinforced configurations. High values correspond to unstable or weakly reinforced configurations. The model does not assume that agents maximize V or directly optimize a known objective. Rather, V summarizes the aggregate effect of local updating, reinforcement, coordination, and institutional stabilization.

In the simulations, the rugged landscape is represented by a stylized two-dimensional potential,

$$V(x, y) = 0.12(x^2+y^2) - 1.40 \exp\left(-\frac{(x+2.1)^2 + (y+1.8)^2}{0.65}\right) - 1.25 \exp\left(-\frac{(x-1.4)^2 + (y-1.0)^2}{0.80}\right)$$

$$-0.95 \exp\left(-\frac{(x+0.3)^2 + (y-1.7)^2}{0.50}\right) + 0.18 \sin(2.5x)\cos(2.2y).$$

The quadratic term confines the landscape. The Gaussian terms create local basins. The sinusoidal term adds small-scale ruggedness. This functional form is not intended as a literal empirical landscape. It is a stylized object used to illustrate how low responsiveness increases local trapping.

The numerical trajectory follows the Euler discretization

$$x_{t+\Delta t} = x_t - z\nabla V(x_t)\Delta t + \sigma\sqrt{\Delta t}\, \varepsilon_t,$$

where

$$\varepsilon_t \sim N(0, I).$$

Higher z produces greater adaptive mobility. Lower z confines the trajectory near local basins.

A.7 Hysteresis Simulation

The hysteresis simulation uses the same bounded responsiveness law,

$$\dot{z}(t) = \eta z(t)^{\gamma} s_t (1 - z(t)) - \rho z(t)(1 - s_t).$$

AI assistance is varied over time. It first rises gradually, then remains high, and then declines. This creates a collapse path and a recovery path. Because responsiveness regeneration depends on existing responsiveness, the recovery path differs from the collapse path.

The simulation uses the common parameter values

$$s_0 = 1, \qquad \alpha = 0.65, \qquad \eta = 2.5, \qquad \rho = 1.0, \qquad \gamma = 0.75.$$

The AI path is

$$AI_t = \begin{cases} AI_{\max}\dfrac{\tau}{0.40}, & \tau \le 0.40, \\ AI_{\max}, & 0.40 < \tau \le 0.60, \\ AI_{\max}\left(1 - \dfrac{\tau - 0.60}{0.40}\right), & \tau > 0.60, \end{cases}$$

where

$$\tau = \frac{t}{T}.$$

The figure in the main text plots $z(t)$ against AI_t. The increasing and decreasing paths do not coincide because responsiveness is a state variable. This is the source of hysteresis.

A.8 Exploratory Mobility on Rugged Landscapes

The second panel of the hysteresis figure measures exploratory mobility on a generic rugged landscape. The simulation uses the same rugged potential $V(x, y)$ as the trapping figure. AI reduces responsiveness according to

$$z_t = \max\{0.05, 1 - 0.70 AI_t\}.$$

The lower bound prevents complete numerical freezing. The trajectory is simulated using the stochastic landscape dynamics. Exploratory mobility is measured as the rolling mean of step sizes:

$$M_t = \frac{1}{m} \sum_{j=t-m+1}^{t} \|x_j - x_{j-1}\|.$$

This mobility measure is noisier than the bounded responsiveness stock because it depends on stochastic realizations and local landscape geometry. The purpose of the panel is to show that path dependence also appears in generic rugged landscapes, even when no special double-well structure is imposed.

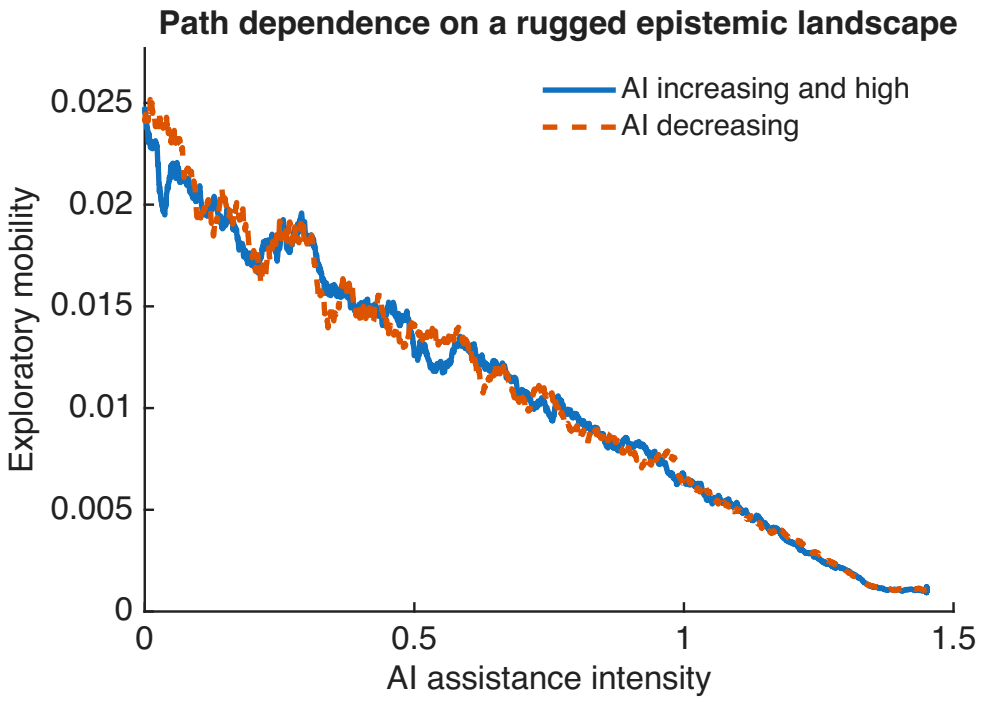


Figure 6: Path dependence in exploratory mobility on a stochastic rugged epistemic landscape. AI assistance is gradually increased, held high, and then reduced while adaptive responsiveness scales exploratory traversal across a rugged potential surface containing multiple locally stable basins and uneven traversal barriers. The trajectories illustrate stochastic movement under declining exploratory responsiveness, where reduced mobility progressively confines the system near existing attractors and lowers the probability of successful basin escape. Although noisier than the bounded responsiveness dynamics presented in the main text, the rugged landscape generates qualitatively similar persistence effects: once responsiveness deteriorates sufficiently, exploratory mobility recovers asymmetrically even after AI intensity declines. The figure therefore illustrates that metastable trapping, delayed recovery, and path-dependent rigidity emerge endogenously from rugged adaptive geometry itself rather than requiring a specialized double-well responsiveness structure.

A.9 Welfare Illustration

The welfare illustration uses

$$W_t = P(AI_t) + \beta z(t)^{\nu}.$$

The productivity term is

$$P(AI) = P_{\max}(1 - \exp(-kAI)).$$

This captures increasing but concave local productivity gains from AI assistance.

The responsiveness term is nonlinear:

$$\beta z(t)^{\nu}, \qquad \nu > 1.$$

The exponent ν captures the idea that highly responsive systems are disproportionately valuable for long-run adaptation. A system with high responsiveness can traverse unfamiliar regions of the

landscape, respond to shocks, and avoid local trapping. A system with low responsiveness may remain locally efficient but globally fragile.

The private objective assigns a lower weight to responsiveness than the social objective. This captures the externality. Individual agents internalize local productivity gains but do not fully internalize the contribution of exploratory activity to system-level adaptive responsiveness. Consequently, the private optimum generally involves greater AI assistance and lower exploration than the social optimum.

A.10 Interpretation of Numerical Illustrations

The numerical figures are illustrative rather than calibrated. They are designed to show the qualitative implications of the model. The common parameter values are chosen to make the mechanisms visually transparent. The key qualitative patterns are robust to moderate parameter variation.

The figures illustrate five linked mechanisms. First, AI assistance reduces responsiveness by substituting for exploratory activity. Second, sufficiently large reductions in responsiveness push the system toward a low-mobility collapse region. Third, low responsiveness increases trapping on rugged landscapes. Fourth, nonlinear responsiveness dynamics generate hysteresis because recovery differs from collapse. Fifth, decentralized adoption can underprovide socially valuable exploration because private agents do not fully internalize responsiveness externalities.

Together, these mechanisms formalize the main claim of the paper. AI-assisted optimization may improve local performance while reducing the exploratory responsiveness required for long-run collective adaptation.