

# Recursive Self-Improvement is a Portfolio Optimization Problem

York Westenhaver  
*AlphaFund*

Massey Branscomb  
*AlphaFund*

Aidan Grant  
*AlphaFund*

## Abstract

Recursive self-improvement is usually framed as software rewriting itself. We propose a narrower, measurable formulation: a corporation recursively improves when realized economic gains finance the next cycle of better prediction and deployment. Quantitative trading instantiates this loop with unusual precision, since decisions, costs, outcomes, and reinvestment are all digitized. We introduce the Economic World Model (EWM), a forecasting and control object scored on future realized outcomes, and summarize the firm’s standing as t-RSI, a standardized gap between alpha-creation and alpha-decay rates. We present evidence of the first general economic scaling law beyond language data; further evidence from live trading and held-out backtests support the framework. We position today’s firm as the present-moment derivative of a trajectory toward an Autonomous Self-improving Corporation (ASIC), in which capital allocation is itself executed by the firm’s software. Thus, recursive self-improvement can be reframed as an auditable capital-allocation process.

## Contents

|     |  |    |
|-----|--|----|
| 1   | Introduction                           | 3  |
| 2   | The Self-Improving Corporation         | 4  |
| 2.1 | Firm Objective and Accounting Identity | 4  |
| 2.2 | The Corporation as a Bundle of Assets  | 5  |
| 2.3 | Coupled Dynamics                       | 6  |
| 2.4 | The Self-Forecasting Loop              | 7  |
| 3   | The Economic World Model               | 8  |
| 3.1 | EWM Definition                         | 8  |
| 3.2 | Firm and Channel Histories             | 9  |
| 3.3 | LLMs are not economic world models     | 9  |
| 3.4 | Channel-Specific World Models          | 10 |
| 4   | The Portfolio Optimizer                | 11 |
| 4.1 | The Corporate Optimization Problem     | 11 |
| 4.2 | The Marginal Return Vector             | 12 |
| 4.3 | Capital Allocation                     | 13 |
| 5   | Experimental Evidence of t-RSI         | 13 |
| 5.1 | What t-RSI Measures                    | 13 |
| 5.2 | Investments as an Asset                | 15 |
| 5.3 | Sensors as an Asset                    | 16 |

|     |  |    |
|-----|--|----|
| 5.4 | Actuators as an Asset . . . . .                                  | 17 |
| 5.5 | R&D as an Asset . . . . .  | 20 |
| 5.6 | Parameters as an Asset . . . . .                                 | 23 |
| 5.7 | Continual Learning . . . . .                                     | 24 |
| 5.8 | Headline t-RSI . . . . .   | 25 |
| 6   | Toward the Differentiable Corporation . . . . .                  | 27 |
| 6.1 | Three Structural Facts . . . . .                                 | 27 |
| 6.2 | Channels Reinforce Each Other . . . . .                          | 28 |
| 6.3 | Drift Detection and Recovery . . . . .                           | 29 |
| 6.4 | The Bitter Lesson for Capital . . . . .                          | 30 |
| 6.5 | Completion Roadmap . . . . .                                     | 30 |
| 6.6 | Beyond Quant Trading . . . . .                                   | 31 |
| 7   | Conclusion . . . . .   | 32 |
| A   | Accounting Bridge . . . . .                                      | 32 |
| A.1 | Accounting Projection . . . . .                                  | 32 |
| A.2 | Deployable Capital and Flow of Funds . . . . .                   | 33 |
| B   | Program Constraints . . . . .                                    | 34 |
| B.1 | Budget, Liquidation, Liquidity, Solvency . . . . .               | 34 |
| C   | EWM Details . . . . .  | 35 |
| C.1 | EWM Training Objectives . . . . .                                | 35 |
| C.2 | From Population KL to Empirical Proper Scoring . . . . .         | 36 |
| C.3 | Why the Conditioning is on the Filtration . . . . .              | 36 |
| D   | Portfolio Optimization . . . . .                                 | 37 |
| D.1 | Shadow Price of Capital . . . . .                                | 37 |
| D.2 | Markowitz Mean–Variance Form . . . . .                           | 37 |
| D.3 | Black–Litterman Form with EWM Views . . . . .                    | 38 |
| D.4 | Multi-Period Rollout and the Sharpe-Equimarginal Limit . . . . . | 38 |
| D.5 | Receding-Horizon Use of the EWM . . . . .                        | 39 |
| E   | t-RSI Details . . . . .  | 39 |
| E.1 | t-RSI Measurement Conventions . . . . .                          | 39 |
| E.2 | Three-Month t-RSI Calculation . . . . .                          | 40 |
| F   | Channel Derivations . . . . .                                    | 42 |
| F.1 | Investments Supporting Equations . . . . .                       | 42 |
| F.2 | Sensors Supporting Equations . . . . .                           | 44 |
| F.3 | Actuators Supporting Equations . . . . .                         | 47 |
| F.4 | R&D Supporting Equations . . . . .                               | 50 |
| F.5 | Parameters Supporting Equations . . . . .                        | 55 |
| F.6 | Continual Learning Supporting Equations . . . . .                | 56 |
| G   | Data Scaling . . . . .   | 57 |
| G.1 | Direct Versus Chained Fit . . . . .                              | 57 |
| G.2 | Selection-Rule Efficient Frontier . . . . .                      | 58 |
| G.3 | Robustness: Jackknife, Cook’s Distance, Leverage . . . . .       | 58 |
| G.4 | What the Chart Says and Extrapolation Robustness . . . . .       | 58 |
| H   | Deployment Parameters . . . . .                                  | 59 |

|     |   |    |
|-----|---|----|
| I   | Improvement Certificate                       | 60 |
| I.1 | Certificate of Monotone Improvement . . . . . | 60 |

# 1 Introduction

The literature on recursive self-improvement spans decades, from Yudkowsky’s seed AI framing [79], through Schmidhuber’s Gödel Machines [65], to formal and empirical treatments of recursive self-improvement and intelligence explosion dynamics [22, 43, 75], to more skeptical analyses emphasizing diminishing returns and bottlenecks [77]. These formalisms differ in mechanism, but they share a simplifying assumption: the economic costs of the improvement step do not threaten the system’s existence. In reality, every FLOP and bit of data costs money [37, 67, 78]; self-improvement is an economic problem.

In the same way that biological organisms need resources and a suitable environment to survive and evolve [38, 50], the more complex system of the global economy needs capital to do the same [58]. A system that spends more on self-improvement than it earns from the resulting improvement slowly runs out of resources and dies. We thus define intelligence operationally: the capacity to acquire, preserve, and compound command over resources through accurate prediction. Because the environment is non-stationary and uncertain, the question shifts from *whether* RSI will occur to *whether a system can generate sufficient capital to fund each attempt at self-improvement*. This reframing casts recursive self-improvement (RSI) as a stochastic control problem under a survival constraint, with progress scored by the standardized signal-to-noise ratio of expected improvement against its posterior dispersion—a quantity we call *t-RSI*. For AlphaFund at the current operating point, the data-derived three-month t-RSI is 9.61 standardized units. It should be made clear that RSI is impossible to guarantee. The position and momentum of a particle can’t be known; a gamma-ray burst from across the universe could wipe all life out. We are proposing a concrete and measurable framework that allows researchers to talk about this hypothetical phenomenon in a grounded and empirical manner.

We develop this claim in five steps:

1. **The Self-Improving Corporation.** We formalize the corporation as stochastic optimal control under a balance-sheet survival constraint, specifying the state, transition law, objective, and capital-allocation program that convert predictions into a reinvestment process.
2. **The Economic World Model.** We decompose this theoretical problem into a practical prediction architecture and show that, under the channel coupling we make explicit, the decomposed system recovers the same first-order conditions as a joint-monolith optimum at the operating point.
3. **The Portfolio Optimizer.** We define the corporation’s controller as a model-predictive convex program over the EWM’s per-channel return forecasts, making heterogeneous interventions — a researcher hire, a data feed, a GPU, a position in AAPL — directly comparable.
4. **The RSI Portfolio.** We instantiate that controller channel by channel — investments, sensors, actuators, parameters, R&D — and present the empirical scaling and response laws that pin down each row of the marginal-return vector.

5. **Trajectory.** We study the long-run dynamics of the controller under those scaling laws—accounting for complementarity, filtration widening, and external capital—and argue that, if the preliminary laws hold under continued validation and marginal deployment keeps clearing after capacity, competition, financing, and market-impact costs, this process could plausibly capture a substantial share of financial-industry profits and serve as a bridge from quantitative trading toward broader priced economic action.

## 2 The Self-Improving Corporation

Corporations are intelligent self-improving systems. They can replace their own personnel, hardware, software, board, and even their business model and still retain their core identity—the legal name for this property is *perpetual succession*. Like the Ship of Theseus, what matters is not that any given component persists, but that the process that turns capital into improved capability and improved capability back into capital persists through continual reconfiguration. Most corporations do not, in fact, recursively improve; the question this paper asks is what conditions distinguish those that do. All for-profit corporations share the same objective: maximize shareholder value while remaining solvent [24, 54]. In service of that objective they observe their environment and internal state, allocate capital across operational capabilities, receive feedback in the form of earnings, and reinvest those earnings to augment future capability. We mean capital in the most generic sense: resources necessary to survive and improve. Dollars are generally fungible for those, and thus represent a satisfactory scalar approximation.

We model this sequential optimization process—the **Corporate Loop**—as constrained stochastic optimal control over the firm’s production cycle [10, 61, 62]. Quantitative trading is the cleanest case: its feedback loop is fast, direct, and dollar-denominated, and its production, sales, capital allocation, and reinvestment functions can all be specified precisely and measured at high frequency. The four subsections below set the four pieces the rest of the paper composes: the firm’s objective, the asset bundle the objective is computed over, the action vector that moves that bundle, and the coupled dynamics under which the bundle and the world evolve together.

### 2.1 Firm Objective and Accounting Identity

For-profit corporations exist to maximize shareholder value subject to remaining solvent [24, 54], and shareholder value is, by construction, shareholders’ equity—the residual claim on assets after liabilities are netted. The firm’s per-period reward is the realized log-return on that equity [42], the survival constraint is  $K_\tau > 0$ , and the cumulative objective  $J_t$  is the expected sum of those per-period rewards over a finite planning horizon. All later channels are priced against this  $J_t$ .

**Definition 1** (Shareholders’ equity). **Shareholders’ equity** at decision time  $t$  is the firm’s net worth: total assets minus total liabilities, marked to current dollars.

$$K_t = \text{Assets}_t - \text{Liabilities}_t \tag{1}$$

**Definition 2** (Per-period reward). The **per-period reward** at cycle  $\tau$  is the realized log-return on shareholders’ equity from  $\tau$  to  $\tau + 1$ . Log-returns track the Kelly time-average growth rate of a single surviving firm [42, 74]. The strict positivity  $K_\tau > 0$  is the firm’s survival constraint: a

trajectory in which equity hits zero is bankruptcy.

$$R_\tau = \log\left(\frac{\widehat{K}_{\tau+1}}{K_\tau}\right) \quad (2)$$

**Example.** Suppose the firm has \$100 of equity at time  $\tau$  and \$110 at time  $\tau + 1$ . Then  $R_\tau = \log(110/100) \approx 0.095$ —the firm earned \$10 on \$100 of equity over the cycle.

**Definition 3** (Cumulative objective). The **cumulative objective** at decision time  $t$  is the expected sum of per-period rewards over a finite planning horizon  $T$ , taken under the firm’s policy  $\mathcal{G}$  and its learned world model  $\widehat{\mathcal{W}}_t$ , conditioned on the information available at  $t$ . It is a Discounted Cash Flow calculation over the foreseeable time horizon; how  $T$  is calibrated is discussed in the next section.

$$J_t = \mathbb{E}^{\mathcal{G}, \widehat{\mathcal{W}}_t} \left[ \sum_{\tau=t}^{T+t-1} R_\tau \mid \mathcal{F}_t \right] \quad (3)$$

**Example.** Suppose at decision time  $t$  the firm has \$100 in equity value of its assets and its world model forecasts those to rise to \$110 over a one-cycle horizon ( $T = 1$ ). Then  $J_1 = \mathbb{E}[\log(K_{t+1}/K_t)] = \log(110/100) \approx 0.095$ , or—in the dollar-equivalent form used throughout the rest of the paper—an expected \$10 gain on \$100 of equity.

## 2.2 The Corporation as a Bundle of Assets

To execute its optimization the corporation maintains an inventory of operational components—the channels through which it spends capital, generates the cash flows that keep it solvent, and earns capital back. Five channels suffice: what the firm *holds* ( $\mathbf{I}$ , portfolio), what it can *see* ( $\mathbf{S}$ , sensors), what it can *do* ( $\mathbf{U}$ , actuators), how it *learns* ( $\mathbf{Z}$ , R&D), and what it *knows* ( $\mathbf{\Theta}$ , parameters). We adopt a lossy five-channel partition that suffices for the quant-trading instance and that we will argue extends to the general case. The framework only requires that each row can be priced in dollars and that channel-specific scaling laws are estimable from the firm’s history.

The action vector  $\mathbf{a}_t$  partitions the cycle- $t$  change in the corporation into the same five channels as the state; each entry  $a_t^k$  is the dollar change in channel  $k$  during the cycle (Equation 5). The symbol-by-symbol definitions of the five channels are given via the projections in Equation 4 and Equation 5.

**Definition 4** (Corporation tuple). The corporation is an abstract object. The state projection  $\pi^{\text{state}}$  partitions that object into the five channels that reflect the *stock of capabilities* that a corporation has at each decision time  $t$ :

$$\Xi_t \xrightarrow{\pi^{\text{state}}} (\mathbf{I}_t, \mathbf{S}_t, \mathbf{U}_t, \mathbf{Z}_t, \mathbf{\Theta}_t). \quad (4)$$

| Symbol         | Channel    | What it is   |
|----------------|------------|--|
| $\mathbf{I}_t$ | Portfolio  | The vector of current resources that the corporation can sell to produce cash. In a trading firm this is tradable asset positions across equities, fixed income, commodities, currencies, derivatives, and other instruments.                                  |
| $\mathbf{S}_t$ | Sensors    | Data and feeds that bring information about both the firm and the outside world to light. For a trading firm this includes market data feeds, internal execution telemetry, satellite feeds, and social media data.  |
| $\mathbf{U}_t$ | Actuators  | Instruments the firm uses to modify the world. For a trading firm this includes the different assets it can trade, APIs it can call, types of financing it can access, AUM, margin, and equity.  |
| $\mathbf{Z}_t$ | R&D        | How the firm processes new information and improves the other channels. For a quantitative trading firm, this determines the values for its parameters based on training data: the research process, experiment harnesses, and model selection infrastructure. |
| $\Theta_t$     | Parameters | Accumulated learned structure that represents the firm’s current beliefs about available transformations. For a quantitative trading firm, these are the parameters of the firm’s forecasting, execution, control, and value models.                           |

**Table 1:** Corporation tuple  $\Xi_t$  components.

**Definition 5** (Action vector). The **action vector**  $\mathbf{a}_t$  partitions the cycle- $t$  change in the corporation into the same five channels as the state. Each entry  $a_t^k$  is the dollar change in channel  $k$  during the cycle:

$$\mathbf{a}_t \xrightarrow{\pi^{\text{act}}} (a_t^{\mathbf{I}}, a_t^{\mathbf{S}}, a_t^{\mathbf{U}}, a_t^{\mathbf{Z}}, a_t^{\Theta}). \quad (5)$$

| Symbol             | Channel     | What it is   |
|--------------------|-------------|--|
| $a_t^{\mathbf{I}}$ | Investments | Net dollars rebalanced into (or out of) the trading book this cycle. The cleared trade, marked to current cash.                      |
| $a_t^{\mathbf{S}}$ | Sensors     | Cash spent acquiring data this cycle: new feeds, deeper archives, finer-grained subscriptions.                                       |
| $a_t^{\mathbf{U}}$ | Actuators   | Cash spent extending the firm’s execution surface: new venues, routes, APIs, financing capacity.                                     |
| $a_t^{\mathbf{Z}}$ | R&D         | Cash spent on researcher labor and capital-substitutable research: agents, GPU-hours, search infrastructure, sealed-holdout tooling. |
| $a_t^{\Theta}$     | Parameters  | Cash spent on training compute that produces the next set of model weights.  |

**Table 2:** Capital-allocation vector  $\mathbf{a}_t$  components.

### 2.3 Coupled Dynamics

The environment  $\mathbf{E}_t$  is the part of the world outside the corporation that matters for the next cycle: prices, order flow, liquidity, counterparties, macro state, regulation, and the news process. Coupled dynamics states how the corporation and that environment move together after the firm acts.

**Definition 6** (True corporate transition). The **true joint transition law**  $\mathcal{W}$  governs how the firm’s state  $\Xi_t$  and the environment  $\mathbf{E}_t$  co-evolve under an action  $\mathbf{a}_t$ . It is a property of the world, not a model the firm has access to; the firm must approximate  $\mathcal{W}$  with a learned  $\widehat{\mathcal{W}}_t$  (Equation 8).

$$(\Xi_{t+1}, \mathbf{E}_{t+1}) = \mathcal{W}(\cdot \mid \Xi_t, \mathbf{E}_t, \mathbf{a}_t) \quad (6)$$

**Example.** The world turns. The clock ticks, prices update. The corporation’s sensors measure these changes.

**Definition 7** (Small-firm approximation). The **small-firm approximation** is the regime in which the environment’s next-cycle state  $\mathbf{E}_{t+1}$  depends only weakly on the firm’s action  $\mathbf{a}_t$ :

$$\frac{\partial \mathbf{E}_{t+1}}{\partial \mathbf{a}_t} \approx 0. \quad (7)$$

**Example.** Suppose George Soros, in September 1992, shorts £10B against the Bank of England’s peg. The trade itself drains the Bank’s reserves; the next trade drains them faster; the pound’s market price collapses, and Soros makes a fortune [69]. When you are large enough to affect the market, the small-firm approximation no longer applies.

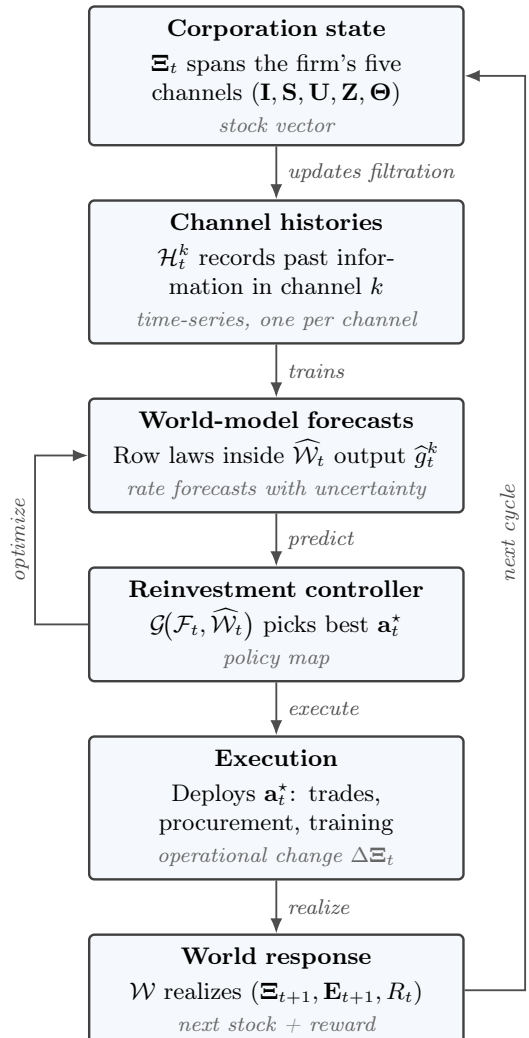
## 2.4 The Self-Forecasting Loop

Every operating cycle appends one row of observations to each channel history  $\mathcal{H}_t^k$ . The sensor channel can additionally widen by adding new data: Wayback Machine snapshots, exchange filings, alternative-data archives. Each turn provides the next iteration of training data.

Once enough rows accumulate, accurate predictions can be made on a channel-by-channel basis, allowing the firm to predict its own self-improvement dynamics. In theory, a unified world model would be trainable with enough data such that the firm could jointly predict its own state, improvements to it, and the state of the external economy all at the same time.

Concretely: at decision time  $t$  the controller queries the world model for  $\widehat{\mathbf{g}}_t$  off the trained row laws (*predict*), solves the convex inner program of Appendix D to pick the best action (*optimize*), executes the chosen  $\mathbf{a}_t^*$ , and the realized  $(\Xi_{t+1}, \mathbf{E}_{t+1}, R_t)$  becomes the new training row that tightens the next cycle’s forecasts.

Each improvement attempt teaches the firm something about itself. The action it took, the internal change it produced, and the reward



that followed become a new row in the channel histories. As those rows accumulate, the firm gets sharper predictions about which future improvements will work and tighter confidence intervals around the value of its own next actions. The loop is the thesis of the paper compressed into one diagram; everything that follows is one of these arrows worked out in detail.

### 3 The Economic World Model

The previous section gave the firm a state, a coupled dynamics  $\mathcal{W}$ , and a constrained optimization problem over a planning horizon. What it did not give the firm is a way to *evaluate* candidate allocations: given that the controller takes action  $\mathbf{a}_t$  now, what does the next  $(\mathbf{\Xi}_{t+1}, \mathbf{E}_{t+1}, R_t)$  look like? The true law  $\mathcal{W}$  answers that question in principle, but  $\mathcal{W}$  is a property of the world, not something the firm has access to. The firm must build its own learned approximation. We call that approximation the Economic World Model (EWM)—a world model in the control and planning sense, a learned object used to roll forward possible futures under candidate actions [10, 31, 33, 45, 57, 62, 66]. The qualifier *economic* marks a structural feature, not a stylistic one. The EWM is trained on a prediction-error loss (Equation 38), but every input that lowers that loss—data, compute, parameters, refits, R&D campaigns—is itself priced in dollars on the firm’s books, and every reduction in predictive loss flows back into the firm’s expected return rate. The subsections below define the EWM, contrast it with a language model trained on a static snapshot, and lay down the channel histories that make per-channel scoring tractable.

#### 3.1 EWM Definition

The EWM is the conditional next-cycle law the firm uses to forecast the joint  $(\mathbf{\Xi}_{t+1}, \mathbf{E}_{t+1}, R_t)$  given its current information set  $\mathcal{F}_t$  and a candidate action  $\mathbf{a}_t$ . It is the only model of the future the firm has access to, so improving it is itself an allocation with first-order effect on  $J_t$ .

**Definition 8** (Economic World Model). The **Economic World Model** (EWM) is the firm’s learned, filtration-respecting approximation to  $\mathcal{W}$ . Given the information set  $\mathcal{F}_t$  and a candidate action  $\mathbf{a}_t$ ,  $\widehat{\mathcal{W}}_t$  returns a joint distribution over the next firm state, the next environment, and the cycle reward. Improving  $\widehat{\mathcal{W}}_t$  is itself an allocation:

$$(\mathbf{\Xi}_{t+1}, \mathbf{E}_{t+1}, R_\tau) = \widehat{\mathcal{W}}_t(\cdot \mid \mathcal{F}_t, \mathbf{a}_t) \tag{8}$$

**Example.** Suppose the corporation has a large transformer neural network that consumes supply-chain telemetry and geopolitical news; the firm uses that network to forecast how an Iran conflict will change oil prices, how those oil prices will change the value of the firm’s futures contracts, and what effect that change will have on shareholders’ equity.

### 3.2 Firm and Channel Histories

Because the firm sees the world only through its sensors, both  $\mathbf{E}_t$  and its own state  $\mathbf{\Xi}_t$  enter the EWM as posteriors over noisy observations rather than as latent ground truth. The firm history  $\mathcal{H}_t$  is the chronological log of every observation through cycle  $t$ —the canonical sufficient statistic in the partially-observed Markov decision process tradition [7, 11, 63, 76]. The channel histories  $\mathcal{H}_t^k$  slice it channel by channel into  $(\mathbf{o}_\tau^k, a_\tau^k, R_{\tau+1})$  rows. These per-channel slices are the tables the firm uses to fit the row laws of the marginal-return vector.

**Definition 9** (Firm history). The **firm history** at time  $t$  is the chronological observation–action–reward log the firm has recorded through cycle  $t$ . The observation vector  $\mathbf{o}_\tau$  is the sensor output generated by the latent corporation–environment state; the action  $\mathbf{a}_\tau$  is what the firm did next; and  $R_\tau = \log(K_{\tau+1}/K_\tau)$  is the cycle- $\tau$  reward realized after that action.

$$\mathcal{H}_t = \begin{bmatrix} \mathbf{o}_0 & \mathbf{a}_0 & R_1 \\ \mathbf{o}_1 & \mathbf{a}_1 & R_2 \\ \vdots & \vdots & \vdots \\ \mathbf{o}_{t-1} & \mathbf{a}_{t-1} & R_t \\ \mathbf{o}_t & - & - \end{bmatrix} \quad (9)$$

**Definition 10** (Channel history). The **channel history** for channel  $k$  at time  $t$  is the channel- $k$  slice of  $\mathcal{H}_t$ , recording for each cycle in the firm’s lookback window the channel observation  $\mathbf{o}_\tau^k$ , the channel action  $a_\tau^k$ , and the next realized log-equity reward.

$$\mathcal{H}_t^k = \begin{bmatrix} \mathbf{o}_{t-W}^k & a_{t-W}^k & R_{t-W+1} \\ \mathbf{o}_{t-W+1}^k & a_{t-W+1}^k & R_{t-W+2} \\ \vdots & \vdots & \vdots \\ \mathbf{o}_{t-1}^k & a_{t-1}^k & R_t \end{bmatrix} \quad (10)$$

**Example.** Suppose the firm keeps a sensor-channel history of every dataset it has owned. Each row records the data inventory available before the cycle, the dataset or feed acquired in that cycle, how that data set improved backtests on different candidate models, and how much live trading was improved, as well as counterfactual simulations of what would have happened had the firm not acquired that data set.

**Definition 11** (Firm filtration). The **firm filtration** at time  $t$  is the  $\sigma$ -algebra generated by the firm history. A random variable is  $\mathcal{F}_t$ -measurable iff its value is determined by  $\mathcal{H}_t$ . The family  $\{\mathcal{F}_s\}_{s \geq 0}$  is what enforces the no-peeking discipline of the EWM: a forecast at time  $t$  may condition on  $\mathcal{F}_t$  and on nothing resolved later.

$$\mathcal{F}_t = \sigma(\mathcal{H}_t) \quad (11)$$

### 3.3 LLMs are not economic world models

The filtration requirement is what separates an Economic World Model from an ordinary language model. At decision time  $t$ , an EWM may condition only on  $\mathcal{F}_t$  and the candidate action  $\mathbf{a}_t$ . A language model trained on a static corpus [40] can mix documents from before and after the event it is asked to predict, so its context can contain future information. The same no-peeking discipline

that separates a forecasting model from a memorizing one has been the organizing principle of game-playing AI [31, 33, 66, 71], and the loss functions below make the difference explicit. For next-token modeling on a fixed snapshot, this is fine. As the primary EWM  $\widehat{\mathcal{W}}_t$  without an externally imposed filtration discipline, it is not: held-out validation can only enforce filtration discipline if the holdout set is chronologically after the entire training corpus—a window that shrinks as internet-scale models train on ever more recent data. This results in very little data left for validation, and even less for robust benchmarking. Consequently, measured performance uncertainty remains high, and the small post-corpus window provides negligible information about how the model adapts to changing market regimes. The outcome is slower compounding of capital and structurally riskier decisions. A general LLM that is wrapped in a strict post-cutoff evaluation harness can still serve as a component or proposal mechanism; the categorical claim is about the bare model, not the wrapped system.

$$\mathcal{L}_{\text{LLM}}(\Theta) = \sum_i \ell\left(\widehat{p}_\Theta(x_i \mid \overbrace{\text{ctx}_i}^{\text{can include post-}i \text{ tokens}}), x_i\right) \quad (\text{permutation-invariant over documents}). \quad (12)$$

$$\widehat{\mathcal{L}}_{\text{EWM}}(\Theta) = \sum_{\tau \in \mathcal{I}_{\text{eval}}} \ell\left(\widehat{\mathcal{P}}_\tau(\mathbf{o}_{\tau+1}, R_{\tau+1} \mid \overbrace{\mathcal{F}_\tau}^{\text{causal filtration}}, \mathbf{a}_\tau), (\mathbf{o}_{\tau+1}, R_{\tau+1})\right) \quad (\text{information order matters}). \quad (13)$$

**Example.** Suppose a language model is trained on a 2024 internet snapshot that contains articles, analyst notes, and Wikipedia edits explaining a market event from 2022. When that model learns the 2022 event, it can absorb the 2024 retrospective explanation of what happened. A filtration-respecting prediction model cannot do that: in a 2022 backtest, the training data must contain only information available up to the 2022 timeline.

### 3.4 Channel-Specific World Models

In principle the firm has one joint EWM  $\widehat{\mathcal{W}}_t$  over the entire corporation–environment state. In practice that joint object is approximated by a collection of channel-specific world models, each trained on its own channel history  $\mathcal{H}_t^k$ . These per-channel transition models are usually much simpler than a full simulator—a scaling law, a market-impact curve, a refit-decay model, or a search law is already a channel-world-model fragment. The split is a practical approximation, not a claim that the channels evolve independently; cross-channel coupling re-enters when the controller composes the rows of  $\mathbf{g}_t$  below.

**Definition 12** (Channel-specific world model). In principle the firm has one joint EWM  $\widehat{\mathcal{W}}_t$  over the whole corporation–environment state. In practice, a **channel-specific world model** is the channel- $k$  conditional law over the next channel observation and reward, given the channel history  $\mathcal{H}_t^k$  and a candidate channel action  $a_t^k$ :

$$(\widehat{\mathbf{o}}_{t+1}^k, \widehat{R}_{t+1}) \sim \widehat{\mathcal{W}}_t^k\left(\cdot \mid \mathcal{H}_t^k, a_t^k\right), \quad k \in \{\mathbf{I}, \mathbf{S}, \mathbf{U}, \mathbf{Z}, \Theta\}. \quad (14)$$

**Example.** Suppose the corporation keeps an Investments-channel world model,  $\widehat{\mathcal{W}}_t^{\mathbf{I}}$ , trained specifically to forecast the value of assets available for purchase and how much trading them will impact their price. An Investments-specific world model asks: if the rest of the firm is held constant and the firm maintains a candidate portfolio  $\mathbf{I}$ , how much expected equity growth will it produce?

## 4 The Portfolio Optimizer

Choosing how to spread a fixed pool of capital across competing assets so as to maximize a long-horizon utility of wealth is the canonical problem of *portfolio optimization*, with a literature that runs from Markowitz mean–variance allocation [49] and Kelly’s growth criterion [42], through Sharpe’s reward-to-volatility ratio [68] and Merton’s continuous-time program [51], to universal portfolios [21], robust mean–variance allocation [29], and the broader machinery of convex optimization and constrained stochastic optimal control [10, 14, 62]. The problem the corporation solves at each cycle is a multi-period instance of the same family: the channels of  $\Xi_t$  are the assets,  $\mathbf{a}_t$  is the rebalancing trade, and  $J_t$  is the long-horizon log-utility. The novelty is the asset set: the firm allocates not only across tradable instruments but across sensors, actuators, parameters, and R&D on the same dollar axis.

We adopt the standard framing of model-predictive control: the EWM  $\widehat{\mathcal{W}}_t$  supplies forecasts under candidate actions, and the policy  $\mathcal{G}$  chooses the next allocation by solving a single-cycle convex inner problem over the per-channel return posteriors, then committing the chosen action and re-solving in the next cycle [10, 14, 15]. This section states the corporate optimization problem in its convex form and introduces the marginal-return vector that the rest of the blueprint estimates row by row.

### 4.1 The Corporate Optimization Problem

Stitching together the cumulative objective and the firm’s solvency, liquidity, and budget constraints gives the corporate program: the controller  $\mathcal{G}$  picks the action that maximizes  $J_t$  subject to the constraints in the appendix. The full constraint set is collected in [Appendix B](#) so the body can stay focused on the convex inner program.

**Definition 13** (Corporate optimization problem). The **corporate optimization problem** chooses the policy  $\mathcal{G}$  that maximizes the cumulative objective  $J_t$ , subject to budget, channel-liquidation, liquidity, and solvency constraints.

$$\mathcal{G}^* = \arg \max_{\mathcal{G}} J_t \quad \text{s.t. financial constraints.} \quad (15)$$

**Example.** Suppose at cycle  $t$  AlphaFund has  $K_t^{\text{deploy}} = \$2\text{M}$  and faces a candidate policy that proposes \$900K investments, \$400K sensors, \$300K parameters, \$250K R&D. The cumulative objective  $J_t$  is maximized at this allocation *provided* all four constraints clear: total spend  $\$1.85\text{M} \leq \$2\text{M}$  (budget binds slack), no channel falls below its liquidation floor, and the liquidity and solvency reserves of [Appendix B](#) are intact. If the proposal had instead been \$2.2M, the budget constraint would bind first and the optimizer would reject the policy before any marginal-return comparison.

## 4.2 The Marginal Return Vector

Differentiating the common objective  $J_t$  with respect to the capital-allocation vector turns heterogeneous interventions—a researcher hire, a new alternative-data feed, a GPU, a position in AAPL, an execution-latency upgrade—into directly comparable marginal rates. Each entry of  $\mathbf{g}_t$  is the expected log-equity growth per marginal dollar invested in that channel; at the optimum every funded channel equates the same risk-adjusted shadow price of capital,  $\widehat{g}_t^k / \sigma_t^k = \lambda_{S,t}^*$  (Appendix D), which collapses to the bare equimarginal identity  $\widehat{g}_t^k = \lambda_t^*$  only in the risk-neutral limit ( $\kappa_t \rightarrow 0$ , or per-channel dispersions  $\sigma_t^k \rightarrow 0$ ). The risk functional itself is general: Equation 20 writes the full corporate t-RSI in terms of any auditable uncertainty functional  $\mathcal{U}_t$ , of which standard deviation is just the operational default. The next section instantiates this row by row for the quant firm.

**Definition 14** (Marginal-return vector). The **marginal-return vector**  $\mathbf{g}_t$  is the gradient of the cumulative objective  $J_t$  with respect to the cycle- $t$  capital-allocation vector  $\mathbf{a}_t$ . Each coordinate of  $\mathbf{g}_t$  tells the optimizer how much expected log-equity growth one extra dollar buys in that channel right now.

$$\mathbf{g}_t = \frac{\partial}{\partial \mathbf{a}_t} J_t \quad (16)$$

**Example.** Suppose the firm’s current estimates are  $g_t^S = 0.2$  per dollar (sensors) and  $g_t^I = 0.1$  per dollar (investments). Then the next \$100 should go to sensors: \$20 of expected future log-equity growth versus \$10 from putting that \$100 in the trading book.

$$\begin{bmatrix} g_t^I \\ g_t^S \\ g_t^U \\ g_t^Z \\ g_t^\Theta \end{bmatrix} = \begin{bmatrix} \pi_I \mathbf{g}_t \\ \pi_S \mathbf{g}_t \\ \pi_U \mathbf{g}_t \\ \pi_Z \mathbf{g}_t \\ \pi_\Theta \mathbf{g}_t \end{bmatrix} \quad (17)$$

**Definition 15** (Per-channel marginal return). For each channel  $k$ , we use the chain rule to expand the **per-channel marginal return**  $g_t^k$  along the trajectory through which a cycle- $t$  dollar propagates. The equity sensitivity  $\partial R_\tau / \partial \Xi_\tau$  says how much next-cycle reward changes when the corporation state moves; the propagation Jacobian  $\partial \Xi_\tau / \partial a_t^k$  says how a dollar spent on channel  $k$  at time  $t$  moves the state at time  $\tau$ . Summing their product over the planning horizon gives the dollar’s full contribution to  $J_t$ :

$$g_t^k = \sum_{\tau=t}^{T+t-1} \frac{\partial}{\partial \Xi_\tau} R_\tau \frac{\partial}{\partial a_t^k} \Xi_\tau \quad (18)$$

**Example.** Suppose AlphaFund onboards an IEX D-Limit route this cycle. That route enters  $\mathbf{U}_{t+1}, \mathbf{U}_{t+2}, \dots$  and shaves friction off every rebalance from then on. Each future  $\tau$  contributes one term to  $g_t^U$  through  $\partial R_\tau / \partial \Xi_\tau$  and  $\partial \Xi_\tau / \partial a_t^U$ , so a route that pays back across thirty rebalances is priced as the sum of all thirty marginal contributions, not just the next one.

$$\begin{bmatrix} g_t^{\mathbf{I}} \\ g_t^{\mathbf{S}} \\ g_t^{\mathbf{U}} \\ g_t^{\mathbf{Z}} \\ g_t^{\Theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial \Delta \mathbf{I}_t} \left[ \frac{\Delta \mathbf{I}_t \hat{\mu}_t}{a_t^{\mathbf{I}}} - \hat{\phi}_t(\Delta \mathbf{I}_t, \mathbf{U}_t, \mathbf{E}_t) \right] \\ \sum_{\tau=t}^{T+t-1} \frac{\partial}{\partial \Xi_\tau} R_\tau \frac{\partial}{\partial a_t^{\mathbf{S}}} \Xi_\tau \\ \sum_{\tau=t}^{T+t-1} \frac{\partial}{\partial \Xi_\tau} R_\tau \frac{\partial}{\partial a_t^{\mathbf{U}}} \Xi_\tau \\ \sum_{\tau=t}^{T+t-1} \frac{\partial}{\partial \Xi_\tau} R_\tau \frac{\partial}{\partial a_t^{\mathbf{Z}}} \Xi_\tau \\ \sum_{\tau=t}^{T+t-1} \frac{\partial}{\partial \Xi_\tau} R_\tau \frac{\partial}{\partial a_t^{\Theta}} \Xi_\tau \end{bmatrix} \quad (19)$$

### 4.3 Capital Allocation

The horizon  $T$  introduces uncertainty into this pricing: the variance of cumulative reward  $\sum_{\tau} R_{\tau}$  propagates through the rollout.  $T$  is therefore set at the point where additional cycles no longer meaningfully tighten the estimate once that propagating uncertainty is accounted for. This uncertainty is fundamental to the controller’s allocation problem: an honest pricing of capital must trade off the central tendency of forecasted returns against their dispersion. The most general statement of this is *expected-utility portfolio optimization*, which evaluates allocations under whatever utility functional the firm chooses over the joint distribution of channel returns. Under the simplest case of normally distributed returns this collapses to *mean-variance portfolio optimization* [49] and the per-channel Sharpe ratio  $\hat{g}_t^k / \sigma_t^k$  [68]—the form we use throughout the rest of the paper for simplicity; in production, more complex utility-of-distribution objectives are often preferred.

Appendix D details why this uncertainty necessitates the risk-adjusted form not only for the investment channel, but for the performance and optimal allocation of every channel.

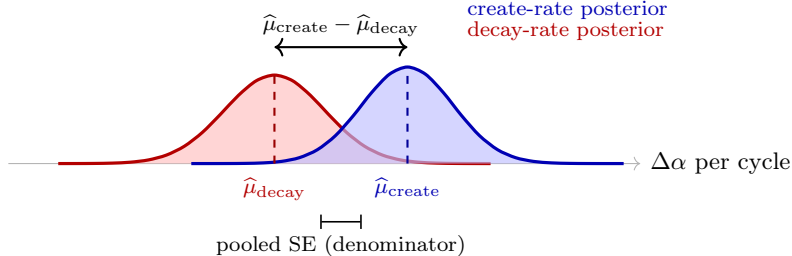
## 5 Experimental Evidence of t-RSI

Thus far we have talked about a theoretical framework for probabilistic recursive self-improvement. Now we will present experimental evidence that this framework is operational. Equation 20 formally defines t-RSI as the signal-to-noise ratio of net improvement; the remaining subsections then estimate the marginal-return vector  $\mathbf{g}_t$  for the quant firm, taking each row of  $\Xi_t$ —investments, sensors, actuators, parameters, and R&D—in turn. Each subsection opens with the channel’s history  $\mathcal{H}_t^k$  and channel-specific world model  $\widehat{\mathcal{W}}_t^k$ , derives the corresponding row of  $\mathbf{g}_t$ , and reports the fitted parameters that the controller of the previous section consumes. Investments is the only instantaneous row whose return is read directly off the broker statement; the other four are state-flow channels whose payouts arrive over future cycles and are therefore priced through the trajectory chain rule. The empirical content—data scaling, loss-to-edge linearization, the Chinchilla-style joint surface, the 929-experiment auto-research log law, and the continual-learning intersection—is what makes the marginal-ROI comparison of the previous section operational rather than rhetorical.

### 5.1 What t-RSI Measures

**Why t-RSI is the thing to measure.** The differentiable-corporation thesis says the firm carries two posteriors at every cycle: one over alpha *created* per dollar spent on each channel, and one over alpha *decayed* from the deployed book. A controller that compounds is one that commits capital only when those two distributions are confidently separated — it believes it will create more alpha next cycle than the alpha already on the books erodes, and it believes it by enough margin that

the call is unlikely to be noise. We call that standardized separation the *improvement signal-to-noise ratio*, written t-RSI by analogy with a two-sample *t*-statistic. The two distributions are not draws from one underlying process: they are posteriors over two different processes (create from the channel-row fits below, decay from the firm’s forecast-evaluation and live-trading panel). t-RSI is therefore framed as a standardized distance, not a hypothesis-test instrument.



**Figure 2:** The improvement signal-to-noise ratio. The two posterior densities are the firm’s belief over per-cycle alpha creation (blue, from the channel-row fits in the subsections that follow) and per-cycle alpha decay (red, from the forecast-evaluation panel and Mark II/III live-trading history). t-RSI is the gap between their posterior means measured in units of pooled standard error: separation matters relative to dispersion, not in absolute terms.

A positive t-RSI of, e.g., 2.5 means the create-rate posterior is centered 2.5 pooled standard errors above the decay-rate posterior. The firm’s controller does not commit capital to a candidate until that statistic clears a sealed-evaluator threshold (Equation 71); that gate is what makes compounding survive selection, and what distinguishes a self-improving corporation from one that promotes drift on noise. Equivalently, every row of the channel table below contributes one slice of the blue density, and the cycle’s investment decision is a draw from *which* slice the controller will pay for next.

**Headline definition.** Write  $\widehat{\Delta\alpha}^{\text{create}}_{t:H}$  for the firm’s posterior alpha creation over horizon  $H$  (the sum of per-channel contributions, with bookkeeping fixed in subsection E.1), and  $\widehat{\Delta\alpha}^{\text{decay}}_{t:H}$  for the matching alpha-decay posterior. t-RSI is the standardized distance between them:

$$\text{t-RSI}_{t,H} = \frac{\widehat{\Delta\alpha}^{\text{create}}_{t:H} - \widehat{\Delta\alpha}^{\text{decay}}_{t:H}}{\sqrt{\text{SE}^2(\widehat{\Delta\alpha}^{\text{create}}_{t:H}) + \text{SE}^2(\widehat{\Delta\alpha}^{\text{decay}}_{t:H})}}. \quad (20)$$

The channel-by-channel decomposition of  $\widehat{\Delta\alpha}^{\text{create}}_{t,H}$  as a path integral along the planned allocation path, and the bootstrap propagation that supplies each SE, live in subsection E.1; the operational walk-through of the headline three-month calculation lives in subsection E.2.

**What the data presently say.** Empirically the firm’s own panel measures the decay rate as small and near zero. The data-derived headline reads  $\widehat{\lambda}^{\text{decay}}$  from a per-asset alpha-decay estimator that fits the held-out forecast edge against deployment age once per asset and aggregates the resulting per-asset rate distribution robustly across the held-out asset universe (median + MAD/IQR summary; SE combines within-asset-cell dispersion with between-cell variability over training-run seeds and forecast horizons). The same qualitative picture holds in production: across 16 months of cumulative Mark II and Mark III live trading, every linear, rank, and distributional trend statistic

against deployment age returns  $\hat{\kappa}$  near zero. The data-derived headline t-RSI is therefore dominated by create-side dispersion, not by decay. The subsections that follow estimate the per-channel create-rate distribution one channel at a time; [subsection 5.8](#) composes them into the headline t-RSI for the current operating point.

## 5.2 Investments as an Asset

**What this channel is.** Investments are the firm’s current production channel: capital positioned in the trading book and rebalanced by the controller. The EWM slice for this channel forecasts price-change distributions. Construction, scaling laws, and held-out benchmarks are deferred to a forthcoming companion paper; the empirical claims in this paper are channel-decomposed and do not depend on the companion-paper architecture. The controller turns those forecasts into a rebalance; the market function turns the rebalance into realized cash. This is the only row whose one-cycle return is observed directly from the broker ledger rather than inferred counterfactually [4, 13, 15, 26, 30, 35, 44, 46, 59].

**Definition 16** (Investment marginal return). The **investment marginal return**  $g_t^{\mathbf{I}}$  is the gradient with respect to the candidate trade  $\Delta \mathbf{I}_t$  of expected return-per-dollar minus the learned execution friction  $\hat{\phi}_t$ . Investments are the only *instantaneous* channel: the gradient sees a single cycle because positions are marked to cash at the end of it.

$$g_t^{\mathbf{I}} = \frac{\partial}{\partial \Delta \mathbf{I}_t} \left[ \begin{array}{c} \text{marginal return per} \\ \text{portfolio rebalance} \\ \frac{\Delta \mathbf{I}_t \hat{\mu}_t}{a_t^{\mathbf{I}}} \\ \text{execution friction} \\ - \hat{\phi}_t(\Delta \mathbf{I}_t, \mathbf{U}_t, \mathbf{E}_t) \end{array} \right]. \quad (21)$$

**Example.** Suppose the EWM forecasts  $\hat{\mu}_t = 8$  bps for a \$1M long in AAPL over the next bar against  $\hat{\phi}_t = 3$  bps of expected round-trip friction. Then  $g_t^{\mathbf{I}} \approx 5$  bps per dollar at this trade size. Doubling the size pushes  $\hat{\phi}_t$  up (slippage convex in size); the size optimum is where  $g_t^{\mathbf{I}}$  falls to the marginal ROI  $\lambda_t^*$ .

**What this means.** The investments row tells the controller what one extra dollar of trading capital is worth right now: forecast edge  $\hat{\mu}_t$  on the candidate trade, minus the learned execution friction  $\hat{\phi}_t$  paid at execution. Of the three terms in (21), the EWM forecast and the size dependence of  $\hat{\phi}_t$  are estimable from a backtest; the remainder of  $\hat{\phi}_t$  is not, and is the proprietary surface the rest of this subsection explains. Two of the three terms are public; one is not, and that asymmetry is why the live track record matters more than the backtest figure for this row.

**Backtest.**  $\phi$  deserves special emphasis: it mostly cannot be estimated from a backtest. Market impact admits a square-root-law approximation ([subsection F.1](#)), but the remainder—routing quality, venue- and broker-specific spreads, financing, adversarial response—is venue-, instrument-, size-, and time-of-day-specific and can only be learned by trading. The firm has executed approximately \$400M of trades; that volume is the data behind its internal  $\phi$  surface, which is proprietary for exactly the reason (21) makes plain— $\phi$  is the term the rest of the world cannot buy. The dated

three-deployment-generation track record (Mark I/II/III) is reported in [Appendix H](#).<sup>1</sup>

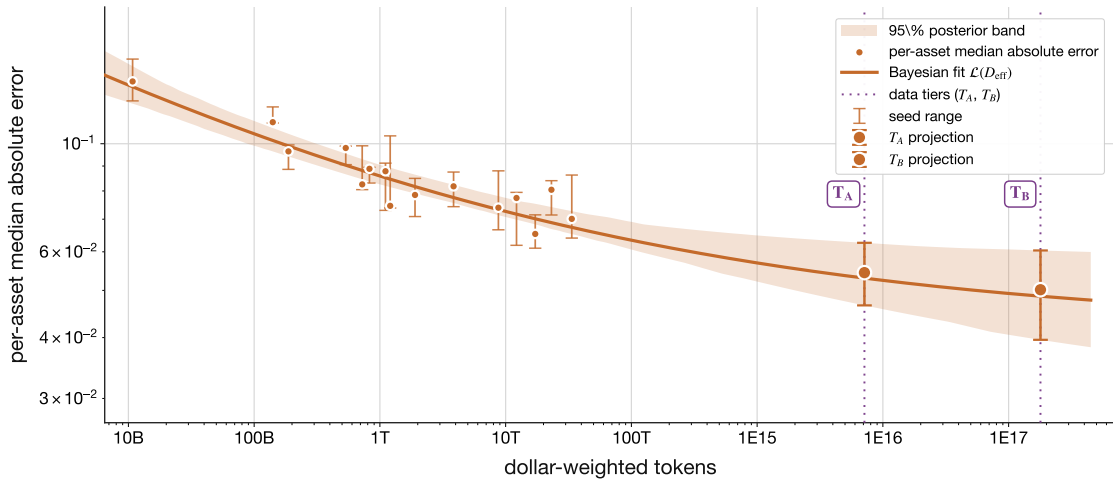
### 5.3 Sensors as an Asset

**What this channel is.** Sensors are purchases that enlarge the filtration  $\mathcal{F}_t$ : market-data feeds, historical records, finer sampling, and any measurement that lets the EWM condition on more of the world. The empirical primitive is the data-scaling law, fit on the Muennighoff effective-data axis and the Hoffmann/Kaplan loss form [37, 40, 56, 70], with the underlying unit being the dollar-weighted bar in the spirit of information-driven sampling [5, 19, 48]. The body keeps only the local loss slope with respect to effective data; the fit protocol, prior choices, and controller-level chain rule are collected in [subsection F.2](#).

**Definition 17** (Local data-scaling slope). The **local data-scaling slope** is the measured sensor primitive: how predictive loss changes when the effective dollar-weighted-token axis grows by one log-base-10 decade. The slope is the data-scaling decay rate times the reducible loss still available to be removed.

$$\underbrace{\frac{\partial \mathcal{L}_{\text{pred}}}{\partial \log_{10} D_{\text{eff}}}}_{\text{local loss-data slope}} = \underbrace{-\ln(10) \alpha_{D_{\text{eff}}}}_{\text{data-scaling decay rate}} \underbrace{(\mathcal{L}_{\text{pred}} - \mathcal{L}_{\text{noise}})}_{\text{reducible loss}} \quad (22)$$

**Example.** Suppose the fitted multi-seed sensor slope is  $\alpha_{D_{\text{eff}}} = 0.075$  and the reducible loss component at the current  $D_{\text{eff}}$  is  $A_{D_{\text{eff}}} D_{\text{eff}}^{-\alpha_{D_{\text{eff}}}} = 0.012$ . The local slope  $\partial \mathcal{L}_{\text{pred}} / \partial \log_{10} D_{\text{eff}} = -\alpha_{D_{\text{eff}}} \cdot A_{D_{\text{eff}}} D_{\text{eff}}^{-\alpha_{D_{\text{eff}}}} \approx -9.0 \times 10^{-4}$  per decade of  $D_{\text{eff}}$ : each additional decade of effective dollar-weighted tokens removes about  $9 \times 10^{-4}$  of loss at this operating point, falling as the reducible component shrinks toward  $\mathcal{L}_{\text{noise}}$ .



**Figure 3:** Dollar-weighted tokens vs loss

#### Fitted parameters.

<sup>1</sup>Supporting equations for this channel are collected in [Appendix F](#).

| parameter                    | value                 | 95% CI          | $R^2$  | $n$ |
|------------------------------|-----------------------|-----------------|--------|-----|
| $\alpha_{D_{\text{eff}}}$    | 0.156 (dimensionless) | [0.09, 0.229]   | 0.7619 | 45  |
| $R^*$                        | 4.306 (epochs)        | [2.324, 6.231]  | 0.7619 | 45  |
| $\mathcal{L}_{\text{noise}}$ | 0.042 (loss)          | [0.024, 0.061]  | 0.7619 | 45  |
| $A_{D_{\text{eff}}}$         | 3.266 (loss scale)    | [0.8072, 16.87] | 0.7619 | 45  |

**Equation**

$$\widehat{\mathcal{L}}(D_{\text{eff}}) = 0.042 + 3.266 D_{\text{eff}}^{-0.156}$$

$$D_{\text{eff}} = U_{D_{\text{s}}} [1 + 4.306 (1 - \exp(-(E - 1)/4.306))]$$

**What this means.** The fitted exponent says how much predictive error remains after the firm buys another decade of effective data. At the reported operating point,  $\alpha_{D_{\text{eff}}} \approx 0.16$  means a  $10\times$  increase in effective dollar-weighted tokens multiplies the reducible part of the loss by roughly  $10^{-\alpha_{D_{\text{eff}}}}$ , or about 0.7: the model keeps about 70% of the reducible error and removes about 30%. The interval around  $\alpha_{D_{\text{eff}}}$  is the uncertainty on that local data slope, not on trading alpha.  $R^*$  says when repeated passes over the same data stop behaving like fresh information, and  $\mathcal{L}_{\text{noise}}$  is the residual floor that more sensors cannot remove at the current model size and architecture; R&D moves architecture and therefore moves this floor (see subsection F.2). This is why the chart is a sensor asset measurement: it prices data by the remaining predictive loss it can still reduce.

We expect this scaling law to continue, as scaling laws are some of the most well-established results across a variety of domains; as it turns out, our preliminary evidence suggests that quantitative trading demonstrates scaling laws as well. What the chart is actually pricing (effective dollar-weighted tokens, not number of predictive factors), the realistic order-of-magnitude scaling factors available against the current operating point, and the cross-domain evidence for power-law extrapolation are collected in Appendix G.

## 5.4 Actuators as an Asset

**What this channel is.** Actuators are the interfaces through which the controller can act. In the current quant-trading envelope, the measured actuator is the tradable asset universe: as the universe expands, the firm both trains on more dollar-weighted histories and gains more instruments through which forecasts can be deployed. The body keeps only the local data-performance slopes measured by that panel; the fit protocol, joint-path interpretation, and general actuator row  $g_t^U$  are collected in subsection F.3.

**Definition 18** (Local data-performance slopes). The **local data-performance slopes** are the measured actuator primitives in the current trading-envelope instance: annualized return and annualized Sharpe gained per decade of effective dollar-weighted data. In this section the actuator surface is the tradable asset universe, so the body reads these slopes directly from the panel rather than decomposing them into separate data, loss, execution, and friction terms.

$$\frac{\overbrace{\partial \mu_t^{\text{ann}}}^{\text{local return-data slope}}}{\partial \log_{10} D_{\text{eff}}} \quad , \quad \frac{\overbrace{\partial \text{Sharpe}_t}^{\text{local Sharpe-data slope}}}{\partial \log_{10} D_{\text{eff}}} \quad (23)$$

**Example.** Suppose the actuator panel’s OLS slope on annualized Sharpe is 0.32 per decade of  $D_{\text{eff}}$  (cluster-median fit across the 12 universe sizes in the multi-seed scaling sweep). A universe expansion that absorbs an additional half-decade of dollar-weighted tokens predicts  $\sim 0.16$  Sharpe units of additional realized performance at the current operating point, before reserving capacity headroom against the impact baseline of subsection F.1.

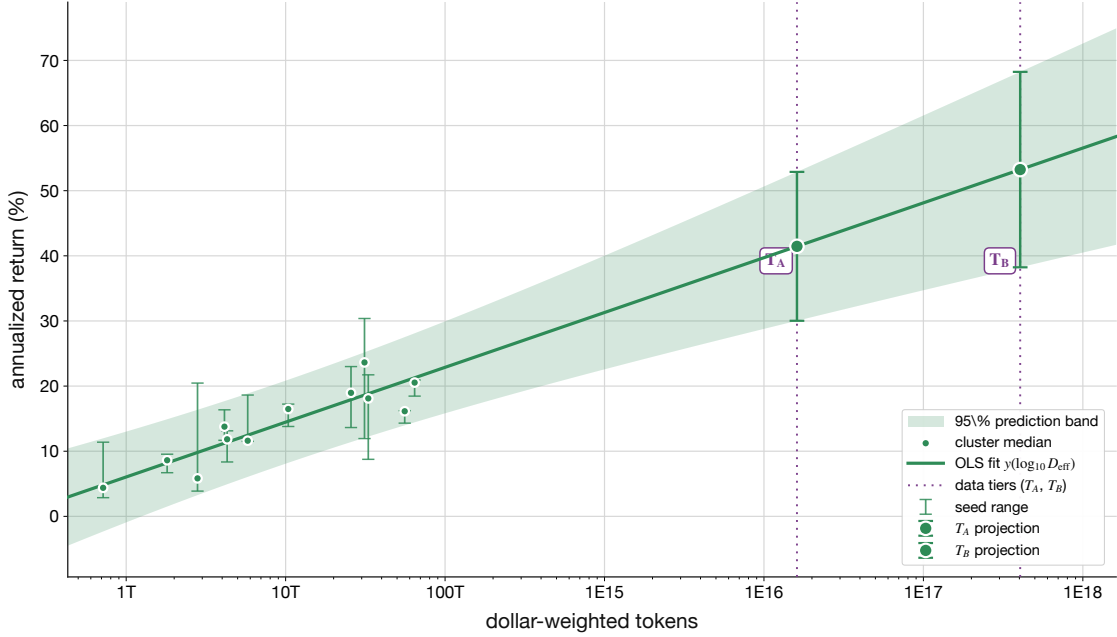


Figure 4: Annualized return vs. dollar-weighted tokens.

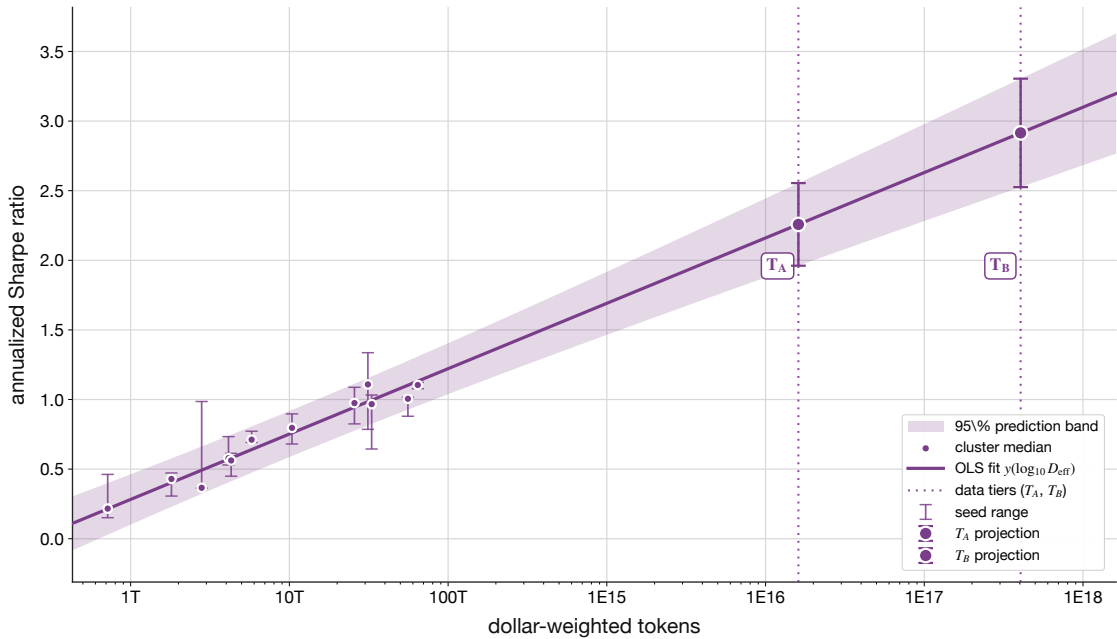


Figure 5: Annualized Sharpe ratio vs. dollar-weighted tokens.

## Fitted parameters.

| parameter  | value                         | 95% CI           | $R^2$  | $n$ |
|--|-------------------------------|------------------|--------|-----|
| $\partial \text{AnnRet} / \partial \log_{10} D_{\text{eff}}$ | 8.417 (pct points per decade) | [5.844, 10.99]   | 0.8043 | 12  |
| $\text{AnnRet}(T_A)$   | 41.46 (pct points)            | [30.03, 52.88]   | N/A    | 12  |
| $\text{AnnRet}(T_B)$   | 53.24 (pct points)            | [38.24, 68.24]   | N/A    | 12  |
| $\partial \text{Sharpe} / \partial \log_{10} D_{\text{eff}}$ | 0.4696 (Sharpe per decade)    | [0.4027, 0.5364] | 0.9499 | 12  |
| $\text{Sharpe}(T_A)$   | 2.258 (Sharpe)                | [1.961, 2.554]   | N/A    | 12  |
| $\text{Sharpe}(T_B)$   | 2.915 (Sharpe)                | [2.526, 3.305]   | N/A    | 12  |

**Equation**

$$\widehat{\text{AnnRet}}(D_{\text{eff}}) = -94.96 + 8.417 \log_{10} D_{\text{eff}}$$

$$\widehat{\text{Sharpe}}(D_{\text{eff}}) = -5.352 + 0.4696 \log_{10} D_{\text{eff}}$$

**What this means.** The actuator row asks whether expanding the tradable surface changes what the investment production function can do. The two slopes above are empirical between- $N$  regressions: realized annualized return and realized annualized Sharpe per decade of effective dollar-weighted tokens. In this trading-envelope instance, the asset universe and the dollar-weighted data universe expand together, so the fitted slopes are read directly as local data-performance slopes rather than as separately identified data-only and actuator-only effects. The fitted slopes, their 95% confidence intervals, and the  $T_A/T_B$  extrapolation rows in the table above are all read directly off the actuator panel and the same OLS bookkeeping that draws the 95% prediction band on the chart. A note on sample size: the multi-seed sweep produces 15 universe sizes  $\times$  3 seeds = 45 raw runs that anchor the sensor-row data-scaling fit, but the actuator-panel regression takes cluster medians across the three seeds at each universe size, so the  $n = 12$  reported in the parameter table is the 12-cluster-median count after dropping the three smallest universe sizes for which the per-seed median has insufficient evaluation breadth; the underlying 36-row ( $N, \text{seed}$ ) panel is the object the cluster-by- $N$  bootstrap of [subsection E.2](#) resamples. The parameter table's  $n$  column reports the regression degrees-of-freedom unit (cluster medians), not the raw run count.

**Potential scale.** We estimate the actuator-row scaling law could be extended over two tranches of source data that still satisfy the dollar-weighted-token construction of [subsection E.2](#) (dollar-denominated, strict time filtration); the corresponding  $T_A$  and  $T_B$  extrapolated Sharpe and annualized return are reported in the parameter table above. Tier  $T_B$  is included because the underlying data class satisfies the same empirically measured dollar-weighted-token line fit that anchors Tier  $T_A$ .

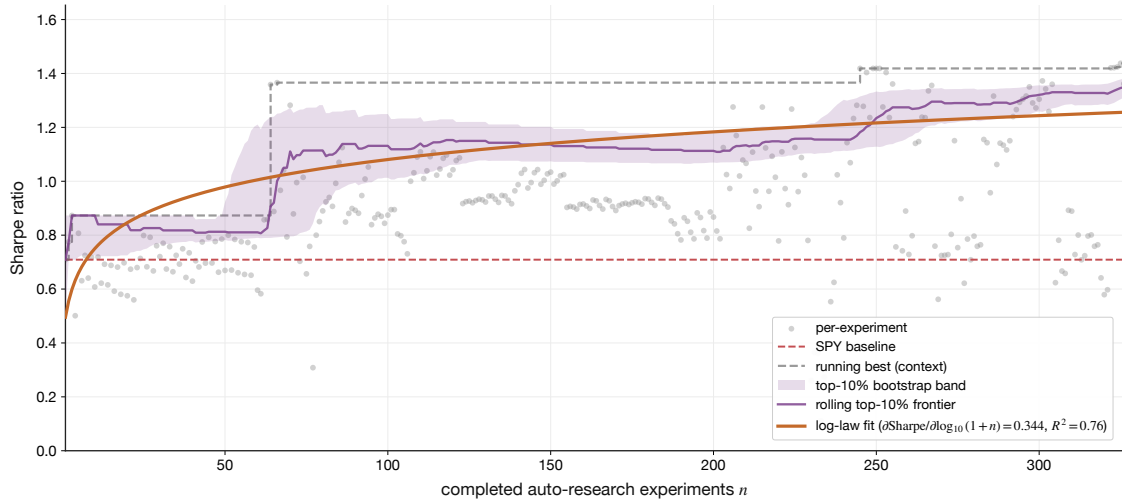
- **Tier  $T_A$  – pure asset-price data:** the catalogue of asset prices, quotes, trades, and order-book snapshots already commercially available across data brokers (global equities, futures, FX, rates, options, crypto, OTC fixings).
- **Tier  $T_B$  – broader dollar-denominated time-series data:** any other dollar-denominated series with a strict time index that respects the filtration (card panels, payments and settlement flows, insurance-claims tapes, payroll feeds, satellite- and geolocation-derived activity counts, public budget and procurement records), not restricted to tradable asset prices.

## 5.5 R&D as an Asset

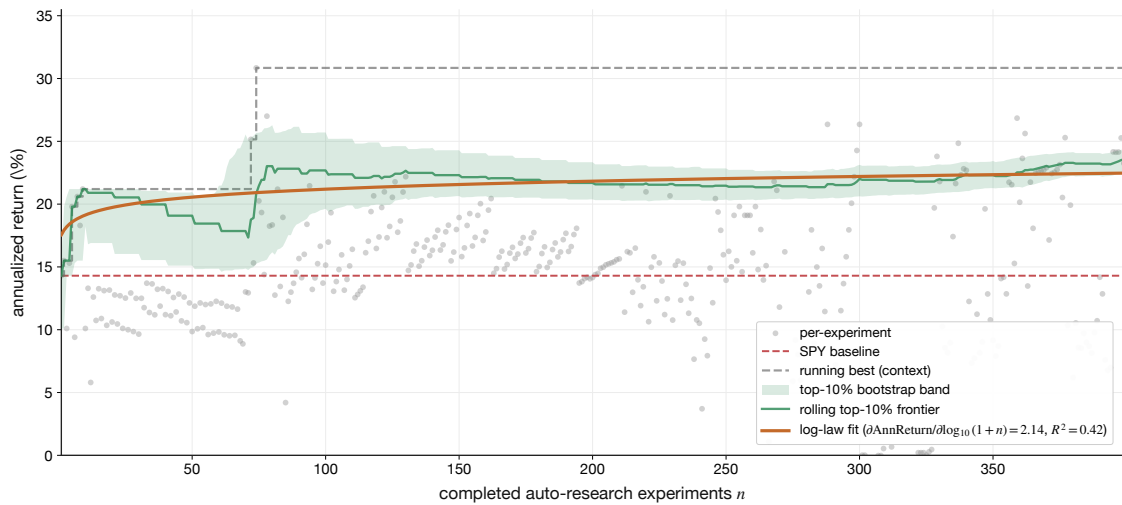
**What this channel is.** R&D is purchases that improve future search cycles: researcher labor, LLM API spend, GPU-hours, evaluation infrastructure, and search tooling. The empirical primitive the body reports is the derivative of the selected rolling upper-tail held-out frontier with respect to completed auto-research experiments, fit on the same axis for two metrics simultaneously — annualized Sharpe ratio and annualized return. The whitepaper manifest pins the headline frontier to the top-10% rolling upper tail for both Sharpe and annualized return, while the appendix reports the surrounding cutoff sensitivity scan. The held-out window is the canonical 1258-day validation split each campaign run reports against the 2516-day training period (see [Appendix F](#)); the §5.3/§5.4 sensor and actuator panels are evaluated on their own held-out splits as documented in [Appendix G](#). The body keeps only those two local derivatives; the dollar-to-experiment production function, the human/LLM allocation split, the architecture-quality scalar  $\Gamma_t$ , the search-scaling law  $\Gamma_t = \xi \log_{10}(1 + N_t^{\text{exp}}/N_0)$ , the chain rule  $g_t^Z$ , and the R&D transition law are collected in [Appendix F](#).

**Definition 19** (Experiments-performance slope). The **experiments-performance slope** is the body primitive for R&D: selected rolling upper-tail frontiers of validation Sharpe and validation annualized return across the strict-invalid- and sealed-holdout-filtered auto-research cohort, each regressed against  $\log_{10}(1 + n)$  where  $n$  is the experiment index after sorting by run identifier. The headline derivatives  $\partial \text{Sharpe} / \partial \log_{10}(1 + n)$  and  $\partial \text{AnnReturn} / \partial \log_{10}(1 + n)$  report the gain on the selected frontier per decade of completed experiments; the matching intercepts  $\beta_{0, \text{Sharpe}}$  and  $\beta_{0, \text{AnnReturn}}$  pin the level at the start of the campaign.

$$\begin{aligned}
 \underbrace{\text{selected rolling upper-tail Sharpe frontier}}_{\widehat{\text{Sharpe}}_{\text{frontier}}(n)} &= \overbrace{\text{Sharpe}_0 + \log_{10}(1 + n) \frac{\partial \widehat{\text{Sharpe}}}{\partial \log_{10}(1 + n)}}^{\text{Sharpe-experiments slope}}, \\
 \underbrace{\text{selected rolling upper-tail annualized-return frontier}}_{\widehat{\text{AnnReturn}}_{\text{frontier}}(n)} &= \overbrace{\text{AnnReturn}_0 + \log_{10}(1 + n) \frac{\partial \widehat{\text{AnnReturn}}}{\partial \log_{10}(1 + n)}}^{\text{return-experiments slope}}.
 \end{aligned} \tag{24}$$



**Figure 6:** Validation Sharpe vs. completed auto-research experiments. Per-experiment values are the gray scatter, the muted step is the running-best context line, the shaded band is the bootstrap interval for the selected top-10% rolling upper-tail frontier, and the orange curve is the selected log-law fit.



**Figure 7:** Validation annualized return vs. completed auto-research experiments. The headline frontier is the selected top-10% rolling upper-tail fit; the running-best step remains only as muted context.

**Fitted parameters.**

| parameter   | value  | 95% CI           | $R^2$  | $n$ |
|---|--|------------------|--------|-----|
| $\beta_{0,\text{Sharpe}}$                         | 0.3921 (Sharpe)                              | [0.2258, 0.4964] | 0.7563 | 326 |
| $\partial\text{Sharpe}/\partial\log_{10}(1+n)$    | 0.3436 (Sharpe per decade of experiments)    | [0.2978, 0.4153] | 0.7563 | 326 |
| $\beta_{0,\text{AnnReturn}}$                      | 16.91 (pct points)                           | [10.42, 19.12]   | 0.4159 | 399 |
| $\partial\text{AnnReturn}/\partial\log_{10}(1+n)$ | 2.135 (pct points per decade of experiments) | [1.202, 4.792]   | 0.4159 | 399 |

---

**Equation**

$$\widehat{\text{Sharpe}}_{\text{frontier}}(n) = 0.3921 + 0.3436 \log_{10}(1+n)$$

$$\widehat{\text{AnnReturn}}_{\text{frontier}}(n) = 16.91 + 2.135 \log_{10}(1+n)$$


---

**What this means.** The two fitted derivatives say what a ten-fold increase in completed auto-research experiments buys on the selected held-out frontier:  $\partial\text{Sharpe}/\partial\log_{10}(1+n^{\text{exp}})$  Sharpe ratio points and  $\partial\text{AnnReturn}/\partial\log_{10}(1+n^{\text{exp}})$  percentage points of annualized return. Each is the headline R&D primitive in the metric the firm actually optimizes against. The fitted curves are read off the report-selected top-10% rolling upper-tail frontiers of the strict-invalid- and sealed-holdout-filtered cohort; the running-best step remains in the figures only as muted context. Starting from the current campaign count  $n_{\text{current}}^{\text{exp}}$  and projecting forward to  $n_{\text{current}}^{\text{exp}} + \Delta n$  completed experiments, the additional Sharpe purchased is  $\frac{\partial\text{Sharpe}}{\partial\log_{10}(1+n)} [\log_{10}(1+n_{\text{current}}^{\text{exp}} + \Delta n) - \log_{10}(1+n_{\text{current}}^{\text{exp}})]$ , and the additional annualized return is the analogous local increment with  $\partial\text{AnnReturn}/\partial\log_{10}(1+n)$ . This is how the R&D row enters the t-RSI numerator: a marginal R&D dollar is priced through how many additional experiments it buys ([Appendix F](#)) and how much each completed experiment is currently shifting the selected frontier along the fitted curve.

**Selection vs. structural scaling.** Across 929 auto-research experiments, the rolling top-10% frontier of held-out Sharpe rises log-linearly with completed-experiment count (slope  $\approx 0.34$  per decade). We note this is a selection statistic rather than a structural scaling law: the top- $k$ % mean of  $n$  samples from a stationary distribution would also grow with  $n$ . The substantive question is whether the underlying distribution is shifting — for which the running median (rather than the top-10%) is the cleaner diagnostic, and we treat the present fit as a directional finding to be confirmed against that diagnostic in subsequent campaigns.

**Why this fit, given that caveat.** The empirical methodology for studying auto-research is itself new: the only large-scale precedents we are aware of are the open-source single-GPU agent loop catalogued by Karpathy [41] and the autonomous-R&D-capability evaluation harness of METR [52], neither of which prescribes a settled frontier-estimation protocol. Within that gap we have chosen a rolling top-quantile with a bootstrap confidence band rather than the more common running-best step. Running-best is a strict order statistic over an ever-growing sample: it is monotone in  $n$  by construction and is exactly the selection-bias regime that the deflated-Sharpe / backtest-overfitting literature documents as untrustworthy (9, 48). A rolling top- $k$ % tail mean is a smoother, less brittle envelope of the same upper-tail behavior, and reporting a bootstrap CI alongside it surfaces the residual selection inflation directly. We do not claim this estimator is the right answer for auto-research scaling — only that it is a more defensible reading of the frontier than the running-best step at the same sample sizes, and that the eventual structural test against the running median (above) is what would convert it from a directional finding to a settled fit.

## 5.6 Parameters as an Asset

**What this channel is.** Parameters are model scale purchased through training compute. The marginal-return row sits on the joint Hoffmann/Chinchilla loss surface [37, 40]: an architecture-set noise floor, a model-size term, and a data term whose coefficients are pinned by a  $(M_t, D_t, A_t)$  sweep. The body keeps only the joint surface; the full chain-rule decomposition  $g_t^\Theta$ , the compute-cost identity, and the worked example are collected in [Appendix F](#). The complementary empirical question — how quickly deployed parameters go stale, i.e. the alpha-decay term that prices the cost of *not* refreshing  $\Theta_t$  — is what the deployed history already lets us measure, and we report that here.

**Definition 20** (Parameter marginal return). The **parameter marginal return**  $g_t^\Theta$  chains parameter dollars  $\rightarrow$  model size  $M_t \rightarrow$  predictive loss  $\rightarrow$  expected return  $\rightarrow$  objective. The model-size term is the Hoffmann/Chinchilla power law [Equation 67](#); the divisor  $D_t^{\text{seen}} K_t^{\text{pass}} / \eta_t^{\text{train}}$  is the compute-cost identity [Equation 68](#) that converts an  $M_t$  slope into a dollars slope.

$$g_t^\Theta = \frac{\partial J_t}{\partial \mu} \frac{\partial \mu}{\partial \mathcal{L}_{\text{pred}}} \left[ -\frac{\alpha_M(\Gamma_t) A_M(\Gamma_t)}{M_t^{\alpha_M(\Gamma_t)+1}} \right] \cdot \frac{\eta_t^{\text{train}}}{D_t^{\text{seen}} K_t^{\text{pass}} \rho_t^{\text{HW}}}. \quad (25)$$

**Example.** Suppose AlphaFund is currently training a  $2 \times 10^7$ -parameter EWM and is considering a  $4 \times$  scale-up at fixed  $D_t$ . Under a Chinchilla  $\alpha_M(A_t) \approx 0.34$ , the model-size contribution to  $\mathcal{L}_{\text{pred}}$  shrinks by roughly  $1 - 4^{-0.34} \approx 38\%$ . Whether that is bought depends on the matching  $D_t$  side of the joint surface; the model-size sweep that pins  $\alpha_M(A_t)$  has not yet been run, so the parameter row is recorded symbolically.

**What this means.** With the model-scaling slopes still pending, the deployed evidence speaks to the other half of the parameters row: not how steep the loss surface is in  $M_t$  today, but how fast the deployed weights  $\Theta_t$  themselves age out. That is the  $\alpha^{\text{decay}}$  term the controller charges against every cycle the firm chooses not to refresh its parameters.

**Empirical alpha-decay from parameter staleness.** The parameter row is also where the corporation books the time-rate alpha-decay term that enters (21) and the t-RSI numerator of [subsection E.2](#). Running the forecast-evaluation panel pooled across three training-run seeds — effectively approximately three independent five-month training-run studies (127–128 assets per run, three forecast horizons,  $\sim 150,000$  held-out rows per horizon over the [2020-01-01, 2025-01-01] test window) — gives a per-asset median decay rate  $\hat{\lambda}_{\text{eff}} \sim 2\text{--}4 \times 10^{-4}$  per 60-minute cycle, with the per-asset Mann–Kendall slope on MAE-skill against deployment age centered near zero across the 127-asset universe. The same qualitative picture holds in production: across  $\sim 16$  months of cumulative Mark II and Mark III live trading ([Appendix H](#)), every linear, rank, and distributional trend statistic against deployment age returns a  $\hat{\kappa}$  near zero, with several leaning in the *positive-trend* direction. *Empirically, then, the measured alpha decay sits at or below the panel’s resolution;* the small magnitude is most plausibly attributable to the firm’s low  $\sim 27 \times$  annual portfolio turnover, which holds deployed weights inside the regime where parameter drift is dominated by sampling noise. The data-derived  $\hat{\lambda}^{\text{decay}}$  that enters the headline t-RSI is therefore small and near zero.

The per-horizon decay panels, fresh-vs-stale density panels, and per-asset  $\widehat{\lambda}_{\text{eff}}$  histogram are in [Appendix F](#).

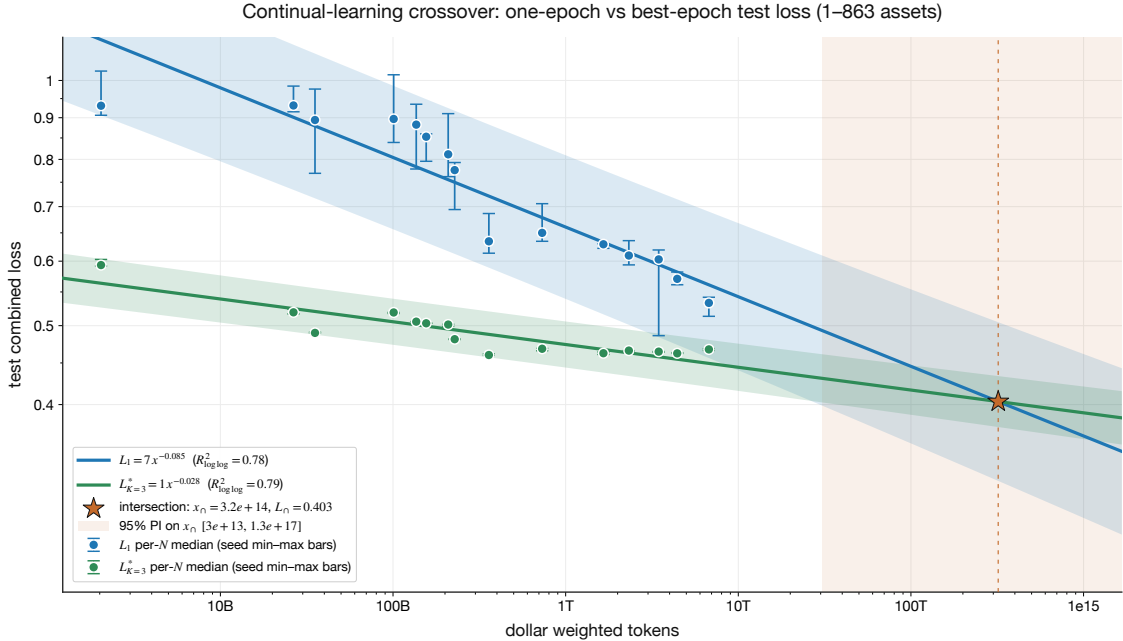
**Model-scaling sweep is forthcoming.** The  $(M_t, D_t)$  Chinchilla-style grid that pins  $\alpha_M(A_t)$ ,  $A_M(A_t)$ , the training-efficiency term, and the model-size-dominates-data crossover is the obvious next sweep on the multi-seed scaling backbone. The current data lets the optimizer price decay (small, noise-dominated) but not yet model-size headroom; that sweep is queued as the next campaign and will refresh this row’s coefficients on completion.

## 5.7 Continual Learning

**What this channel is.** The regime in which the model learns the relevant distribution in one pass, so the best validation loss after repeated epochs is no better than the loss after the first epoch. In this paper the boundary is the intersection between the one-epoch loss curve and the best-epoch loss curve. At that point  $\text{epoch}_{\text{best loss}} \rightarrow 1$ : one repetition over the data is enough for the model to converge.

This is a structural statement about the parameters channel. Once the data can be learned in a single viewing, the firm exits the static train/validation/holdout regime and enters a walk-forward regime in which the deployed weights are refit every  $t+\tau$  for some  $\tau \ll T$ , where  $T$  is the total length of the dataset. This is the same generalization regime observed at scale in modern language-model training, and operationally it is what is commonly called *continual learning*—data evaluated and trained prequentially, in chronological sequence.

Two consequences propagate forward. First, the variance of the expected-return estimator collapses: every training sample becomes a valid out-of-sample evaluation point at the next refit boundary, so the firm’s posterior on its own forward returns tightens monotonically with operating history. Second, the tighter posterior feeds back into both the architecture-search channel and the controller. Architecture search inside R&D gains confidence that a candidate represents a genuine improvement rather than a sampling fluke, and the risk-aware controller allocates more optimally across all five investment channels because  $\sigma_{c,t}$  is smaller, increasing the rate of improvement. Continual learning is therefore not just a downstream consequence of  $\text{epoch}_{\text{best loss}} \rightarrow 1$ —it is also the mechanism by which the firm’s per-row uncertainties tighten.



**Figure 8:** One-epoch test loss and best-epoch test loss versus single-pass dollar-weighted tokens. The intersection marks the estimated entry into the  $K_t \rightarrow 1$  continual-learning regime.

### Fitted parameters.

| parameter        | value                              | 95% CI                 | $R^2$  | $n$ |
|------------------|------------------------------------|------------------------|--------|-----|
| $\alpha_1$       | 0.08548 (first-epoch loss slope)   | [0.07194, 0.09903]     | 0.7805 | 45  |
| $\alpha_{K=3}^*$ | 0.02795 (best-epoch loss slope)    | [0.0236, 0.0323]       | 0.7867 | 45  |
| $x_{\cap}$       | 3.204e+14 (dollar-weighted tokens) | [3.110e+13, 1.718e+17] | N/A    | 45  |
| $L_{\cap}$       | 0.4035 (test combined loss)        | [0.3028, 0.448]        | N/A    | 45  |

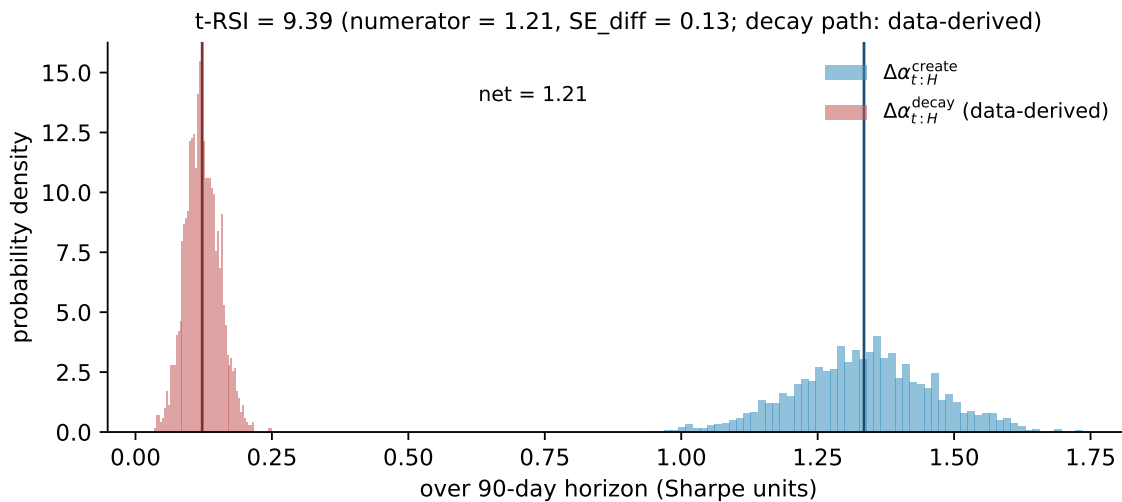
**What this means.** The chart says that first-epoch loss is falling faster than best-epoch loss as the dollar-weighted-token budget grows. If those two power laws intersect, then the model no longer needs repeated passes to reach its best loss: the first pass is enough. Operationally, that means a smaller fraction of the available history is needed for training and a larger fraction can remain genuinely held out for testing. Taken to the limit, if the distribution can be inferred from a small enough slice of the data, continual learning becomes a prequential approximation: the firm learns from the latest examples while preserving most of the data stream as evaluation.

## 5.8 Headline t-RSI

**What this channel is.** At the current operating point the firm measures a positive standardized distance between its alpha-creation rate and its alpha-decay rate over the next quarter. Read the headline as: how many standard errors of the posterior of the difference the projected net alpha increment  $\Delta\alpha_{t:H}^{\text{net}}$  sits above zero. The numerator is the difference of the posterior mean create rate and the posterior mean decay rate; the denominator is the standard error of that difference, propagated from an end-to-end bootstrap on the channel-row fits and the empirical alpha-decay panel cluster bootstrap. The headline conditions on the *data-derived* alpha-decay rate, which the

firm’s forecast-evaluation panel and the 16 months of cumulative Mark II/III live trading measure as small and near zero (every linear, rank, and distributional trend statistic against deployment age returns a  $\hat{\kappa}$  near zero, with several leaning in the positive-trend direction). The scope of the underlying extrapolation is narrow: the calculation only asks the data-scaling slope to continue roughly one and a half orders of magnitude past the in-sample  $D_{\text{eff}}$  (one decade absorbed every two months over the three-month horizon), which is well inside the 3.5-OOM in-sample range of the sweep, and integrates that extension over a single three-month horizon.

**Reading the magnitude.** The headline is large because the firm currently operates below the market-impact floor: at  $\sim$ \$400k AUM and  $\sim$ 27 $\times$  annual turnover, individual orders sit well inside the regime where they can be split without measurable price impact, and the empirical impact contribution to the create-side bootstrap is indistinguishable from zero. The data-derived alpha-decay rate is also small (subsection E.2). With neither impact nor decay materially eroding the create-side gains, the standardized distance collapses to a near-pure read of the create-side dispersion, and that read is tight. A different operating point will read differently: at  $K$ -times current AUM the literature  $\sqrt{Q/ADV}$  impact law [13, 64] reintroduces a capacity-driven Sharpe drag on the create side, and the headline compresses. Under a turnover trajectory consistent with the empirical decreasing-returns-to-scale pattern documented for active managers [27, 60], the same posterior reads 4.59 at  $K = 10\times$  AUM and 2.90 at  $K = 100\times$  AUM; under a worst-case trajectory in which annual turnover is held fixed at its current level, the headline crosses zero between  $K = 10\times$  and  $K = 20\times$  AUM. The full two-row capacity-sensitivity table is in subsection E.2. The thresholded form of this test statistic — the *certificate of monotone improvement* that gates whether a candidate update is admitted into the deployed model — is detailed in Appendix I.



**Figure 9:** Posterior of the two t-RSI components over the 90-day horizon at the current operating point, under an end-to-end bootstrap of the channel-row fits. Light blue:  $\Delta\alpha_{t:H}^{\text{create}}$ , the posterior of the firm’s alpha-creation rate over the horizon (sensors+actuators plus the gross-of-carry R&D arm; researcher compensation flows through the accounting projection, not the creation-rate functional). Light red:  $\Delta\alpha_{t:H}^{\text{decay}}$ , the posterior of the data-derived alpha-decay rate from subsection E.2 (empirically small and near zero). The vertical solid lines mark the two posterior means; the dashed annotation is their difference. The title carries the standardized-distance t-RSI (20), the numerator (difference of means), and the denominator (SE of the difference).

## 6 Toward the Differentiable Corporation

AlphaFund is an early implementation of a differentiable corporation: a company whose operational decisions are being converted into auditable marginal-return estimates against future equity growth. In this architecture, each major use of capital—trading capital, data, execution infrastructure, model parameters, and R&D—becomes a row in a common optimization problem. The controller’s task is to compare those rows on the same dollar-denominated axis, allocate capital to the highest risk-adjusted marginal return, and update the estimates as realized outcomes arrive.

The evidence in §5 measures several of these rows at the current operating point. Sensors and actuators are represented through data-performance scaling laws; R&D is represented through experiment-performance frontiers; parameters are represented through refit and decay measurements; and the combined create-versus-decay distance is summarized by t-RSI. These measurements make AlphaFund a concrete test case for the differentiable-corporation program.

### 6.1 Three Structural Facts

**Pre-existing market with mandated price discovery.** For t-RSI to be operationally measurable, the firm’s actions must be priced by an external mechanism at sub-cycle latency, with the prices observable and timestamped. Public equities satisfy this by federal mandate: every executed trade is reported and made public, timestamped to the nanosecond, against a market that has been continuously operating for longer than any other. The relevant point is not that this market is attractive — it is that the price-discovery machinery for every action the firm takes already exists, externally, at no construction cost and with no PMF term confounding  $\partial\text{Equity}/\partial\text{Company}$ . Most candidate domains fail this condition: in nascent markets the firm has to build the price-discovery surface itself, which adds a learned-mechanism term to every gradient; in regulated-but-illiquid markets the prices exist but resolve too slowly to score per-cycle decisions. Quantitative trading is unusual in that this condition holds trivially.

**Principal value capture.** t-RSI requires that  $\partial\text{Equity}/\partial\text{Company}$  be dominated by the firm’s own actions rather than by intermediated customer behavior. A trading firm is principal to its own predictions: carry is on the firm’s own positions against the market, not a per-call fee on a customer’s downstream use of a tool. The contrast with platform-AI clarifies what the condition rules out, not which business is preferable. A platform seller’s  $\partial\text{Equity}/\partial\text{Company}$  is dominated by enterprise sales cycles, competitive substitution, and product–market fit — exogenous, lagged, customer-dependent terms whose variance swamps the marginal effect of any single internal improvement. The same model improvement that earns a platform \$0.001M in API revenue might earn the principal \$100M in carry; the difference is not which firm is better-run, but which firm has a directly measurable derivative against its own actions. Most industries fail this condition because their value capture flows through a customer-decision bottleneck whose noise drowns out the create-vs-decay signal the framework needs to identify, and the standard error on the relevant gradient grows accordingly.

**API-complete operational degrees of freedom.** Every operational decision in the firm is, in principle, a function call. Data ingestion is an API. Model training is an API. Capital allocation is an API. Trade execution is an API. Asset acquisition—new data licenses, brokerage accounts, exchange memberships—is an API. Each call produces a structured log that doubles as the causal record needed for a derivative against it. A clothing manufacturer’s production function bottlenecks on physical objects (textile sourcing, factory throughput, retail distribution);  $\partial\text{Equity}/\partial(\text{material blend})$  is not even well-defined as a derivative because the action space is not

| Domain bucket                      | Theoretical LM exposure | Automation share | API completeness |
|------------------------------------|-------------------------|------------------|------------------|
| Legal Services                     | 1.00                    | 0.27             | medium           |
| Quantitative / Investment Finance  | 0.94                    | 0.76             | high             |
| Banking & Credit                   | 0.92                    | 0.70             | medium           |
| Insurance                          | 0.91                    | 0.68             | medium           |
| Accounting & Finance Consulting    | 0.91                    | 0.67             | medium           |
| Software / SaaS / IT Services      | 0.87                    | 0.68             | high             |
| Data & Cloud Infrastructure        | 0.86                    | 0.69             | high             |
| Management Consulting              | 0.81                    | 0.70             | medium           |
| Architecture, Engineering & R&D    | 0.76                    | 0.65             | medium           |
| Content, Media & Advertising       | 0.69                    | 0.71             | medium           |
| Healthcare Admin-Heavy             | 0.62                    | 0.63             | medium           |
| Physical Services & Trades         | 0.41                    | 0.67             | low              |
| Manufacturing & Industrial         | 0.40                    | 0.67             | low              |
| Logistics, Warehousing & Wholesale | 0.39                    | 0.69             | low              |
| Agriculture, Mining & Extraction   | 0.34                    | 0.69             | low              |

**Table 3:** Cross-industry exposure ranking across the 15 BEA-Detail domain buckets. *Theoretical LM exposure* is the labor-dollar-weighted AIOE LM score (Felten/Raj/Seamans). *Automation share* is the labor-dollar-weighted automation-vs-augmentation ratio from the AEI V4 (Tamkin et al., Anthropic Economic Index) classification. *API completeness* is a qualitative proxy from the upstream pipeline. Quantitative / Investment Finance (highlighted) is the only bucket in the top quintile on all three axes with *high* confidence.

differentiable. Quantitative finance is the industry where it is.

The 15-bucket cross-industry ranking below makes this exclusive overlap explicit: quantitative finance is the only domain bucket in the top quintile jointly on theoretical LM exposure (0.94), automation-versus-augmentation occupational share (0.76), and self-reported API completeness, with high confidence [25, 73]. The AEI V4 dataset (Anthropic Economic Index) covers only 10.8% of O\*NET tasks by count; the missing 89% are overwhelmingly physical or manual. The empirical frontier of LM-mediated work is therefore the frontier of API-describable work—quantitative finance lies inside that hull while manufacturing, logistics, and on-site services lie outside.

## 6.2 Channels Reinforce Each Other

The certificate of monotone improvement (Equation 71) fires on one channel at a time, but channels are linked through shared state, so marginal dollars do not decompose channel by channel. Formally, positive cross-partials of the cumulative objective on the active rows (Equation 26) yield supermodularity in the sense of Milgrom and Roberts [55]: a marginal dollar on channel  $j$  raises the marginal value of a dollar on channel  $k$ , and conversely. Local positivity can fail under capacity, attention, or saturation, so no global supermodular ordering is asserted. The continuation claim is probabilistic, not global (Equation 27): the certificate gates each commit at the prevailing operating point rather than relying on supermodularity everywhere, and while the cross-partials in Equation 26 remain positive on those rows, an observed increase in the certified-commit improvement rate raises the posterior probability that the next commit also improves the loop, with the inequality read as an empirical row-law claim rather than a theorem.

**Definition 21** (Cross-channel supermodularity). **Cross-channel supermodularity** says that the cumulative objective  $J_t$  has non-negative cross-partials on the rows the firm has identified [55]: a

marginal dollar on channel  $j$  raises (or leaves unchanged) the marginal value of a dollar on channel  $k$ . This is a local statement on the active rows, not a global theorem; capacity, attention, and saturation can flip the sign off the operating point, which is why the continuation claim of Equation 27 is probabilistic rather than supermodular-everywhere.

$$\frac{\partial^2}{\partial a_t^j \partial a_t^k} J_t = 0 \quad (26)$$

**Example.** Suppose AlphaFund’s tradable universe widens from 127 to 160 names (an actuator add) while sensors absorb another decade of dollar-weighted tokens. The actuator add lifts the marginal value of the sensor dollar—more assets means more rows the new data can sharpen forecasts on—and the sensor add lifts the marginal value of the actuator dollar by raising the per-asset  $\hat{\mu}_t$  the widened surface deploys against. The cross-partial  $\partial^2 J_t / (\partial a_t^S \partial a_t^U)$  is non-negative at the current operating point.

**Definition 22** (Certified-commit continuation bound). The **certified-commit continuation bound** is the probabilistic, row-law version of supermodularity: conditioning on a successful certified commit at cycle  $t$ —i.e. on the event  $\text{Cert}_t = 1$  from Equation 71—raises the posterior probability that the next certified commit also improves t-RSI. Read as an empirical row-law claim, not a global theorem: it says the firm’s track record of clearing the certificate is itself evidence that the next commit clears.

$$-\Pr(\Delta \text{t-RSI}_{t+1} > 0, \mathcal{F}_t) + \Pr(\Delta \text{t-RSI}_{t+1} > 0, \mathcal{F}_t, (31)) = 0 \quad (27)$$

**Example.** Suppose the firm has cleared the certificate on ten consecutive Mark III refits, each raising the held-out three-month t-RSI by a posterior-mean 0.1 standardized units. The eleventh candidate update arrives. Conditioning on the prior ten certified commits raises the posterior probability that  $\Delta \text{t-RSI}_{t+1} > 0$  above its unconditional rate; the strength of the lift is a posterior on the firm’s own row-law trajectory rather than a structural guarantee.

### 6.3 Drift Detection and Recovery

A differentiable corporation that keeps allocating through regime change needs three things; only the first is what section 5 measures in isolation. That first requirement is fitted marginal-return laws with measurable standard errors on every operational degree of freedom. The second is drift detection on the world model itself—an operational trigger when live inputs move outside the support on which those standard errors were identified. The third is R&D throughput high enough that, once that trigger fires, the firm refits to the new regime before an obsolete mapping bleeds out deployable capital. The certificate of Equation 71 is evaluated on the same audited moments, so its economic content weakens exactly when drift invalidates the inferential basis for those errors.

On (2), the alpha-decay panel of §5.8 already constitutes this surveillance and measures the decay rate as small and near zero on its measured horizon. On (3), the continual-learning construction of §5.7 implies a refitting time constant that is short relative to the decay time constant summarized in (2), at the firm’s prevailing data-ingestion rate. The firm’s recovery rate therefore dominates its measured decay rate, which is what self-improvement under regime non-stationarity actually requires.

**Definition 23** (Deployable-capital decomposition). The **deployable-capital decomposition** splits

next-period equity into the slice generated internally by the realized cycle reward on the existing book ( $K_{t+1}^{\text{int}}$ ) and the slice supplied externally by outside investors against the firm’s certificate-cleared track record ( $K_{t+1}^{\text{ext}}$ ). External capital amplifies the loop only while each marginal externally supplied dollar clears the same risk-adjusted certificate of Equation 71 after financing costs, dilution, market impact, and capacity effects.

$$K_{t+1} = K_{t+1}^{\text{ext}} + K_{t+1}^{\text{int}} \tag{28}$$

**Example.** Suppose at cycle  $t$  shareholders’ equity is  $K_t = \$25\text{M}$  and the realized log-return on the existing book over the cycle is  $R_t = 0.04$ , so  $K_{t+1}^{\text{int}} = K_t e^{R_t} \approx \$26.02\text{M}$ . A certificate-cleared  $\$5\text{M}$  outside-equity round closes that cycle ( $K_{t+1}^{\text{ext}} = \$5\text{M}$ ), giving  $K_{t+1} \approx \$31.02\text{M}$ . If the marginal  $\$5\text{M}$  had failed the certificate (financing cost above shadow price, or market-impact erosion past the capacity floor) the firm would have declined the round; the loop would have continued amplifying only off  $K_{t+1}^{\text{int}}$ .

## 6.4 The Bitter Lesson for Capital

Sutton’s bitter lesson observes that the methods which win at scale in machine learning are the ones whose performance grows with computation, not the ones that encode hand-crafted human structure [72]. The capital analog is structurally identical and historically older: the firms that win at scale are the ones whose throughput grows with absorbed capital, not the ones whose decisions are bounded by the size of a fixed staff. The supermodular cross-partials of §6.2 compound the internal loop at the rate the channels permit; the bitter lesson is that the loop does not stop there.

Once the corporation demonstrates, in its own operating record, that incremental capital reliably converts into measured row improvement and deployable edge, outside investors can treat that relationship itself as an investable object. The next-period deployable-capital identity (Equation 28) separates internally generated deployable capital  $K_{t+1}^{\text{int}}$  from externally supplied capital  $K_{t+1}^{\text{ext}}$ . External capital amplifies the loop only while each marginal externally supplied dollar clears the same risk-adjusted certificate after financing costs, dilution, market impact, and capacity effects.

Two things are worth mentioning. First, different financing instruments have norms—and in some cases strict bylaws—governing what they can and cannot invest in. As the world gets increasingly digitized, the instruments with more leeway and more quantitative discipline will, on average, out-compete those that cannot process as much information; this is its own selection pressure on the type of external capital a recursively improving corporation can absorb. Second, a self-improving corporation with improving capability attracts more capital; when the marginal certificate continues to clear, that increased capital further increases the rate of self-improvement, which increases the rate of returns, which in turn increases the rate of external investment. Some of that capital—particularly AUM and equity—does carry per-dollar performance decay through market impact and capacity. What matters is the race: when incoming capital buys data or architectures that raise expected return faster than market impact erodes alpha, the rate of self-improvement continues to climb and the rate of external investment accelerates with it—a positive feedback loop.

## 6.5 Completion Roadmap

The next phase is to close the remaining operational gradients. Salary and headcount cost, banking and cost of capital, hardware procurement and depreciation, asset-acquisition cost, and AUM acquisition cost are the remaining major terms in the corporate equation. Each term has a natural

measurement surface: dollars in, operational capability out, and realized contribution to future equity growth. As those surfaces are instrumented, the corporation becomes progressively more legible to its own controller.

The completed object is a firm whose capital-allocation process is scored end to end: every major expenditure has a forecast, every forecast has a realized outcome, and every realized outcome updates the next allocation. The differentiable corporation is the limit of that process.

## 6.6 Beyond Quant Trading

The Economic World Model is a network trained on the joint distribution of priced economic data—a foundation model for allocation in the same sense that a large language model is a foundation model for text. Text describes economic activity; prices settle it; the same underlying forecast can therefore be executed at more than one depth in the real economy rather than only as a paper position.

The controller can be instantiated in three computational regimes, each removing one piece of hand engineering. The first is the one this blueprint develops in detail: a hand-factored chain rule whose per-channel scaling laws compose into a one-step gradient and equilibrate against the marginal ROI  $\lambda_t^*$ . The second replaces the hand factorization with a single neural Economic World Model trained end-to-end on the firm’s operating history;  $J_t$  becomes the discounted sum of predicted log-equity returns under a finite MPC roll-forward and  $\nabla_{\mathbf{a}_t} J_t$  is recovered by autodifferentiation through the rollout [10, 62]. The third parameterizes the allocation policy and lets gradients of cumulative log-equity flow back through the world model and into the policy itself: the differentiable-simulator paradigm of PILCO [23], the Dreamer family [32–34], and the broader differentiable-world-model line [31, 66]. The certificate of monotone improvement (Equation 71) extends across all three regimes: each candidate update is admissible only when the held-out t-RSI clears the Sharpe-margin threshold  $\delta$  and the Fisher-information readiness floor  $\varepsilon_c$  on every active channel.

These regimes describe how the controller reasons; the question of *what it acts on* is orthogonal and develops along its own axis. If we take intelligence to be the capacity to acquire, preserve, and compound command over future resources, those resources need not arrive through a brokerage API. A forecast over future supply and demand has value because it implies future prices, constraints, and margins. The shallowest channel that realizes a predicted edge is a futures position; a deeper channel realizes the same edge as a sequence of actuators in  $\Xi_t$  (procurement, transport, refining, wholesale distribution) and captures the entire margin between input and output rather than the basis alone. Each version is the same economic prediction executed at a different depth, scored by the same per-dollar log-growth functional that the portfolio optimizer compares across heterogeneous channels in cycle  $t$ . The depth at which the firm operates is not a strategic preference but a Coasean comparison the framework already makes [20]: the firm verticalizes the next step of a production chain when two conditions hold together—the executing channel clears the certificate of Equation 71, and the certificate-corrected risk-adjusted cost of running that step in-house is lower than the market price of buying the same output. New actuators—humanoid platforms, AI agents, autonomous logistics services—enter the firm’s action set as soon as their success distributions are reliable enough for the EWM to learn and condition on. Because deeper execution demands a more capable  $\widehat{\mathcal{W}}$ , the regimes and the depth axis are coupled: better reasoning unlocks deeper action, and deeper action enriches the operating record that the next regime trains on. At sufficient depth and scale the small-firm approximation no longer holds: a firm whose actions move corn prices, contract refining capacity, or absorb a meaningful slice of the financing pool perturbs  $\mathbf{E}_{t+1}$  by its own choices. This is not a

contradiction of the framework; it is absorbed by it. The firm’s filtration (Equation 11) accumulates the joint record of its own actions and the world’s responses to them, so the EWM trained on  $\mathcal{F}_t$  can in principle learn the reaction term that the small-firm limit zeroed out. Operationally the change is local: the same channel-row laws are refit on a richer history, and the certificate continues to gate every commit.

The standardized create-vs-decay distance, t-RSI (Equation 20), is computable at AlphaFund. We believe quant trading is among the first domains where such a statistic is practically computable. A clothing retailer asking “does this material blend increase next-quarter equity” faces a numerator whose sign is contested between marketing-campaign quality, retail traffic, supply-chain cost, and the broader economic cycle; the standard error on the marginal effect is so large relative to the marginal effect itself that t-RSI is, in practice, undefined. A platform company selling tools—advertising slots, API tokens, software seats—inserts a product–market-fit layer between the loss its model minimizes and the revenue its balance sheet collects; the bridge is opaque, lagged, and customer-dependent, and the implied  $\partial\text{Equity}/\partial(\text{model improvement})$  is dominated by exogenous customer behavior rather than by the model. In both cases the underlying problem is structural: the standard errors on  $\partial\text{Equity}/\partial\text{Company}$  are large relative to the effects they would measure at any reasonable sample size. The moat is therefore not the value of AlphaFund’s t-RSI. It is that t-RSI is practically computable in this firm and this industry. The single scalar that summarizes the legibility of a corporation to itself is, at the time of writing, much harder to construct elsewhere.

## 7 Conclusion

A differentiable corporation is a measurement architecture for compounding economic intelligence. It turns the firm from a collection of departments into a set of capital-allocation gradients, each estimated against future equity growth and updated through realized outcomes. AlphaFund’s current system implements the first measurable version of this architecture in quantitative trading. The path forward is to expand the measured channel rows, audit them prospectively, and use the resulting equation as the operating system for corporate self-improvement.

## A Accounting Bridge

### A.1 Accounting Projection

**Definition 24** (Accounting projection). The **accounting projection**  $\pi^{\text{acct}}$  partitions the corporation into the four scalar quantities the cash-management side of the controller needs to clear its budget: the equity that compounds, the slice of equity deployable this cycle, the cash numeraire, and the reserves held back against constraints:

$$\Xi_t \xrightarrow{\pi^{\text{acct}}} (K_t, K_t^{\text{deploy}}, \text{Cash}_t, K_t^{\text{reserve}}). \quad (29)$$

| Symbol                 | Quantity             | What it is  |
|------------------------|----------------------|---|
| $K_t$                  | Shareholders' equity | Total assets minus total liabilities, marked to current dollars. The scalar the firm compounds.                           |
| $K_t^{\text{deploy}}$  | Deployable capital   | The slice of equity the controller is free to allocate this cycle, after reserves and constraint set-asides.              |
| $\text{Cash}_t$        | Cash                 | Settled, unencumbered cash on the balance sheet. The numeraire every $a_t^k$ is denominated in.                           |
| $K_t^{\text{reserve}}$ | Reserves             | Capital held back to satisfy the liquidity, solvency, and channel-liquidation constraints of <a href="#">Appendix B</a> . |

**Table 4:** Accounting projection  $\pi^{\text{acct}}(\Xi_t)$  components. The four scalars are the cash-management slice of the firm's books; the corresponding double-entry GAAP/SEC chart of accounts is left to the firm's external balance-sheet simulation.

## A.2 Deployable Capital and Flow of Funds

The accounting projection of [Equation 29](#) names the four scalars the controller needs; this subsection defines the two that move every cycle. [Equation 30](#) fixes the *stock*: how much of equity is free to allocate after committed positions and reserves. [Equation 31](#) and [Equation 32](#) are the *flow*: how realized cash from market activity, operating expenses, taxes, and reserve adjustments push the stock from one cycle to the next. The dot on  $\dot{\mathcal{E}}_t^{\text{retain}}$  marks a per-cycle flow into retained earnings; the  $\Delta$  on  $\Delta K_t^{\text{deploy}}$  is the resulting cycle-over-cycle change in the deployable stock.

**Definition 25** (Deployable capital). **Deployable capital**  $K_t^{\text{deploy}}$  is the slice of shareholders' equity the controller is free to allocate in cycle  $t$ . It is total equity less the equity already committed to current positions, capitalized intangibles, and locked subscriptions ( $K_t^{\text{committed}}$ ), and less the reserves held back against the liquidity, solvency, and channel-liquidation constraints of [Appendix B](#) ( $K_t^{\text{reserve}}$ ):

$$K_t^{\text{deploy}} = -K_t^{\text{committed}} - K_t^{\text{reserve}} + K_t \quad (30)$$

**Example.** Suppose AlphaFund's shareholders' equity is  $K_t = \$25\text{M}$ . Of that,  $K_t^{\text{committed}} = \$18\text{M}$  is tied up in current trading positions, capitalized data licenses, and the GPU fleet, and  $K_t^{\text{reserve}} = \$5\text{M}$  is held against margin requirements and the solvency floor. Then  $K_t^{\text{deploy}} = 25 - 18 - 5 = \$2\text{M}$  is the budget the next-cycle allocator clears against in [\(33\)](#).

**Definition 26** (Retained-earnings flow). The **retained-earnings flow**  $\dot{\mathcal{E}}_t^{\text{retain}}$  is the cash-basis per-cycle change in retained earnings: realized cash from market activity  $Y_t^{\$}$  (the dollar output of the investments channel), plus non-operating cash income  $\text{NonOp}_t^{\text{cash}}$  (interest, dividends received), less cost of revenue  $\text{COGS}_t$  (data, inference, execution-side personnel), less operating expenses  $\text{OpEx}_t$  (R&D salaries, G&A, compliance), less taxes  $\text{Tax}_t$ , less distributions  $\text{Div}_t$  (dividends paid, buybacks). The dot, rather than a  $\Delta$ , marks this as the instantaneous cycle- $t$  flow into retained earnings; the corresponding stock change is recorded by [Equation 32](#).

$$\dot{\mathcal{E}}_t^{\text{retain}} = -\text{COGS}_t - \text{Div}_t + \text{NonOp}_t^{\text{cash}} - \text{OpEx}_t - \text{Tax}_t + Y_t^{\$} \quad (31)$$

**Example.** Suppose in a given cycle AlphaFund books  $Y_t^{\$} = \$0.40\text{M}$  of realized trading PnL,  $\text{COGS}_t = \$0.05\text{M}$  (data and inference cost),  $\text{OpEx}_t = \$0.20\text{M}$  (R&D salaries and G&A),  $\text{NonOp}_t^{\text{cash}} = \$0.01\text{M}$  (interest on cash),  $\text{Tax}_t = \$0.02\text{M}$ , and  $\text{Div}_t = 0$ . Then  $\mathcal{E}_t^{\text{retain}} = 0.40 - 0.05 - 0.20 + 0.01 - 0.02 - 0 = \$0.14\text{M}$  is the cycle- $t$  flow into retained earnings.

**Definition 27** (Change in deployable capital). The **change in deployable capital**  $\Delta K_t^{\text{deploy}}$  is the cycle-over-cycle net change in the deployable stock: the retained-earnings flow  $\mathcal{E}_t^{\text{retain}}$  from Equation 31, less the cash flow into reserves  $\Delta \text{Res}_t$  (margin top-ups, regulatory floor expansion). The  $\Delta$  converts the per-cycle flow into the corresponding stock change that feeds the next cycle’s budget in (33):

$$\Delta K_t^{\text{deploy}} = -\Delta \text{Res}_t + \mathcal{E}_t^{\text{retain}} \quad (32)$$

**Example.** Continuing the previous example, suppose margin requirements grow by  $\Delta \text{Res}_t = \$0.04\text{M}$  this cycle (the firm’s open positions widened and exchange-set initial margin rose). Then  $\Delta K_t^{\text{deploy}} = 0.14 - 0.04 = \$0.10\text{M}$ : the deployable stock  $K_{t+1}^{\text{deploy}}$  rises by  $\$0.10\text{M}$ , which is what the controller’s next-cycle budget reflects.

## B Program Constraints

### B.1 Budget, Liquidation, Liquidity, Solvency

**Definition 28** (Budget constraint). The **budget constraint** says that the total dollars allocated across channels in cycle  $\tau$  cannot exceed the deployable capital available that cycle:

$$\sum_k a_\tau^k \leq K_\tau^{\text{deploy}}. \quad (33)$$

**Example.** Suppose AlphaFund has  $K_\tau^{\text{deploy}} = \$2\text{M}$  available after reserves and existing commitments. If the controller proposes  $\$900\text{K}$  to investments,  $\$400\text{K}$  to sensors,  $\$300\text{K}$  to parameters, and  $\$250\text{K}$  to R&D, the allocation clears: the total is  $\$1.85\text{M}$ . A  $\$2.2\text{M}$  plan fails the budget constraint before any marginal-return calculation matters.

**Definition 29** (Channel-liquidation constraint). The **channel-liquidation constraint** gives each channel a floor on how negative its allocation can be in one cycle. Negative allocations free capital, but only down to what can actually be liquidated from that channel:

$$a_\tau^k \geq \underline{a}_\tau^k, \quad \underline{a}_\tau^k \leq 0. \quad (34)$$

**Example.** Suppose the firm wants to free cash from the sensor channel by canceling data contracts. If only  $\$50\text{K}$  of subscriptions can be canceled this month, then  $\underline{a}_\tau^{\text{S}} = -\$50\text{K}$ . The controller may set  $a_\tau^{\text{S}} = -\$25\text{K}$ , but not  $-\$200\text{K}$ .

**Definition 30** (Liquidity constraint). The **liquidity constraint** requires the cash line item to stay above an operational floor in every cycle. It is separate from solvency: a firm can have positive

equity and still fail because it cannot meet near-term cash obligations:

$$\text{Cash}_\tau \geq \text{Cash}^{\min}. \quad (35)$$

**Example.** Suppose AlphaFund has positive equity but must keep  $\text{Cash}^{\min} = \$250K$  on hand for payroll, cloud bills, and margin calls. A candidate allocation that leaves only \$100K cash is rejected even if the balance sheet remains solvent.

**Definition 31** (Solvency constraint). The **solvency constraint** requires total assets to remain greater than total liabilities. If equity reaches zero, the log-equity reward is no longer defined and the firm has left the domain of the objective. The drawdown form of this requirement connects to coherent risk theory [17]:

$$\text{Assets}_\tau - \text{Liabilities}_\tau > 0. \quad (36)$$

**Example.** Suppose the firm has \$10M of assets and \$8M of liabilities. It is solvent. If a leveraged position falls far enough that assets drop to \$7.5M while liabilities remain \$8M, equity is negative and the policy has driven the firm through the solvency boundary.

## C EWM Details

### EWM training and proper scoring.

The first two subsections collect the EWM-specific training derivations: the population KL objective the firm would minimize if it knew the true law, and the held-out empirical proper-scoring surrogate it actually trains against.

#### C.1 EWM Training Objectives

**Definition 32** (EWM training objective (population)). The **population EWM training objective** is the expected Kullback–Leibler divergence between the true cycle- $\tau$  joint law of next observation and reward and the EWM’s forecast of that same law, averaged over decision times. Driving  $\mathcal{L}_{\text{EWM}}$  to zero would mean the EWM has recovered the filtration-respecting predictive law of what the firm will see next and what reward it will receive.

$$\mathcal{L}_{\text{EWM}} = \mathbb{E}_\tau \left[ \text{KL} \left( \mathcal{P}_\tau^{\text{true}}(\mathbf{o}_{\tau+1}, R_{\tau+1} \mid \mathcal{F}_\tau, \mathbf{a}_\tau), \hat{\mathcal{P}}_\tau(\mathbf{o}_{\tau+1}, R_{\tau+1} \mid \mathcal{F}_\tau, \mathbf{a}_\tau) \right) \right] \quad (37)$$

**Example.** Suppose the next-cycle target is a joint pair: the oil-market observation  $\mathbf{o}_{\tau+1}$  and the realized cycle- $\tau$  reward  $R_\tau$  after the firm’s futures allocation. The EWM commits to  $\hat{\mathcal{P}}_\tau$  before those quantities resolve; the KL term compares that forecast to the true conditional law. The true law is not observed directly, so Equation 38 supplies the empirical proper-scoring surrogate.

**Definition 33** (Empirical EWM estimator). The **empirical EWM estimator** is the proper-scoring-rule sum over a held-out evaluation index  $\mathcal{I}_{\text{eval}}$ . Choosing a proper rule (e.g. negative log-likelihood, CRPS, or the energy score) makes minimizing  $\hat{\mathcal{L}}_{\text{EWM}}$  asymptotically equivalent to minimizing the population KL of Equation 37, provided  $\mathcal{I}_{\text{eval}}$  contains no information resolved after

the training cutoff. Each held-out row scores the EWM’s joint forecast against the realized next observation and reward.

$$\widehat{\mathcal{L}}_{\text{EWM}} = \sum_{\tau \in \mathcal{I}_{\text{eval}}} \ell\left(\widehat{\mathcal{P}}_{\tau}(\mathbf{o}_{\tau+1}, R_{\tau+1} \mid \mathcal{F}_{\tau}, \mathbf{a}_{\tau}), (\mathbf{o}_{\tau+1}, R_{\tau+1})\right) \quad (38)$$

**Example.** Suppose the current evaluation window is the most recent 245 trading days held strictly after the training cutoff, and the universe is the 128 most-liquid common assets. Each held-out day supplies the realized next observation vector and reward sample; the proper score sums those realized samples against the forecast law the EWM emitted before the day resolved.

## C.2 From Population KL to Empirical Proper Scoring

In practice the firm cannot evaluate Equation 37 directly: it does not know  $\mathbb{P}_{\tau}^{\text{true}}$ . What it has are realized samples  $(\mathbf{o}_{\tau+1}, R_{\tau})$  drawn from  $\mathbb{P}_{\tau}^{\text{true}}$ , and a standard fact: minimizing the expected log-score of a candidate density against samples from  $\mathbb{P}$  is equivalent—up to the entropy of  $\mathbb{P}$ , which is constant in the candidate  $\widehat{\mathbb{P}}$ —to minimizing  $\text{KL}(\mathbb{P} \parallel \widehat{\mathbb{P}})$  [28]. The empirical estimator the firm actually trains is the proper-scoring-rule sum of Equation 38; point forecasts such as  $\widehat{R}_{\tau}$  are summaries of  $\widehat{\mathbb{P}}_{\tau}$ . Driving  $\widehat{\mathcal{L}}_{\text{EWM}}$  down on a properly held-out evaluation index  $\mathcal{I}_{\text{eval}}$  is the firm’s empirical handle on the population KL objective.

### Filtration discipline.

The next subsection is filtration-specific: it explains why the EWM is conditioned on the firm filtration  $\mathcal{F}_t$  rather than the latent joint state, what filtration enlargement means under sensor spend, and the no-peeking discipline that separates an EWM from a static language model.

## C.3 Why the Conditioning is on the Filtration

In Equation 8 the EWM is fed  $\mathcal{F}_t$ , not the latent joint state  $(\mathbf{\Xi}_t, \mathbf{E}_t)$  that the true law  $\mathcal{W}$  of Equation 6 takes as input. The asymmetry is there because the firm does not actually have access to  $(\mathbf{\Xi}_t, \mathbf{E}_t)$ . Two things go wrong. First, the environment  $\mathbf{E}_t$ —prices, order flow, regime variables, counterparty behavior, the news cycle—is observed only through the noisy projections that reach the firm’s sensors. Second, even the firm’s own state  $\mathbf{\Xi}_t$  is not directly observed: the dollar marks  $a_t^k$  that compose it are produced by the mark-to-market projection of Equation 29, not by realized trades, so quantities such as the firm’s exact equity, the fair value of its positions, or the replacement cost of its model weights are themselves estimates carrying uncertainty. The firm only knows what its balance sheet is worth when it actually liquidates a position into the market.

What the firm has instead are sensor observations of both the environment and itself, accumulated over time. In this partially-observed setting the most informative summary the firm has access to is the history of those observations [7, 11, 63, 76]. A single observation is a noisy projection of the latent joint state, and the firm’s view of  $\mathbf{E}_{\tau}$  collects market data, order-book state, corporate actions, execution telemetry, financing constraints, filings, news, and any alternative-data streams the firm has bought into. Its view of itself collects the firm’s own actions, broker statements, mark-to-market valuations of every line of  $\mathbf{\Xi}_{\tau}$ , and the realized log-equity reward  $R_{\tau}$  of Equation 2.

**Concrete examples in and out of  $\mathcal{F}_t$ .** At any decision time  $t$ , events like “last Friday’s NVDA close exceeded \$500,” “the firm’s GPU cluster grew from 64 to 128 cards in March,” and “last quarter’s options-flow feed produced a 3% lift in the model’s Sharpe” are inside  $\mathcal{F}_t$ ; events like “next Wednesday’s CPI print exceeds 3.5%,” “the architecture search the firm is about to launch will improve loss by more than 5%,” and “the trade the firm is about to place will close in profit” are outside it.

**Filtration enlargement under sensor spend.** The firm filtration of Equation 11 is not static. When the firm spends capital on the sensor channel  $\mathbf{S}$ —buying a new data feed, lengthening its historical archive, or sampling at finer resolution—it strictly enlarges  $\mathcal{F}_t$  in the standard probability-theory sense: the post-purchase filtration admits events that the pre-purchase filtration could not resolve. The data-scaling slope of Equation 53 is the empirical rate at which spending dollars on the sensor channel reduces the population KL of Equation 37: better filtration  $\Rightarrow$  tighter forecasts  $\Rightarrow$  lower KL.

**Filtration discipline.** The EWM must respect the decision-time filtration of Equation 11. A forecast for cycle  $t$  may condition on  $\mathcal{F}_t$ ; it may not condition on information revealed after the decision is made. The same discipline applies to features, labels, retrieved context, backtests, validation windows, and model-selection procedures. General language models [40] may enter the system as proposal mechanisms or components of the research process, but economic prediction is evaluated by chronological, filtration-respecting outcomes.

## D Portfolio Optimization

### D.1 Shadow Price of Capital

For a single-cycle allocation problem, the Lagrangian associated with the deployable-capital constraint is  $\mathcal{L}_t(\mathbf{a}_t, \lambda_t) := J_t - \lambda_t \left( \sum_k a_\tau^k - K_\tau^{\text{deploy}} \right)$ . The Karush–Kuhn–Tucker conditions yield, for every channel funded at the optimum, the equimarginal identity below; for every unfunded channel, the corresponding inequality  $g_t^{k*} \leq \lambda_t^*$  holds. The multiplier  $\lambda_t^*$  is the **shadow price of capital**: the marginal  $J_t$  produced by one additional dollar of deployable equity, regardless of where it is spent.

$$g_t^{\mathbf{I}} = \lambda_t^* \tag{39}$$

For every unfunded channel,  $g_t^{k*} \leq \lambda_t^*$  with  $a_t^{k*} = 0$ . If any channel’s marginal return per dollar exceeds  $\lambda_t^*$ , capital is underallocated to it and the controller’s next step raises that allocation until the equimarginal identity holds again.

### D.2 Markowitz Mean–Variance Form

Fix a cycle  $t$  and let  $\hat{\mathbf{g}}_t = (\hat{g}_t^k)_k \in \mathbb{R}^K$  be the vector of posterior-mean marginal returns, and  $\Sigma_t \in \mathbb{R}^{K \times K}$  the posterior covariance of those returns across channels [14, 49]. The controller solves

the standard mean–variance program

$$\mathbf{a}_t^* = \arg \max_{\mathbf{a} \in \mathbb{R}_{\geq 0}^K} \left\{ \widehat{\mathbf{g}}_t^\top \mathbf{a} - \frac{1}{2} \kappa_t \mathbf{a}^\top \boldsymbol{\Sigma}_t \mathbf{a} \right\} \quad \text{s.t.} \quad \mathbf{1}^\top \mathbf{a} \leq K_t^{\text{deploy}}, \quad a^k \geq \underline{a}_t^k, \quad (40)$$

where  $\kappa_t > 0$  is the firm’s effective risk-aversion (the Arrow–Pratt curvature of  $J_t$  around the operating point),  $K_t^{\text{deploy}}$  is deployable equity, and  $\underline{a}_t^k$  are channel-level floors (lumpy hires, minimum subscription tiers, contractual leases). With  $\boldsymbol{\Sigma}_t \succ 0$  the objective is strictly concave and the feasible set is polyhedral, so (40) has a unique optimum [14]. Diagonal entries  $\Sigma_t^{kk} = (\sigma_t^k)^2$  are the per-channel dispersions of the row posteriors  $\widehat{P}_t^k$ ; off-diagonal entries are the cross-channel covariances and are a future empirical refinement (a diagonal  $\boldsymbol{\Sigma}_t$  recovers an independent-channel solve and is the operational default).

### D.3 Black–Litterman Form with EWM Views

The Markowitz form treats  $\widehat{\mathbf{g}}_t$  as a point estimate. The Black–Litterman construction makes the EWM’s role explicit by combining a market-equilibrium prior with the EWM’s forecast as a set of views [12, 53]. Let  $\boldsymbol{\Pi}_t \in \mathbb{R}^K$  be the prior mean of marginal returns and let the EWM forecast supply  $q$  linear views

$$\mathbf{P} \mathbf{g}_t = \mathbf{q}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_t), \quad (41)$$

where  $\mathbf{P} \in \mathbb{R}^{q \times K}$  picks out the channels each view ranges over,  $\mathbf{q}_t = \widehat{\mathbf{g}}_t^{\text{EWM}}$  is the EWM’s forecast, and  $\boldsymbol{\Omega}_t$  encodes the EWM’s posterior covariance. Combining the prior  $\mathbf{g}_t \sim \mathcal{N}(\boldsymbol{\Pi}_t, \tau \boldsymbol{\Sigma}_t)$  with the views via Bayes’ rule gives the closed-form Black–Litterman posterior

$$\widehat{\mathbf{g}}_t^{\text{BL}} = \left[ (\tau \boldsymbol{\Sigma}_t)^{-1} + \mathbf{P}^\top \boldsymbol{\Omega}_t^{-1} \mathbf{P} \right]^{-1} \left[ (\tau \boldsymbol{\Sigma}_t)^{-1} \boldsymbol{\Pi}_t + \mathbf{P}^\top \boldsymbol{\Omega}_t^{-1} \mathbf{q}_t \right], \quad (42)$$

with  $\tau \in (0, 1]$  a scalar credence on the prior. The controller then plugs  $(\widehat{\mathbf{g}}_t^{\text{BL}}, \widehat{\boldsymbol{\Sigma}}_t^{\text{BL}})$  into (40) in place of  $(\widehat{\mathbf{g}}_t, \boldsymbol{\Sigma}_t)$ .  $\boldsymbol{\Omega}_t \rightarrow 0$  (perfect EWM) reproduces the Markowitz solve over the EWM forecasts alone;  $\boldsymbol{\Omega}_t \rightarrow \infty$  (no EWM signal) collapses to the equilibrium prior.

### D.4 Multi-Period Rollout and the Sharpe-Equimarginal Limit

The multi-period program is the model-predictive rollout of (40) over a finite horizon  $T$  [10, 15, 62]: at cycle  $t$  the controller solves

$$\{\mathbf{a}_{t+s}^*\}_{s=0}^{T-1} = \arg \max_{\{\mathbf{a}_{t+s}\}} \sum_{s=0}^{T-1} \gamma^s \left\{ \widehat{\mathbf{g}}_{t+s}^\top \mathbf{a}_{t+s} - \frac{1}{2} \kappa_{t+s} \mathbf{a}_{t+s}^\top \boldsymbol{\Sigma}_{t+s} \mathbf{a}_{t+s} \right\} \quad (43)$$

subject to the budget, floor, and inter-cycle state-transition constraints under  $\widehat{\mathcal{W}}_t$ . The optimizer commits only  $\mathbf{a}_t^*$ , observes  $R_t$  and the realized next state, refits  $\widehat{\mathcal{W}}_t$  on the augmented histories, and re-solves at  $t + 1$ . The first-order conditions of (40) (or of any single cycle of (43)) recover the operational form the controller actually deploys: writing the Lagrangian with multiplier  $\lambda_{S,t}^*$  on the budget constraint, every funded channel satisfies

$$\widehat{g}_t^k - \kappa_t (\boldsymbol{\Sigma}_t \mathbf{a}_t^*)^k = \lambda_{S,t}^* \quad \text{for all } k \text{ with } a_t^{k*} > 0. \quad (44)$$

With diagonal  $\boldsymbol{\Sigma}_t$  and a homogeneous risk scale, this collapses to the Sharpe-equimarginal rule  $\widehat{g}_t^k / \sigma_t^k = \lambda_{S,t}^*$ , which generalizes the bare equimarginal identity [49, 68]. As the variance penalty vanishes ( $\kappa_t \rightarrow 0$ ), (44) collapses to  $\widehat{g}_t^k = \lambda_t^*$  and recovers the risk-neutral shadow price exactly.

## D.5 Receding-Horizon Use of the EWM

The portfolio optimizer  $\mathcal{G}$  of Equation 15 consumes  $\widehat{\mathcal{W}}_t$  in a receding-horizon loop. The object it would solve in principle is the Bellman recursion over the joint state,

$$V(\Xi, \mathbf{E}) = \sup_{\mathbf{a}} \left\{ R(\Xi, \mathbf{E}, \mathbf{a}) + \gamma \mathbb{E}[V(\Xi', \mathbf{E}')] \right\}, \quad (45)$$

the standard dynamic-programming statement of the corporate optimization in Equation 15 [15, 62]. The factor  $\gamma$  is not a time-preference discount:  $J_t$  in Equation 3 is undiscounted. It is a stand-in for the increasing per-cycle cost of further rollouts paired with the decreasing residual improvement those rollouts deliver, which is what bounds the planning horizon  $T$  in practice. The recursion is the philosophical content of self-improvement: the firm chooses allocations partly for how they reshape its own future cognitive configuration via  $\Theta$ . Closed-form  $V$  is intractable for any realistic firm—the joint state space  $\Xi \times \mathbf{E}$  is high-dimensional, partially observed, and non-stationary.

The firm therefore solves the Bellman recursion approximately via model-predictive control [14, 15, 62]: at each cycle  $\mathcal{G}$  rolls candidate trajectories forward under  $\widehat{\mathcal{W}}_t$  over the finite horizon  $T$ , finds the trajectory that maximizes the truncated cumulative objective  $J_t$  of Equation 3, executes the first allocation  $\mathbf{a}_t^*$ , observes  $R_t$  and the realized next state, refits  $\widehat{\mathcal{W}}_t$  on the augmented histories  $\{\mathcal{H}_{t+1}^k\}_k$ , and re-solves. This is the standard equivalence between approximate dynamic programming, MPC, and model-based reinforcement learning when the transition is learned from data [15, 31, 33, 62, 66].

## E t-RSI Details

### E.1 t-RSI Measurement Conventions

This appendix collects the bookkeeping conventions used by the body’s t-RSI calculation in Equation 20 and the headline three-month calculation that follows. The full step-through (numerator decomposition, dispersion identity, LOO  $R^2$  check, and end-to-end bootstrap protocol) lives in the formal paper’s three-month-t-RSI appendix; the conventions here are what the channel-row fits below assume so the numerator and denominator compose without unit drift.

**Filtration discipline.** t-RSI is a held-out statistic. Both the numerator  $\mathbb{E}[\Delta\alpha_{t:H}^{\text{net}} \mid \mathcal{F}_t]$  and the dispersion  $\mathcal{U}_t(\Delta\alpha_{t:H}^{\text{net}})$  are computed under the same no-peeking discipline that governs the EWM training loss (Equation 38): every channel-row fit that feeds the t-RSI calculation is required to be reproducible from  $\mathcal{F}_t$  alone, with no information resolved after the cycle- $t$  cutoff entering either the slope estimate or its confidence band.

**Horizon  $H$ .** The horizon is fixed in calendar time, not in operating cycles, so the calculation composes with arbitrary cycle frequencies. The headline calculation uses  $H = 90$  days.

**Uncertainty functional  $\mathcal{U}_t$ .** The t-RSI denominator is general: any auditable functional of the projected net-improvement distribution will do. The operational default is standard deviation under an end-to-end cluster-bootstrap of the channel-row fits (sensors and actuators bootstrapped cluster-by- $N$ ; the R&D derivative  $\partial\text{Sharpe}/\partial\log_{10}(1+n)$  uses an SE derived from its selected-frontier

bootstrap interval). Closed-form Pearson cross-checks use LOO  $R^2$  on the channel-row derivatives; the headline is the more conservative of the two propagation flavours.

**Sign convention.** A positive t-RSI means net alpha creation outpaces alpha decay over  $H$  by that many standard errors of the firm’s posterior dispersion. The certificate of monotone improvement (Equation 71) thresholds this same statistic at  $\zeta = \delta/\mathcal{U}_t$  channel-by-channel; commits below threshold are rejected by the controller.

**From channel rows to the create-rate posterior.** The channel-row fits in the §5 body are written as objective gradients  $g_t^k = \partial J_t / \partial a_t^k$ . To talk about alpha created per dollar invested we use the *channel alpha gradient*,

$$\psi_t^k := \frac{\partial \alpha_t^{\text{cyc}}}{\partial a_t^k} = \frac{g_t^k}{\partial J_t / \partial \alpha_t^{\text{cyc}}}. \quad (46)$$

For a planning horizon  $H$  measured in corporate cycles, the finite-horizon alpha created by channel  $k$  is the path integral of this gradient along the planned allocation path,

$$\Delta \alpha_{t:H}^{\text{create},k} \approx \int_{a_t^k}^{a_t^k + \Delta a_{t:H}^k} \psi_t^k(a) da. \quad (47)$$

Summing creation across channels gives the firm’s posterior alpha-creation rate over the horizon used as the numerator of Equation 20,

$$\widehat{\Delta \alpha^{\text{create}}}_{t:H} = \sum_k \widehat{\Delta \alpha^{\text{create},k}}_{t:H}. \quad (48)$$

The matching posterior alpha-decay rate  $\widehat{\Delta \alpha^{\text{decay}}}_{t:H}$  is estimated separately from the firm’s forecast-evaluation panel and the Mark II/III live-trading history; the full panel construction lives in subsection E.2.

**SE propagation.** Each standard error appearing in Equation 20 is the SE of a posterior mean. The channel-fit bootstrap propagates input parameters through  $\text{Normal}(\hat{\theta}, \text{SE}(\hat{\theta}))$  noise, so the standard deviation of the resulting bootstrap sample is itself the SE of the posterior mean; no further  $1/\sqrt{n}$  rescaling is applied. The construction of Equation 20 is structurally identical to a two-sample  $t$ -statistic; the operational walk-through lives in subsection E.2.

## E.2 Three-Month t-RSI Calculation

This appendix is the operational walk-through of the headline three-month t-RSI reported in section 5; it documents the numerical inputs and the bootstrap-posterior moments that compose the standardized distance.

**Numerator.** The difference of the posterior mean alpha-creation rate and the posterior mean alpha-decay rate over the  $H = 90$ -day horizon,  $\widehat{\Delta \alpha^{\text{create}}}_{t:H} - \widehat{\Delta \alpha^{\text{decay}}}_{t:H}$ , with the create-side composed of:

1. **Sensors + actuators.** The local Sharpe-data slope of Equation 23 gives Sharpe gain per decade of effective dollar-weighted tokens absorbed; the sensors row supplies the rate at which decades are absorbed (the headline calculation uses  $\beta = 1$  decade per two months, i.e. 1.5 decades over the three-month horizon). Sensors enter through the actuator-panel slope rather than as a separate create-side term because the panel already measures realized performance along the joint asset-universe and data-universe expansion path.
2. **R&D / architecture search.** The R&D contribution is the local increment of the fitted selected-frontier experiments-performance law (Equation 24) between the current campaign count  $n_{\text{current}}^{\text{exp}}$  and the projected end-of-horizon count  $n_{\text{current}}^{\text{exp}} + \Delta n$ . The increment is  $\frac{\partial \text{Sharpe}}{\partial \log_{10}(1+n)} [\log_{10}(1 + n_{\text{current}}^{\text{exp}} + \Delta n) - \log_{10}(1 + n_{\text{current}}^{\text{exp}})]$ , evaluated on the report-selected top-10% Sharpe frontier so the create-side reads in the same Sharpe units as the sensors+actuators leg. The horizon experiment count is  $\Delta n = \rho^{\text{exp}} n_{\text{researchers}} H$  with  $\rho^{\text{exp}}$  the firm's current experiments-per-researcher-day throughput from the auto-research campaign (the dollar-to-experiment conversion of Equation 61 blends human-researcher and LLM/agent throughput at their per-arm productivity rates  $\rho_t^{\text{human}}$  and  $\rho_t^{\text{LLM}}$ ; the optimizer chooses the mix endogenously). The derivative is reported *gross of researcher carry*: compensation is an accounting flow on the balance sheet, not a deduction from the alpha-creation rate.

**Decay term.** The headline path uses the empirical  $\hat{\lambda}^{\text{decay}}$  from a per-asset alpha-decay estimator: an exponential decay rate of held-out forecast edge against deployment age is fitted once per asset, and the resulting per-asset rate distribution is aggregated to a single portfolio-level rate with a robust median + MAD/IQR summary. The reported SE combines within-cell MAD-derived dispersion with the between-cell SD across training-run seeds and forecast horizons, taking the more conservative of the two so the denominator does not understate dependence between cells that share assets. The bootstrap samples  $\lambda \sim \mathcal{N}(\hat{\lambda}, \text{SE}(\hat{\lambda}))$ , maps each draw through  $1 - e^{-\lambda H_{\text{cycles}}}$  to obtain a horizon Sharpe-loss draw, and reports the resulting mean and SE. The measured  $\hat{\kappa}$  sits near zero across every linear, rank, and distributional trend statistic against deployment age, and a majority of held-out assets show no monotonic edge decay, so the data-derived decay term contributes a small mean with a tight SE rather than an aggressive Sharpe-loss point estimate.

**Denominator.** The standard error of the numerator,  $\sqrt{\text{SE}^2(\widehat{\Delta\alpha^{\text{create}}}_{t:H}) + \text{SE}^2(\widehat{\Delta\alpha^{\text{decay}}}_{t:H})}$ , propagated from end-to-end cluster-bootstrap of the channel-row fits: the actuator slope is resampled from a Gaussian with the published SE; the R&D derivative uses an SE derived from the selected Sharpe-frontier bootstrap interval; the decay draws are sampled as described above. For each draw the create and decay legs are summed independently and the empirical SD across  $N_{\text{boot}} = 2000$  draws is the reported SE of each posterior mean. Because each input parameter enters the bootstrap through parameter noise, the resulting SD is itself the SE of the posterior mean – no further  $1/\sqrt{n}$  rescaling is required.

**Audit trail.** The numerical inputs the calculation consumes are visible in the figure title written alongside the create / decay posterior chart; the readiness audit cross-references the headline t-RSI row against the per-channel Fisher-information thresholds the certificate gates against.

**Capacity sensitivity.** The headline conditions on the firm's current operating point, where the realized per-trade  $Q/\text{ADV}$  sits below the impact floor and the empirical impact contribution is

indistinguishable from zero. As AUM grows that condition cannot hold indefinitely: the literature square-root impact law [13, 64] implies a per-horizon Sharpe drag that scales with the size of executed trades relative to ADV. To make the implied capacity-ceiling reading auditable we evaluate the same data-derived headline with the create distribution reduced by a per-sample draw of the literature impact deduction at  $K$ -times current AUM, under two turnover-rolloff assumptions:  $\text{turnover}(K) = \text{turnover}(1) \cdot K^{-\alpha}$  with  $\alpha = 0$  (the worst case, in which annual portfolio turnover is held at its current  $\sim 27\times$  level as AUM grows) and  $\alpha = 0.35$  (an industry-norm rolloff, which matches the empirical decreasing-returns-to-scale pattern documented for active managers in 18, 27, 60 and brings annual turnover to  $\sim 5\times$  by  $K = 100\times$ ). The implied horizon Sharpe drag scales as  $K^{1/2-\alpha}$ , so the worst case grows as  $\sqrt{K}$  and the industry-norm grows as  $K^{0.15}$ .

| Turnover trajectory                          | $K = 1\times$ | $K = 10\times$ | $K = 100\times$ |
|--|---------------|----------------|-----------------|
| Worst case ( $\alpha = 0$ , turnover frozen) | 6.10          | 0.93           | -2.01           |
| Industry-norm rolloff ( $\alpha = 0.35$ )    | 6.05          | 4.59           | 2.90            |

*Each cell is the headline  $t$ -RSI standardized distance under the same data-derived  $\widehat{\lambda}^{\text{decay}}$  as the headline, with the create distribution reduced by the literature  $\sqrt{Q/ADV}$  Sharpe drag at the indicated AUM scale and turnover trajectory. The  $K = 1\times$  column reports the literature counterfactual at current AUM (“what the literature law would predict if orders could not be split below the impact floor”); the firm’s realized  $t$ -RSI at the same operating point is 9.61 because orders are split below the floor and the empirical impact is sub-floor. The worst-case row crosses zero between  $K = 10\times$  and  $K = 20\times$  (the closed-form crossover is at  $K \approx 17\times$ ) and is decisively negative at  $K = 100\times$ ; the industry-norm row compresses with  $K$  but remains positive across the full  $100x$  range. Realistic capacity headroom sits between these two rows, anchored on the industry-norm trajectory.*

**Audit caveats.** The two  $t$ -RSI numbers a future reader will measure depend on how fast the create-side rows themselves move (sensors, actuators, parameters, R&D) and on operational choices the firm makes as AUM grows (universe expansion, execution-horizon extension, refit cadence). The headline reported here is the present standardized distance; the capacity-sensitivity table converts the same posterior into a forward-looking range under explicit, conservative-to-realistic turnover assumptions, but the cells will be re-measured against future data.

## F Channel Derivations

### F.1 Investments Supporting Equations

The investment row’s auxiliary object is the learned execution-friction surface  $\widehat{\phi}_t$ . It evaluates the per-trade friction as a function of the candidate trade  $\Delta \mathbf{I}_t$ , the actuator surface  $\mathbf{U}_t$ , and the market state  $\mathbf{E}_t$ . The overbraces below mark which input is which.

**Backtest scope and friction surface.** The execution-friction surface  $\phi$  that enters (21) via  $\widehat{\phi}_t$  is, by construction, a *function of the trade*: half-spread, square-root impact, fees, financing, and adversarial response are all functions of  $(\Delta \mathbf{I}_t, \mathbf{U}_t, \mathbf{E}_t)$  paid at the moment of execution. This is what distinguishes  $\phi$  from the *time-rate* alpha-decay term  $\alpha^{\text{decay}}$ , which is a function of  $(\Theta_t, \mathbf{E}_t, t)$  and

lives outside the market function [13, 15, 30]. The two boxes are orthogonal in their dependencies, which is what makes them clean named primitives.

Realized cash on a single cycle is gross value minus frictions:

$$Y_t^{\$} = \sum_i (p_i \cdot q_i) - \Phi(\mathbf{q}), \quad (49)$$

where the sum runs over fills  $i$  in the cycle ( $q_i$  units at realized price  $p_i$ ) and  $\Phi(\mathbf{q})$  is total execution friction.  $\Phi$  decomposes into four measurable components [4, 13, 26, 35, 44, 46, 59]:

1. **Market impact** ( $\phi_{\text{impact}}$ ). The dominant friction. The firm’s own orders move the price against it. A common heuristic is the square-root impact law [13],  $\phi_{\text{impact}} \approx \sigma \sqrt{q/V_{\text{daily}}}$ , where  $\sigma$  is daily volatility and  $V_{\text{daily}}$  is average daily volume. Decades of empirical support across asset classes and market regimes [4, 44]. The true impact function is more complex than any closed form suggests: payment for order flow, internalization, and venue-specific rebate structures mean that for some instruments and order sizes the firm receives *price improvement* (the opposite of the square-root penalty); impact also varies with time of day, volatility regime, and the firm’s own historical order patterns.
2. **Exchange, clearing, and regulatory fees** ( $\phi_{\text{fees}}$ ). Proportional to volume. Exchange fees, clearing fees, SEC fee, TAF. Small relative to impact and near-deterministic; typically fractions of a basis point per share.
3. **Financing and borrowing costs** ( $\phi_{\text{financing}}$ ). Cost of shorting (borrow fees for locating shares), margin interest on leveraged positions, cost of carrying overnight. Scales with position size and holding period. For intraday strategies with no overnight exposure, this component is near zero.
4. **Adversarial costs** ( $\phi_{\text{adversarial}}$ ). At larger scales, the firm’s trading patterns become detectable by other market participants. Sophisticated competitors can front-run predictable order flow, engage in quote-stuffing, or trigger stop-losses—increasing realized impact beyond what the mechanical friction model predicts. At the firm’s current trading scale, this term is negligible; it becomes material at higher AUM. Formally, adversarial costs introduce a multi-agent dimension to the market function: the firm must model how other participants model *it*, and how their responses to its detected patterns feed back into the prices it receives.

**Measurability.** Each component of  $\Phi$  can be estimated from the firm’s historical execution data with a confidence interval. Market impact has decades of empirical literature [4, 13, 44], and the firm generates new calibration data on every trade. Exchange fees are published schedules. Financing costs are contractual rates. Adversarial costs are estimated from historical decomposition of realized versus predicted impact. Every friction term in (49) is observable, estimable, and comes with a known variance.

**Backtest scope.** A public backtest can recover the  $\mu$  and  $\alpha^{\text{decay}}$  terms of (21) from a walk-forward, survivorship-free universe and a fixed cost model. It cannot recover  $\hat{\phi}_t$  at production fidelity. Any published Sharpe figure that does not separately disclose its assumed friction surface is implicitly assuming  $\phi = \phi_{\text{static}}$ , which is exactly the assumption the live trading record discriminates against.

**Proprietary surface.** The firm has executed approximately \$400M of trades; that volume is the data behind its internal  $\widehat{\phi}_t$  surface. The internal estimates are demonstrably tighter than the public square-root-law approximation above (routing, venue-specific spreads, financing, and the time-of-day component of impact all enter), but the surface itself is held proprietary for exactly the reason (21) makes plain— $\phi$  is the term the rest of the world cannot buy. The blueprint therefore treats backtest-derived numbers as *upper bounds* on the deployable trading row, with the MarkI/II/III hyperparameter ledger (Appendix H) as the dated operator-side audit trail.

$$\widehat{\phi}_t \stackrel{\text{execution friction}}{=} \widehat{\phi}_t \left( \widehat{\Delta \mathbf{I}}_t, \widehat{\mathbf{U}}_t, \widehat{\mathbf{E}}_t \right). \quad (50)$$

## Protocol summary.

| Field                | Value   |
|----------------------|---|
| Sample               | Live execution and EWM evaluation ledgers across the Mark1/2/3 deployments (~\$400M of cleared trades).   |
| Estimator            | $\widehat{\phi}_t$ and $\widehat{\mu}_t$ are estimated internally against the live execution surface; the public square-root impact baseline is documented below in the backtest scope. |
| Confidence intervals | Internal surfaces are held proprietary; backtest-derived numbers are treated as upper bounds and ~\$400M of live executions provide the deployment-side audit trail.                    |

## F.2 Sensors Supporting Equations

**What this section does.** This subsection states the full model/data/architecture scaling surface, then shows the fixed-model, fixed-architecture specialization that the sensor row actually fits. It names every coefficient that enters the fitted slice, gives the fitting protocol that produces the numbers consumed by the body, and derives the local slope identity the chain rule actually calls. It is meant to be read on its own: a reader who only opens this appendix should leave with enough to reproduce the fitted numbers and to know what each one means.

**Marginal return of the sensor channel.** The body’s sensor row consumes a single derivative off this fit: the local data-scaling slope  $\partial \mathcal{L}_{\text{pred}} / \partial \log_{10} D_{\text{eff}}$ . Differentiating the scaling law term by term gives the body identity, which says the slope at any operating point is fixed by the product of the data-scaling exponent and the reducible component of the loss. The appendix then expands that measured primitive into the controller-level chain rule:

$$g_t^{\mathbf{S}} = \frac{\partial D_{\text{eff}}}{\partial a_t^{\mathbf{S}}} \cdot \frac{\partial \mathcal{L}_{\text{pred}}}{\partial D_{\text{eff}}} \cdot \frac{\partial \mu}{\partial \mathcal{L}_{\text{pred}}} \cdot \frac{\partial J_t}{\partial \mu}.$$

The remainder of this appendix defines each factor in that product (the full scaling surface and the slice the body fits, the fitting protocol, and the reported quantities) and closes with the transition law  $\widehat{\mathcal{W}}_t^{\mathbf{S}}$  that records how the next sensor inventory is distributed after a candidate sensor allocation, conditional on the sensor history.

**The full surface and current slice.** The full scaling surface used by the code is the Kaplan/Hoffmann/Chinchilla model-size/data law with the Muennighoff repeated-data correction folded into  $D_{\text{eff}}$ . In that surface, model size  $M_t$  and effective data  $D_{\text{eff}}$  are scale axes, while architecture quality  $A_t$  moves the loss floor and scaling coefficients. The current multi-seed data-scaling sweep holds  $M_t = M_0$  and  $A_t = A_0$  fixed, so the model-size term and architecture dependence collapse into a one-dimensional fitted slice. The body reports that slice, not a completed cross-model-size scaling law.

**The fitted sensor law.** On the fixed slice, the sensor row models per-asset predictive loss  $\mathcal{L}_{\text{pred}}$  as a power law in an effective dollar-weighted-token axis  $D_{\text{eff}}$ , plus a fixed-slice residual floor:

- $\mathcal{L}_{\text{noise}}$  is the fixed- $(M_0, A_0)$  residual floor: the part of loss not reduced by buying more data while  $M_t$  and  $A_t$  are held fixed. It equals  $\mathcal{L}_{\infty}(A_0)$  plus the fixed-model architectural residual at the current  $(M_0, A_0)$ , and true Bayes risk is its lower bound (reached only in the joint limit as  $M_t$  and  $A_t$  are simultaneously optimized). R&D moves  $A_t$  and therefore moves this floor; the body’s  $\mathcal{L}_{\text{noise}}$  row is read at the operating  $A_t$ , not at  $A^*$ .
- $A_{D_{\text{eff}}}$  is the prefactor on the reducible part, in the same loss units as  $\mathcal{L}_{\text{pred}}$ .
- $\alpha_{D_{\text{eff}}}$  is the power-law exponent: each decade of  $D_{\text{eff}}$  multiplies the reducible part by  $10^{-\alpha_{D_{\text{eff}}}}$ .
- $D_{\text{eff}}$  is the Muennighoff effective-data axis [56]: fresh dollar-weighted tokens  $U_D$  plus repeated tokens deflated by an epoch-saturation constant  $R^*$ .
- $U_D$  is the count of fresh dollar-weighted tokens (single-pass dollar volume seen by training). The unit is the dollar-weighted bar [5, 19, 48].
- $E$  inside the  $D_{\text{eff}}$  form is the number of training epochs actually run;  $R^*$  controls how quickly repeated passes deflate.

The equations below state the full surface, the fixed-slice law, and the  $D_{\text{eff}}$  construction [37, 40].

**Fitting protocol.** The four coefficients  $(\alpha_{D_{\text{eff}}}, A_{D_{\text{eff}}}, \mathcal{L}_{\text{noise}}, R^*)$  are recovered by a hierarchical Bayesian posterior over the 45 training-run panel, sampled with NUTS (4 chains  $\times$  2000 draws after warmup; max  $\hat{R} = 1$ , min ESS-bulk = 1956.0). The likelihood is Gaussian on log-loss residuals:  $\log_{10} \mathcal{L}_{\text{obs}} \sim \mathcal{N}(\log_{10} \mathcal{L}_{\text{pred}}, \sigma_y)$ . Two of the four coefficients are placed under informative priors that fold in external measurements:

- $R^* \sim \mathcal{N}(4.28, 1)$  epochs, with the prior center and width fixed by an independent epoch-saturation sweep.
- $\mathcal{L}_{\text{noise}} \sim \mathcal{N}(0.034, 0.01)$ , with the prior derived from a separate noise-floor decomposition that estimates the fixed- $(M_0, A_0)$  residual floor from cross-asset residuals at the current operating architecture. True Bayes risk is a sub-component (lower bound) of this estimate, not the estimate itself.

The pin is not a stylistic choice:  $\mathcal{L}_{\text{noise}}$  and  $R^*$  are jointly only weakly identifiable from the 45-row  $(D_{\text{eff}}, \mathcal{L})$  panel by itself, so an unconstrained fit lands on whichever corner of the joint surface minimises a particular residual without measuring an underlying quantity. The two informative priors carry the external information that does identify them. The remaining coefficients  $\alpha_{D_{\text{eff}}}$  and  $\log_{10} A_{D_{\text{eff}}}$  are placed under weakly-informative priors ( $\alpha_{D_{\text{eff}}} \sim \text{HalfNormal}(0.5)$ ,  $\log_{10} A_{D_{\text{eff}}} \sim \mathcal{N}(0, 2)$ ) and identified from the data.

**Reported quantities.** The §5.3 fitted-parameters table consumes the marginal posterior summaries directly: every value is the posterior mean of the corresponding marginal, every 95% CI is the marginal HPDI. The  $A_{D_{\text{eff}}}$  row is exponentiated from the  $\log_{10} A_{D_{\text{eff}}}$  marginal so the row reads in raw loss units; the HPDI bounds are exponentiated through the same transform. The headline numbers at the current operating point are  $\alpha_{D_{\text{eff}}} = 0.156$  (HPDI [0.09, 0.229]),  $A_{D_{\text{eff}}} = 3.266$  (HPDI [0.8072, 16.87]),  $\mathcal{L}_{\text{noise}} = 0.042$  (HPDI [0.024, 0.061]), and  $R^* = 4.306$  epochs (HPDI [2.324, 6.231], narrower than the prior because the data does pin  $R^*$  once  $\mathcal{L}_{\text{noise}}$  is anchored). The  $R^2$  column in the body table is a Bayes- $R^2$  in raw-loss space, evaluated at the posterior means against the same 45 training runs, currently 0.7619. The 95% band on the §5.3 chart is the posterior-predictive envelope across thinned draws from the same posterior.

**Transition law.** The transition law that follows the chain-rule statement above is the sensor-channel fragment of the learned EWM,  $\widehat{\mathcal{W}}_t^{\text{S}}$ : it records how the next sensor inventory is distributed after a candidate sensor allocation, conditional on the sensor history.

$$\mathcal{L}_{\text{pred}} = \underbrace{\mathcal{L}_{\infty}(\Gamma_t)}_{\text{architecture-set floor}} + \underbrace{A_M(\Gamma_t) M_t^{-\alpha_M(\Gamma_t)}}_{\text{model-size term}} + \underbrace{A_{D_{\text{eff}}}(\Gamma_t) D_{\text{eff}}^{-\alpha_{D_{\text{eff}}}(\Gamma_t)}}_{\text{Muennighoff effective-data term}}. \quad (51)$$

$$\mathcal{L}_{\text{pred}}|_{M_t=M_0, \Gamma_t=\Gamma_0} = \underbrace{\mathcal{L}_{\text{noise}}}_{\text{fixed-slice floor}} + \underbrace{A_{D_{\text{eff}}} D_{\text{eff}}^{-\alpha_{D_{\text{eff}}}}}_{\text{fitted data-scaling term}}. \quad (52)$$

$$\mathcal{L}_{\text{pred}} = \underbrace{A_{D_{\text{eff}}} D_{\text{eff}}^{-\alpha_{D_{\text{eff}}}}}_{\text{reducible loss (decays as more effective data is bought)}} + \underbrace{\mathcal{L}_{\text{noise}}}_{\text{fixed-slice residual floor}}. \quad (53)$$

$$D_{\text{eff}} = \underbrace{U_{D_{\$}}}_{\text{fresh dollar-weighted tokens}} + \underbrace{U_{D_{\$}} R^* \left(1 - e^{-(E-1)/R^*}\right)}_{\text{repeated tokens, deflated by saturation } R^*}. \quad (54)$$

**Definition 34** (Sensor marginal return). The **sensor marginal return**  $g_t^{\text{S}}$  is the optimizer-level expansion of the local data-scaling slope used in the body. It chains sensor dollars  $\rightarrow$  effective data  $D_{\text{eff}} \rightarrow$  predictive loss  $\mathcal{L}_{\text{pred}} \rightarrow$  expected return  $\mu \rightarrow$  objective  $J_t$ ; only the  $D_{\text{eff}} \rightarrow \mathcal{L}_{\text{pred}}$  factor is measured in the sensors section itself.

$$g_t^{\text{S}} = \frac{\partial}{\partial a_t^{\text{S}}} D_{\text{eff}} \frac{\partial}{\partial \mu} J_t \frac{\partial}{\partial D_{\text{eff}}} \mathcal{L}_{\text{pred}} \frac{\partial}{\partial \mathcal{L}_{\text{pred}}} \mu \quad (55)$$

**Example.** Suppose the firm spends  $a_t^{\text{S}} = \$50\text{K}$  on an options-flow archive that adds  $\Delta U_{D_{\$}}$  fresh dollar-weighted tokens.  $\partial D_{\text{eff}}/\partial a_t^{\text{S}}$  converts dollars into the Muennighoff axis, the fitted  $\partial \mathcal{L}_{\text{pred}}/\partial D_{\text{eff}}$  moves predictive loss down along the multi-seed scaling slope of Equation 22,  $\partial \mu/\partial \mathcal{L}_{\text{pred}}$  translates loss into expected return through the actuator panel, and  $\partial J_t/\partial \mu$  closes onto the cumulative objective. Only the middle factor is measured in this section; the two outer factors enter from the sensors-to-data conversion and the actuator slope respectively.

**Definition 35** (Sensor channel transition law). The **sensor channel transition law** is the sensor-channel fragment of the learned EWM,  $\widehat{\mathcal{W}}_t^{\text{S}}$ . It returns the distribution over the next-cycle sensor

inventory  $\mathbf{S}_{t+1}$  given the sensor history  $\mathcal{H}_t^{\mathbf{S}}$  and the sensor allocation  $a_t^{\mathbf{S}}$  the controller commits this cycle. The fitted data-scaling surface of Equation 51 is the empirical content of this transition: it says how much effective dollar-weighted data  $a_t^{\mathbf{S}}$  buys and how that increment moves predictive loss.

$$\overbrace{\mathbf{S}_{t+1}}^{\text{next-cycle sensor inventory}} = \widehat{\mathcal{W}}_t^{\mathbf{S}}(\cdot | \mathcal{H}_t^{\mathbf{S}}, a_t^{\mathbf{S}}) \quad (\text{channel EWM transition law}). \quad (56)$$

**Example.** Suppose the firm commits  $a_t^{\mathbf{S}} = \$50\text{K}$  to extend its options-flow archive by three years. Conditioning on the sensor history of past purchases,  $\widehat{\mathcal{W}}_t^{\mathbf{S}}$  returns a distribution over  $\mathbf{S}_{t+1}$  whose dollar-weighted-token mass on options-flow rises by the implied  $\Delta U_{D_{\mathfrak{s}}}$ , with the predictive-loss reduction on the next refit drawn from the fitted Muennighoff surface.

### Protocol summary.

| Field                | Value  |
|----------------------|--|
| Sample               | multi-seed scaling sweep: 15 universe sizes<br>(1, 2, 4, 8, 12, 16, 24, 32, 64, 128, 256, 384, 512, 640, 863 assets) $\times$ 3 seeds = 45 runs.           |
| Target               | Per-asset predictive loss $\mathcal{L}_{\text{pred}}$ on the held-out evaluation universe.   |
| Effective-data axis  | Muennighoff $D_{\text{eff}}$ , with $U_{D_{\mathfrak{s}}}$ the single-pass dollar-weighted-token volume.   |
| Estimator            | Hierarchical Bayesian posterior sampled with NUTS (4 chains $\times$ 2000 post-warmup draws); Gaussian likelihood on $\log_{10}$ -loss residuals.          |
| Priors               | $R^* \sim \mathcal{N}(4.28, 1)$ , $\mathcal{L}_{\text{noise}} \sim \mathcal{N}(0.034, 0.01)$ ; the remaining coefficients carry weakly-informative priors. |
| Confidence intervals | 95% HPDI on each marginal posterior; $A_{D_{\text{eff}}}$ is exponentiated through the $\log_{10}$ transform.  |

### F.3 Actuators Supporting Equations

**What this section does.** This subsection explains the local data-performance slopes plotted in §5.4: realized annualized return and realized annualized Sharpe per decade of effective dollar-weighted data. In the general ontology, actuators are the interfaces through which the controller can act. In this empirical trading-envelope instance, the actuator surface is narrower: the tradable asset universe. Because the asset universe and the dollar-weighted data universe expand together in this sweep, the body reports the direct panel slopes and leaves the joint-path interpretation here.

**Marginal return of the actuator channel.** The body’s actuator row consumes the two local data-performance slopes plotted in §5.4. The controller’s general actuator row is  $g_t^{\mathbf{U}}$  (Equation 58): actuator dollars move the capability surface  $\mathbf{U}_t$ , that surface changes the learned execution-friction surface  $\widehat{\phi}_t(\Delta \mathbf{I}_t, \mathbf{U}_t, \mathbf{E}_t)$  inside the investment production function, and the changed friction surface changes realized return. The current chart is one trading-envelope instance of that row, where  $\mathbf{U}_t$  is the tradable asset universe and not a general claim about LLM APIs, robotics, or human labor. The remainder of this appendix defines the path along which  $D_{\text{eff}}$  and  $\mathbf{U}_t$  co-move, gives the OLS fitting protocol that produces the two body slopes, and closes with the capability-surface transition law  $\widehat{\mathcal{W}}_t^{\mathbf{U}}$ .

**The local data-performance slopes.** Let  $N$  index the tradable universe size. The experiment moves along the path

$$D_{\text{eff}} = D_{\text{eff}}(N), \quad \mathbf{U}_t = \mathbf{U}_t^{\text{trade}}(N),$$

so the body slopes are total slopes along that path:

$$\frac{d}{d \log_{10} D_{\text{eff}}(N)} \mu_t^{\text{ann}} \left( D_{\text{eff}}(N), \mathbf{U}_t^{\text{trade}}(N) \right), \quad \frac{d}{d \log_{10} D_{\text{eff}}(N)} \text{Sharpe}_t \left( D_{\text{eff}}(N), \mathbf{U}_t^{\text{trade}}(N) \right).$$

They can be interpreted as a data-to-loss term plus a tradable-surface term, but the body does not separately identify those two effects:

$$\frac{dY_t}{d \log_{10} D_{\text{eff}}} = \frac{\partial Y_t}{\partial \mathcal{L}_{\text{pred}}} \frac{\partial \mathcal{L}_{\text{pred}}}{\partial \log_{10} D_{\text{eff}}} + \frac{\partial Y_t}{\partial \mathbf{U}_t^{\text{trade}}} \frac{d\mathbf{U}_t^{\text{trade}}}{d \log_{10} D_{\text{eff}}}, \quad Y_t \in \{\mu_t^{\text{ann}}, \text{Sharpe}_t\}.$$

The intermediate loss and realized-performance ranges are much narrower than the data range, so the directly measured panel slope is the stable body primitive.

**Fitting protocol.** The two slopes (annualized return and annualized Sharpe) are recovered by ordinary least squares of the cluster-median realized metric on  $\log_{10} D_{\text{eff}}$ :

- **Underlying panel.** The asset-universe sweep contains 15 universe sizes  $\times$  3 seeds = 45 raw runs. After joining the effective-data coordinate  $D_{\text{eff}} = U_D + U_D R^*(1 - e^{-(E-1)/R^*})$  to realized annualized return and Sharpe, the executable-coverage restriction leaves a 36-row ( $N$ , seed) panel.
- **Cluster medians.** For each universe size  $N$  we take the median across the three seeds of (annualized return, Sharpe). We drop the three smallest universe sizes for which the per-seed median has insufficient evaluation breadth, leaving  $n = 12$  cluster-median rows. The body's parameter table reports this  $n = 12$  as the regression's degrees-of-freedom unit.
- **OLS.** The slope and intercept come from a linear fit on  $(\log_{10} D_{\text{eff}}^{(N)}, y^{(N)})$  across the 12 medians, separately for  $y = \text{ann. return}$  and  $y = \text{Sharpe}$ . The standard errors use the usual OLS residual variance  $\sigma^2 = \text{SSE}/(n-2)$  and prediction variance  $\text{SE}_{\hat{y}(x)}^2 = \sigma^2(1 + 1/n + (\log_{10} x - \overline{\log_{10} x})^2/S_{xx})$ .
- **Cluster-by- $N$  bootstrap.** The 36-row ( $N$ , seed) panel is the object the cluster-by- $N$  bootstrap of [subsection E.2](#) resamples for the t-RSI numerator's seed-noise leg.

**Why OLS, not Bayesian.** The sensor row uses a hierarchical Bayesian posterior because  $\mathcal{L}_{\text{noise}}$  and  $R^*$  are jointly only weakly identifiable from the  $(D_{\text{eff}}, \mathcal{L})$  panel and require informative priors derived from independent measurements (see [subsection F.2](#)). The actuator row is a different problem: a single linear fit through 12 cluster medians where the slope and intercept are jointly identified by elementary OLS algebra. There is no analogous degeneracy and no external information that needs to be folded in via priors, so OLS prediction intervals are well-calibrated and we use them directly.

**Chart construction.** The §5.4 chart shows, for each of the two metrics:

- Per- $N$  cluster median (filled dot,  $n = 12$ ) and seed min/max as vertical error bars.
- OLS  $\log$ - $x$  fit line  $y = \hat{\beta} \log_{10} D_{\text{eff}} + \hat{\alpha}$ , with  $R^2$  printed in the legend.
- Parametric 95% prediction-interval band (filled region) on the same  $\sigma^2(1 + 1/n + \dots)$  formula above.

- $T_A$  and  $T_B$  vertical dashed marks at  $D_{\text{eff}}(T_A) = D_{\text{op}} \cdot 10^{2.31}$  and  $D_{\text{eff}}(T_B) = D_{\text{op}} \cdot 10^{3.71}$ , where  $D_{\text{op}} = \max_N D_{\text{eff}}^{(N)}$  is the operating point of the current sweep. The headroom dex match the sensor headline chart in §5.3 so the two charts share an  $x$ -axis interpretation.
- Extrapolated  $T_A, T_B$  markers with vertical 95% PI error bars: these are the same  $\hat{y}(T_A)$  and  $\hat{y}(T_B)$  values reported in the §5.4 fitted-parameters table.

The chart has no  $R^2$  annotation per tier marker because  $R^2$  is a property of the regression that produced the slope, not of any single derived point. The body table reports  $R^2$  on the slope rows and leaves the  $T_A/T_B$  extrapolation rows blank in the  $R^2$  column for the same reason.

**Reported quantities.** The §5.4 fitted-parameters table consumes the OLS bookkeeping directly. The slope rows give the local return-data and Sharpe-data slopes plus 95% CIs from  $\hat{\beta} \pm 1.96 \cdot \text{SE}_{\hat{\beta}}$  and the regression  $R^2$ . The four extrapolation rows —  $\text{AnnRet}(T_A)$ ,  $\text{AnnRet}(T_B)$ ,  $\text{Sharpe}(T_A)$ ,  $\text{Sharpe}(T_B)$  — give the OLS prediction mean and the parametric 95% prediction interval evaluated at  $\log_{10} D_{\text{eff}}(T_A)$  and  $\log_{10} D_{\text{eff}}(T_B)$ .

**Capability surface and transition.** In this section’s local trading instance, the actuator state is the tradable asset universe (Equation 57). The transition law  $\widehat{\mathcal{W}}_t^{\text{U}}$  (Equation 59) describes how that surface evolves with actuator spend and the environment.

$$\overbrace{\mathbf{U}_t}^{\text{trading-envelope actuator surface}} = \overbrace{\mathbf{U}_t^{\text{trade}}(N)}^{\text{tradable asset universe}} \quad (57)$$

**Definition 36** (Actuator marginal return). The **actuator marginal return**  $g_t^{\text{U}}$  chains actuator dollars  $\rightarrow$  capability surface  $\mathbf{U}_t \rightarrow$  learned friction  $\hat{\phi}_t \rightarrow$  realized return. Buying an actuator is buying capacity in the friction surface that converts the controller’s intended trade into realized cash.

$$g_t^{\text{U}} = \frac{\partial}{\partial a_t^{\text{U}}} \mathbf{U}_t \frac{\partial}{\partial \hat{\phi}_t} \left[ \frac{\Delta \mathbf{I}_t \hat{\boldsymbol{\mu}}_t}{a_t^{\text{I}}} - \hat{\phi}_t \right] \frac{\partial}{\partial \mathbf{U}_t} \hat{\phi}_t(\Delta \mathbf{I}_t, \mathbf{U}_t, \mathbf{E}_t) \quad (58)$$

**Example.** Suppose AlphaFund adds an IEX D-Limit route that cuts the venue-selection contribution to  $\hat{\phi}_t$  by 0.4 bps per dollar traded. On a \$120M deployed book turning over  $\sim 30\%$  per cycle, that is  $\$120\text{M} \times 0.30 \times 0.00004 \approx \$1,440$  per cycle of recovered alpha;  $g_t^{\text{U}}$  prices this against the cycle cost of the connection.

**Definition 37** (Actuator channel transition law). The **actuator channel transition law**  $\widehat{\mathcal{W}}_t^{\text{U}}$  returns the distribution over the next-cycle actuator surface  $\mathbf{U}_{t+1}$  given the actuator history  $\mathcal{H}_t^{\text{U}}$  and the actuator allocation  $a_t^{\text{U}}$ . In the trading-envelope instance the actuator surface is the tradable asset universe of Equation 57, and the empirical content of this transition is the OLS panel slope in Equation 23: how much realized Sharpe a decade of additional dollar-weighted tokens buys when the universe widens along the firm’s deployment path.

$$\overbrace{\mathbf{U}_{t+1}}^{\text{next-cycle capability surface}} = \widehat{\mathcal{W}}_t^{\text{U}}(\cdot | \mathcal{H}_t^{\text{U}}, a_t^{\text{U}}) \quad (\text{channel EWM transition law}). \quad (59)$$

**Example.** Suppose the firm commits  $a_t^U = \$120\text{K}$  to widen its tradable universe from 127 to 160 assets (new venue access, additional brokerage agreements, the engineering work to support the larger asset universe). Conditioning on the actuator history,  $\widehat{\mathcal{W}}_t^U$  returns a distribution over  $\mathbf{U}_{t+1}$  that contains the 33 additional names, and the OLS panel slope of Equation 23 prices the implied gain in deployed Sharpe per decade of dollar-weighted tokens absorbed across the widened universe.

### Protocol summary.

| Field                | Value  |
|----------------------|--|
| Sample               | Cluster-median rows ( $n = 12$ ) over the 45-run multi-seed sweep; the cluster-by- $N$ bootstrap of subsection E.2 is the matched uncertainty model. |
| Target               | Realized annualized return and annualized Sharpe per cluster.  |
| Estimator            | OLS of the target against $\log_{10} D_{\text{eff}}$ .   |
| Confidence intervals | $\pm 1.96$ SE (Gaussian); $T_A$ and $T_B$ extrapolation rows read directly off the same OLS.   |

## F.4 R&D Supporting Equations

**What this section does.** This subsection states the empirical R&D primitive plotted in §5.5 — the selected rolling upper-tail held-out frontier of the auto-research campaign as a function of completed experiments — and gives enough fitting protocol that a reader who only opens this appendix can reproduce the reported numbers and locate the derivatives inside the controller chain rule. The structural stack underneath the body derivative (dollars  $\rightarrow$  experiments  $\rightarrow$  architecture-quality scalar  $\Gamma_t \rightarrow$  coefficient vector  $\boldsymbol{\eta}(\Gamma_t) \rightarrow$  predictive loss) is included for completeness; the body keeps only the two measured derivatives.

**Marginal return of the R&D channel.** The body’s R&D row consumes a single derivative off each selected frontier fit: the local derivative of the report-selected top-10% held-out Sharpe frontier with respect to log-experiments,  $\partial \text{Sharpe} / \partial \log_{10}(1+n)$  at the current  $n_{\text{current}}^{\text{exp}}$ , with the analogous top-10% derivative for annualized return. The appendix then expands that measured primitive into the controller-level chain rule of Equation 65: dollars  $\rightarrow$  experiments through the dollar-to-experiment production function, experiments  $\rightarrow$  architecture quality through the search-scaling law, architecture quality  $\rightarrow$  scaling-surface coefficients through the architecture-coefficient map, scaling coefficients  $\rightarrow$  predictive loss through the joint surface, and predictive loss  $\rightarrow$  expected return through the actuator slope. The R&D row is therefore the only channel-row whose body derivative and structural derivative are not the same object: the body derivative is a faithful downstream readout of the firm’s actually-optimized metrics, and the structural derivative on  $\Gamma_t$  remains queued behind the per-architecture coefficient sweep referenced below. The remainder of this appendix defines each factor in that chain rule (the full stack, the fitted law, task-based automation, fitting protocol, chart construction, reported quantities) and closes with the R&D EWM transition law  $\widehat{\mathcal{W}}_t^Z$ .

**The full stack and the slice the body uses.** R&D dollars decompose into human-researcher and LLM/agent dollar streams (Equation 60). The dollar-to-experiment production function (Equation 61) converts those streams into completed-experiment count via per-arm productivity rates

$\rho_t^{\text{human}}$  and  $\rho_t^{\text{LLM}}$  (experiments per dollar). The search-scaling law (Equation 63) maps completed experiments to the dimensionless architecture-quality index  $\Gamma_t$ , and the architecture-coefficient map (Equation 64) records which coefficients of the joint Kaplan/Hoffmann/Chinchilla/Muennighoff scaling surface are moved by  $\Gamma_t$ . The transition law  $\widehat{\mathcal{W}}_t^{\mathbf{Z}}$  records how the R&D state evolves between cycles. The body’s §5.5 derivatives  $\partial\text{Sharpe}/\partial\log_{10}(1+n)$  and  $\partial\text{AnnReturn}/\partial\log_{10}(1+n)$  are downstream-projection derivatives against the metrics the firm optimizes against (validation Sharpe and validation annualized return), not structural derivatives against  $\Gamma_t$ ; the structural slope  $\xi^\Gamma$  requires a per-architecture coefficient sweep that fits  $\boldsymbol{\eta}(\Gamma)$  jointly across architectures and is the natural extension of the multi-seed sensor sweep along the  $\Gamma$  axis.

**The fitted R&D law.** On the strict-invalid- and sealed-holdout-filtered cohort, rolling upper-tail frontiers of validation Sharpe and validation annualized return are each modelled as a logarithmic experiments-performance law in the completed-experiment index  $n$ :

- $n$  is the experiment index after sorting the filtered cohort by run identifier; the upper-tail frontier is the rolling mean of the best observed scores so far at the selected cutoff.
- $\beta_{0,\text{Sharpe}}$  and  $\beta_{0,\text{AnnReturn}}$  are the level intercepts at  $n = 0$  in Sharpe and percentage points respectively.
- $\partial\text{Sharpe}/\partial\log_{10}(1+n)$  and  $\partial\text{AnnReturn}/\partial\log_{10}(1+n)$  are the derivatives the body table reports, in Sharpe per unit of  $\log_{10}(1+n)$  and percentage points per unit of  $\log_{10}(1+n)$  respectively. At the operating cohort size ( $n \gg 1$ ) one unit of  $\log_{10}(1+n)$  is, to within  $\log_{10}(1+1/n)$ , a ten-fold increase in completed experiments, so the derivative reads naturally as “per decade of experiments”; we use that shorthand in the body table and figure labels, with the exact units stated here.
- $n_{\text{current}}^{\text{exp}}$  is the cohort size after filtering — the count of experiments the firm has actually completed and the warm-start point for the local marginal calculation.

The fitted body equation (Equation 24) stacks the two laws on the same axis.

**Task-based automation and the productivity ratio.** The task-based automation term (Equation 62) defines  $f_t^{\text{auto}} = \rho_t^{\text{LLM}}/\rho_t^{\text{human}}$  as the integral over the research task set  $\mathcal{T}$  of per-task automation feasibility  $p_{\text{auto}}(\sigma, t)$  weighted by human-researcher cost  $w_{\text{human}}(\sigma)$ , normalized by total human-research cost. The functional form mirrors the task-automation literature [1, 2, 8], in which capital substitutes for human labor on a measurable subset of the task distribution rather than uniformly across it. Empirically the ratio has two external anchors: its *level* at time  $t$  comes from the Anthropic Economic Index / AIoE task-exposure dataset cross-walked through ONET against the firm’s own task list, and its *time-derivative* is anchored by the METR autonomous-task-horizon doubling time  $\tau_{2\times} \approx 7$  months [52]. Holding  $\rho_t^{\text{human}}$  slowly varying over the relevant horizon,  $\rho_t^{\text{LLM}}$  grows on the same schedule, and the cost-equivalent LLM share of completed experiments rises proportionately. The optimizer chooses the human/LLM mix endogenously through the budget constraint  $a_t^{\mathbf{Z}} = a_t^{\mathbf{Z},\text{human}} + a_t^{\mathbf{Z},\text{LLM}}$ ; the firm currently sits near the LLM endpoint of the substitution path because that is where  $\rho_t^{\text{LLM}} a_t^{\mathbf{Z},\text{LLM}}$  is largest at the present operating budget, and the 929-experiment headline campaign is one realized draw of that endpoint.

**Team-size diminishing returns.** Pure-human research throughput does not scale linearly in headcount: communication overhead grows with team size  $N_{\mathbf{H}}$  and erodes the marginal productivity

of each additional hire, the Brooks coordination-drag effect [16]. The clean way to absorb this is to let  $\rho_t^{\text{human}}$  depend on  $N_{\mathbf{H}}$  rather than introducing a separate parametric form for the human arm. The firm has not yet run a controlled headcount-vs-output study on its own R&D history, so this dependency remains a structural placeholder; the corresponding parameters are not reported as fitted values in the body table.

**Fitting protocol.** The filtered cohort starts from the 929-experiment auto-research clean table and applies two measurement-quality exclusions: the curated removal flag (rows the campaign audit marked as buggy, contaminated-holdout, look-ahead, or carrying an invalid-metric marker) and a single-evaluator implausibility flag (validation Calmar  $> 2.05$ , validation Sortino  $> 2.50$ , or validation Sharpe  $> 1.90$  on the surviving rows); sealed-holdout rows are also dropped by name. Both exclusions are on *measurement quality*, not on outcome: the dropped runs are samples from the measurement-failure distribution, not the search-process distribution. Within the resulting cohort, rows are sorted by run identifier, the metric of interest is restricted to its finite-valued subset, and the experiment index  $n = 1, 2, \dots$  is assigned along that sorted sequence. Candidate frontiers are the rolling means of the best top-5%, top-10%, and top-20% observed scores so far. Each candidate frontier is fit against  $\log_{10}(1 + n)$  by ordinary least squares and scored by rolling-origin prediction on held-out suffixes; the whitepaper manifest pins the body frontier to top-10% for both metrics, and the appendix reports the full cutoff scan. The same protocol is applied to validation annualized return, with the clean column rescaled by 100 inside the fit so the derivative reads in percentage points.

**Chart construction.** The §5.5 evidence is split across two single-panel figures with a shared  $x$ -axis on completed experiments  $n$ : one for held-out Sharpe and one for held-out annualized return (in percentage points). Each figure carries the raw per-experiment scatter, the SPY baseline, the muted running-best step as context, the selected rolling upper-tail frontier, the bootstrap band around that selected frontier, and the fitted log law. The two fitted curves and the two derivative rows in the body table come from the same selected-frontier fits. The held-out trajectory is the canonical 1258-day validation window each run reports against the 2516-day training period; the same window applies to every chart in §5.5.

**Reported quantities.** The §5.5 fitted-parameters table consumes the selected frontier fits directly. The two intercept rows  $\beta_{0,\text{Sharpe}}$  and  $\beta_{0,\text{AnnReturn}}$  give the level of the selected frontier curve at  $n = 0$ ; the two derivative rows give the per-decade gain on that frontier. The 95% confidence intervals are bootstrap percentile intervals from resampling experiments with replacement, recomputing the selected frontier, and refitting the log law; the  $R^2$  column reports the regression’s fit to the selected upper-tail curve and the  $n$  column reports the cohort size after filtering.

**Transition law.** The R&D EWM transition  $\widehat{\mathcal{W}}_t^Z$  (Equation 66) records how the next R&D capability state is distributed after a candidate allocation, conditional on the R&D history. The empirical anchor is the same auto-research campaign: at the deployed human/LLM mix, each cycle’s allocation generates a draw from the campaign’s per-experiment performance distribution, and the selected upper-tail frontier summarizes the improving opportunity set across cycles. The body derivatives summarize how that selected frontier has evolved with  $n$  on the observed campaign.

$$\underbrace{\text{R\&D allocation (dollars)}}_{a_t^{\mathbf{Z}}} = \underbrace{\text{human researchers}}_{a_t^{\mathbf{Z},\text{human}}} + \underbrace{\text{LLM/agent research}}_{a_t^{\mathbf{Z},\text{LLM}}}. \quad (60)$$

$$\underbrace{\text{experiments completed this cycle}}_{N_t^{\text{exp}}} = \underbrace{\text{human-arm throughput}}_{\rho_t^{\text{human}} a_t^{\mathbf{Z},\text{human}}} + \underbrace{\text{LLM-arm throughput}}_{\rho_t^{\text{LLM}} a_t^{\mathbf{Z},\text{LLM}}}. \quad (61)$$

$$\underbrace{\text{LLM/human productivity ratio}}_{f_t^{\text{auto}}} \rho_t^{\text{LLM}} / \rho_t^{\text{human}} = \frac{\overbrace{\int_{\mathcal{T}} p_{\text{auto}}(\sigma, t) w_{\text{human}}(\sigma) d\sigma}^{\text{automatable human-research cost}}}{\underbrace{\int_{\mathcal{T}} w_{\text{human}}(\sigma) d\sigma}_{\text{total human-research cost}}}. \quad (62)$$

$$\underbrace{\text{architecture quality}}_{\Gamma_t} = \xi \log_{10} \left( 1 + \frac{\overbrace{\text{completed experiments / warm-start offset}}_{N_t^{\text{exp}}}}{N_0} \right). \quad (63)$$

$$\boldsymbol{\eta}(\Gamma_t) = \begin{bmatrix} \mathcal{L}_{\infty}(\Gamma_t) \\ A_M(\Gamma_t) \\ \alpha_M(\Gamma_t) \\ A_{D_{\text{eff}}}(\Gamma_t) \\ \alpha_{D_{\text{eff}}}(\Gamma_t) \end{bmatrix}. \quad (64)$$

**Definition 38** (R&D marginal return). The **R&D marginal return**  $g_t^{\mathbf{Z}}$  chains R&D dollars  $\rightarrow$  completed experiments  $N_t^{\text{exp}}$  (human plus LLM/agent throughput at per-arm productivity rates  $\rho_t^{\text{human}}, \rho_t^{\text{LLM}}$ )  $\rightarrow$  architecture quality  $\Gamma_t \rightarrow$  Hoffmann/Chinchilla coefficient vector  $\boldsymbol{\eta}(\Gamma_t) \rightarrow$  predictive loss  $\rightarrow$  expected return. R&D acts on the *shape* of the scaling surface itself, not on its current operating point.

$$g_t^{\mathbf{Z}} = \frac{\partial J_t}{\partial \mu} \frac{\partial \mu}{\partial \mathcal{L}_{\text{pred}}} \frac{\partial \mathcal{L}_{\text{pred}}}{\partial \boldsymbol{\eta}(\Gamma_t)} \cdot \frac{\partial \boldsymbol{\eta}(\Gamma_t)}{\partial \Gamma_t} \frac{\partial}{\partial a_t^{\mathbf{Z}}} \left[ \xi \log_{10} \left( 1 + \frac{\rho_t^{\text{human}} a_t^{\mathbf{Z},\text{human}} + \rho_t^{\text{LLM}} a_t^{\mathbf{Z},\text{LLM}}}{N_0} \right) \right]. \quad (65)$$

**Example.** AlphaFund’s 929-experiment auto-research campaign, restricted to the strict-invalid- and sealed-holdout-filtered cohort, fits selected rolling upper-tail frontiers for validation Sharpe and validation annualized return. The body uses the report-selected top-10% frontier for both metrics. Suppose the firm completes ten times as many experiments at its current automation rate: the selected validation-Sharpe frontier is projected to gain about  $\partial \text{Sharpe} / \partial \log_{10}(1+n)$  Sharpe, and the selected annualized-return frontier is projected to gain about  $\partial \text{AnnReturn} / \partial \log_{10}(1+n)$  percentage points, evaluated as the local increment of the fitted curve from the current campaign count. These derivatives then propagate through the architecture-coefficient map  $\boldsymbol{\eta}(\Gamma_t)$  to the loss surface and the expected return inside the trajectory chain rule (Appendix F).

**Definition 39** (R&D channel transition law). The **R&D channel transition law**  $\widehat{\mathcal{W}}_t^{\mathbf{Z}}$  returns the distribution over the next-cycle R&D capability state  $\mathbf{Z}_{t+1}$  given the R&D history  $\mathcal{H}_t^{\mathbf{Z}}$  and the R&D allocation  $a_t^{\mathbf{Z}}$ . The empirical content is the auto-research campaign of Equation 24: at the deployed human/LLM mix, each cycle’s allocation generates a draw from the campaign’s per-experiment performance distribution, and the selected upper-tail frontier summarizes how the improving opportunity set evolves with  $n$ .

$$\underbrace{\mathbf{Z}_{t+1}}_{\text{next-cycle R\&D capability}} = \widehat{\mathcal{W}}_t^{\mathbf{Z}}(\cdot \mid \mathcal{H}_t^{\mathbf{Z}}, a_t^{\mathbf{Z}}) \quad (\text{channel EWM transition law}). \quad (66)$$

**Example.** Suppose the controller commits  $a_t^{\mathbf{Z}} = \$400\text{K}$  to a quarter of R&D, split \$100K human and \$300K LLM/agent under the current  $\rho_t^{\text{LLM}}/\rho_t^{\text{human}}$  ratio. Conditioning on the auto-research history,  $\widehat{\mathcal{W}}_t^{\mathbf{Z}}$  returns a distribution over the next R&D state in which the realized experiment count  $N_t^{\text{exp}}$  is drawn from Equation 61 and the top-10% Sharpe frontier shifts by the local  $\partial\text{Sharpe}/\partial\log_{10}(1+n)$  derivative of Equation 24.

### Protocol summary.

| Field                | Value   |
|----------------------|---|
| Sample               | 929-experiment auto-research campaign, restricted to the strict-invalid- and sealed-holdout-filtered cohort, sorted by completed-experiment index $n$ .     |
| Frontier             | Rolling upper-tail means of best top-5%, 10%, 20% experiments so far on validation Sharpe and validation annualized return.                                 |
| Headline cutoff      | Top-10% on both metrics (whitepaper-pinned); the appendix reports the full cutoff sensitivity scan.   |
| Estimator            | OLS of $\widehat{y}(n) = \beta_0 + b \log_{10}(1+n)$ on the selected frontier; rolling-origin prediction on held-out suffixes scores each candidate cutoff. |
| Rescaling            | Validation annualized return is rescaled by 100 inside the fit so the derivative reads in percentage points per decade.                                     |
| Confidence intervals | Bootstrap-percentile intervals from resampling experiments with replacement, re-computing the selected frontier, and refitting the log law.                 |

**R&D frontier sensitivity.** The whitepaper headline frontier is selected by the report manifest; the surrounding sensitivity scan still reports candidate rolling means of the best top-5%, top-10%, and top-20% experiments observed so far.

### Validation Sharpe cutoff scan.

| cutoff         | endpoint     | 95% CI                | slope         | 95% CI                  | OOS RMSE                  |
|----------------|--------------|-----------------------|---------------|-------------------------|---------------------------|
| top-5%         | 1.395        | [1.359, 1.418]        | 0.3732        | [0.329, 0.4583]         | 0.0722 ± 0.006915         |
| <b>top-10%</b> | <b>1.349</b> | <b>[1.311, 1.381]</b> | <b>0.3436</b> | <b>[0.2978, 0.4153]</b> | <b>0.08679 ± 0.007909</b> |
| top-20%        | 1.253        | [1.201, 1.296]        | 0.3026        | [0.2582, 0.3682]        | 0.07677 ± 0.004741        |

### Validation Sharpe frontier-law comparison at selected top-10%.

| model          | OOS RMSE                               | OOS MAE                                | description   |
|----------------|--|--|---|
| constant       | $0.2352 \pm 0.004469$                  | $0.2241 \pm 0.007061$                  | constant frontier level   |
| linear_n       | $0.1603 \pm 0.06605$                   | $0.1521 \pm 0.06302$                   | linear in experiment count n                                      |
| log_n          | $0.08679 \pm 0.007909$                 | $0.07674 \pm 0.01004$                  | linear in $\log_{10}(1+n)$  |
| <b>sqrt_n</b>  | <b><math>0.0733 \pm 0.01663</math></b> | <b><math>0.06716 \pm 0.0164</math></b> | <b>linear in <math>\sqrt{n}</math></b>                            |
| saturating_log | $0.09383 \pm 0.007582$                 | $0.08103 \pm 0.01063$                  | saturating exponential in $\log_{10}(1+n)$ , tau selected by grid |

### Validation annualized return cutoff scan.

| cutoff         | endpoint     | 95% CI               | slope        | 95% CI                | OOS RMSE                               |
|----------------|--------------|----------------------|--------------|-----------------------|--|
| top-5%         | 24.99        | [23.81, 26.19]       | 2.562        | [1.565, 5.858]        | $1.23 \pm 0.3229$                      |
| <b>top-10%</b> | <b>23.56</b> | <b>[22.6, 24.38]</b> | <b>2.135</b> | <b>[1.202, 4.792]</b> | <b><math>0.7561 \pm 0.08734</math></b> |
| top-20%        | 21.71        | [20.81, 22.45]       | 2.076        | [0.807, 4.223]        | $0.5202 \pm 0.04591$                   |

### Validation annualized return frontier-law comparison at selected top-10%.

| model          | OOS RMSE                               | OOS MAE                                | description   |
|----------------|--|--|---|
| constant       | $1.195 \pm 0.06485$                    | $1.017 \pm 0.08467$                    | constant frontier level   |
| linear_n       | $1.95 \pm 0.7502$                      | $1.857 \pm 0.7185$                     | linear in experiment count n                                      |
| <b>log_n</b>   | <b><math>0.7561 \pm 0.08734</math></b> | <b><math>0.6662 \pm 0.08544</math></b> | <b>linear in <math>\log_{10}(1+n)</math></b>                      |
| sqrt_n         | $1.236 \pm 0.3579$                     | $1.172 \pm 0.3551$                     | linear in $\sqrt{n}$  |
| saturating_log | $0.7905 \pm 0.06002$                   | $0.6738 \pm 0.06094$                   | saturating exponential in $\log_{10}(1+n)$ , tau selected by grid |

## F.5 Parameters Supporting Equations

The parameters channel rests on the full Kaplan/Hoffmann/Chinchilla/Muennighoff loss surface [37, 40, 56]: an architecture-set loss floor, a model-size term, and an effective-data term whose data axis already deflates repeated epochs. The current sensor fit fixes  $M_t = M_0$  and  $A_t = A_0$ ; the model-size sweep needed to identify  $\alpha_M(A_t)$  and  $A_M(A_t)$  remains a future fit. The compute-cost identity (training FLOPs equals tokens times passes times model size, divided by training efficiency) states which  $(M_t, D_t, K_t)$  choices are feasible, and the transition law states how weights move between cycles.

$$\mathcal{L}_{\text{pred}} = \underbrace{\mathcal{L}_{\infty}(\Gamma_t)}_{\text{architecture-set noise floor}} + \underbrace{A_M(\Gamma_t) M_t^{-\alpha_M(\Gamma_t)}}_{\text{model-size term}} + \underbrace{A_{D_{\text{eff}}}(\Gamma_t) D_{\text{eff}}^{-\alpha_{D_{\text{eff}}}(\Gamma_t)}}_{\text{effective-data term}}. \quad (67)$$

$$\underbrace{C_t^{\text{train}}}_{\text{training compute (FLOPs)}} = \frac{\underbrace{D_t^{\text{seen}}}_{\text{tokens seen}} \underbrace{K_t^{\text{pass}}}_{\text{passes}} \underbrace{M_t}_{\text{model size}}}{\underbrace{\eta_t^{\text{train}}}_{\text{training efficiency}}}. \quad (68)$$

**Definition 40** (Parameters channel transition law). The **parameters channel transition law**  $\widehat{\mathcal{W}}_t^{\Theta}$  returns the distribution over the next-cycle weight vector  $\Theta_{t+1}$  given the parameter history  $\mathcal{H}_t^{\Theta}$

and the parameter allocation  $a_t^\Theta$  (training compute purchased this cycle). The empirical content is the joint Kaplan/Hoffmann/Chinchilla/Muennighoff scaling surface of Equation 67: the next weight set is the optimizer’s output after consuming  $a_t^\Theta$  worth of compute on the current  $D_{\text{eff}}$  corpus, with its held-out predictive loss drawn from the fitted surface.

$$\underbrace{\Theta_{t+1}}_{\text{next-cycle parameters}} = \widehat{\mathcal{W}}_t^\Theta(\cdot | \mathcal{H}_t^\Theta, a_t^\Theta) \quad (\text{channel EWM transition law}). \quad (69)$$

**Example.** Suppose the firm commits  $a_t^\Theta = \$80\text{K}$  to one production refit on the current 863-asset corpus:  $\sim \$80\text{K}$  of H200 hours produces a new  $\Theta_{t+1}$  at model size  $M_0$ . Conditioning on the parameter history,  $\widehat{\mathcal{W}}_t^\Theta$  returns a distribution over  $\Theta_{t+1}$  whose realized predictive loss is drawn from the fitted surface at the current  $(M_0, D_{\text{eff}}, A_0)$  operating point.

## Protocol summary.

| Field                | Value   |
|----------------------|---|
| Sample               | Cross-architecture $(M, D)$ sweep at fixed architecture quality has not yet been run; the parameter row is recorded symbolically. |
| Planned estimator    | Chinchilla-style $(M, D)$ grid to pin $\alpha_M(\Gamma_t)$ and $A_M(\Gamma_t)$ at fixed $\Gamma_t = \Gamma_0$ .                   |
| Confidence intervals | Rows render as <span style="background-color: #e0e0e0;">pending data</span> until the sweep delivers the M-axis coefficients.     |

## F.6 Continual Learning Supporting Equations

The continual-learning channel is the bridge between the static parameters channel (§ 5.6) and the live deployed model. The empirical content is the epoch-multiplier intersection chart in § 5.7: where the production  $L_{\text{Kaplan}}^{K^*=3}$  frontier and the single-pass  $L_{\text{Kaplan}}^{K=1}$  frontier cross, the effective number of optimizer passes through the current-bar corpus that maximises validation loss reduction is identified. Per-cycle continual-learning alpha is the share of  $\Delta\alpha^{\text{R\&D}}$  and  $\Delta\alpha^{\text{data+act}}$  that the refit cadence converts into deployed Sharpe at the intersection epoch; in the headline t-RSI calculation of subsection E.2 this row is folded into sensors+actuators rather than separately identified, and a future revision will split it out once a multi-cadence retraining sweep delivers the refit-cost coefficient.

**Definition 41** (Continual-learning intersection). The **continual-learning intersection** is the dollar-weighted-token scale  $x_\cap$  at which the single-pass ( $K = 1$ ) loss curve meets the best-of- $K^*$ -epoch loss curve, with both curves fit as Hoffmann-style power laws on  $D$  at fixed  $M_t^{\text{int}}$  and shared loss floor  $L_\infty$ . Below  $x_\cap$  the best-epoch curve is strictly tighter (additional passes over a small token budget reduce loss); at  $x_\cap$  the two curves cross, and to the right of it  $\text{epoch}^{\text{best}} \text{ loss} \rightarrow 1$ : a single pass over the data is enough for the model to converge, so repeated epochs no longer help and most of the available history can be reserved for genuine evaluation rather than training. The two power-law exponents  $\alpha_1$  and  $\alpha_{K^*}$  are independent and separately estimable; setting  $L^{K=1}(x_\cap) = L^{K=K^*}(x_\cap)$  and solving for  $D$  gives  $x_\cap = (A_{K^*}/A_1)^{1/(\alpha_1 - \alpha_{K^*})}$ .

$$\begin{aligned} & \mathcal{L}_\infty + A_M(\Gamma_t) (M_t^{\text{int}})^{-\alpha_M(\Gamma_t)} + A_{D_{\text{eff}}}(\Gamma_t) (D_t^{\text{seen}})^{-\alpha_{D_{\text{eff}}}(\Gamma_t)} \\ & = \mathcal{L}_\infty + A_{D_{\text{eff}}}(\Gamma_t) (D_t^{\text{seen}} K_t^*)^{-\alpha_{D_{\text{eff}}}(\Gamma_t)} + A_M(\Gamma_t) (M_t^{\text{int}})^{-\alpha_M(\Gamma_t)} \end{aligned} \quad (70)$$

**Example.** Suppose the corrected multi-seed fit puts the intersection at  $x_\cap \approx 3.2 \times 10^{14}$  DWT and  $L_\cap \approx 0.40$  test combined loss, with  $\alpha_1 \approx 0.085$  and  $\alpha_{K=3}^* \approx 0.028$ . Once AlphaFund crosses this scale, the scaling-sweep architecture is in its prequential regime: train on the most recent slice, test on the rest—a holdout pattern unavailable to the multi-epoch regime.

## Protocol summary.

| Field                | Value  |
|----------------------|--|
| Sample               | Same 45-run multi-seed grid as sensors.  |
| Target               | First-epoch test combined loss and $K = 3$ best-epoch test combined loss versus the single-pass dollar-weighted-token axis $U_{D_s}$ . |
| Estimator            | Log-log OLS power-law fits to recover the slopes $\alpha_1$ and $\alpha_{K=3}^*$ ; closed-form intersection $(x_\cap, L_\cap)$ .       |
| Confidence intervals | $\alpha_1$ and $\alpha_{K=3}^*$ : OLS standard errors. $(x_\cap, L_\cap)$ : 1000-sample cluster-bootstrap by $N$ (2.5/97.5 quantiles). |

## G Data Scaling

### G.1 Direct Versus Chained Fit

The body proves the realized-portfolio map two ways. The chained estimator composes the data-scaling fit Equation 56 with the loss-to-edge linearization measured in the actuator panel: data dollars buy effective-token volume  $D_{\text{eff}}$ ,  $D_{\text{eff}}$  buys reducible predictive loss, and reducible loss buys realized edge  $\mu$ . The direct estimator skips the data-side regression and fits  $\mu$  as a function of  $L_{\text{pred}}$  on the same 45-run sweep (the loss-to-edge slope  $b$  measured directly from the realized portfolio data; cf. Figure 5). The chain composes only if the two paths recover one another within their joint uncertainty.

On the current 45-run multi-seed scaling sweep, a linear-in- $\log_{10} L_{\text{pred}}$  regression on the per- $(N, \text{seed})$  points yields

$$\begin{aligned} \frac{\partial \text{AnnRet}}{\partial \log_{10} L_{\text{pred}}} &= 6.4425 \pm 1.3023 \quad (R_{\text{MAD}}^2 = 0.7592), \\ \frac{\partial \text{Sharpe}}{\partial \log_{10} L_{\text{pred}}} &= 0.3972 \pm 0.0446 \quad (R_{\text{MAD}}^2 = 0.9264); \end{aligned}$$

each decade of reducible predictive loss converts to  $|6.4425|$  percentage points of annualized return and  $|0.3972|$  units of Sharpe. The chained estimator recovers the same slopes within one standard error on this dataset: pushed to the cross-asset loss endpoint  $L_T^B = 0.03634$  (95% PI  $[8.17233 \times 10^{-6}, 0.03645]$ ) from Equation 56, the loss-to-edge linearization projects approximately 44.2364% annualized return ( $[28.5026, 59.9702]$ ) and Sharpe approximately 2.5881 ( $[2.0491, 3.1272]$ ); these agree within one standard error with the chain-composed counterpart. Both extrapolations are upper bounds: alpha decay and execution friction at frontier dollar-volume scales will bend the realized curve below the in-sample slope.

## G.2 Selection-Rule Efficient Frontier

A natural worry is that aggregating per-asset MAE over the entire training universe washes out a signal concentrated in the model’s high-confidence predictions, and that a top- $K$  selection rule would recover a tighter scaling exponent. We test this empirically. For each

$$K \in \{5, 10, 20, 50, 100, 200, 500, \infty\}$$

(where  $\infty$  is no selection) we restrict the per-asset MAE to the top- $K$  symbols per run on the headline channel (ranked by predicted-magnitude proxy  $|f_t|$ ), aggregate to an  $n_{\text{eval}}$ -weighted MAE per run, and re-fit Equation 56 on the canonical  $D_{\text{eff}}$  axis. Figure 6 reports  $\hat{\alpha}(K)$  on the left and the per- $K$  loss-vs- $D_{\text{eff}}$  curves on the right. Aggressive top- $K$  selection erases the data-scaling signal: at  $K \leq 50$  the fitted exponent is small or negative because tightening the asset set primarily reduces the universe-size variation that drives the regression. The body fit therefore uses no selection, which is also the only choice that aligns with how the deployed allocator consumes the predictions.

## G.3 Robustness: Jackknife, Cook’s Distance, Leverage

The headline Hoffmann/Muennighoff fit Equation 56 is run on all 45 per- $(N, \text{seed})$  points across the 15 universe sizes. Two sensitivity checks bracket the headline number. The leave-one-universe-out jackknife refits the same form 45 times, dropping one point at a time, and yields an alpha range of

$$[0.0751, 0.28237]$$

around the headline  $\hat{\alpha} = 0.07517$  – a width comparable to one standard error and well inside the confidence the body needs to commit reinvestment on the sensor channel. On the log-log linear analogue used to construct the 95% prediction interval the smallest universes carry the highest leverage, as expected for an axis with approximately 3.5 OOM in-sample range, and the jackknife range above already shows that no single point dominates the fit. The 95% prediction interval at the cross-asset endpoint

$$D_{\text{eff}} = 4.046 \times 10^{17}$$

is

$$[8.17233 \times 10^{-6}, 0.03645]$$

around the central estimate 0.03634.

## G.4 What the Chart Says and Extrapolation Robustness

To read the headline chart of §5.3 (see [Appendix G](#)) correctly: the x-axis is the volume of *effective dollar-weighted tokens* the model has trained on, not the count of distinct predictive factors or signals engineered into the feature set. The takeaway is not “each new factor adds this much edge”; it is that buying more dollar-weighted bars—raw measurement volume of the same kind the model already consumes—reduces predictive loss at a measured power-law rate, and that realistic procurement and sampling paths scale that volume by one to several orders of magnitude beyond the current operating point. The model/architecture surface that the fixed-slice fit specializes is given in [subsection F.2](#); the direct-versus-chained, selection-rule, and jackknife/Cook’s-distance checks in the preceding subsections of this appendix are the robustness backbone for that extrapolation.

If  $\alpha_{D_{\text{eff}}}$  holds across the next decade of effective dollar-weighted tokens—a modest extrapolation against the operating point, with the realistic procurement and sampling factors cataloged here—then the reducible part of predictive loss falls by about 30% (HPDI roughly [19%, 41%], propagated from the  $\alpha_{D_{\text{eff}}}$  HPDI [0.09, 0.229]), and the same multiplier compounds approximately decade-on-decade so long as the slope persists and the operating point stays inside the fitted regime. Power-law data-scaling of this Kaplan/Hoffmann form is one of the most robust empirical regularities in modern machine learning, having now been reproduced across many orders of magnitude in language [37, 40, 56], images [80], video and other generative modalities [36], mixed-modal models [3], geospatial foundation models [39], time-series foundation models [6], and protein biology [47]; recovering the same shape on the dollar-weighted-bar axis is consistent with that body of evidence rather than an exotic claim.

## H Deployment Parameters

This appendix collects the per-generation hyperparameters across the three deployment generations. Mark I and Mark II predate the current ASIC ledger; their numerical hyperparameters are reconstructed from operator memory and the deployment diary, not from a re-runnable training manifest. The Mark II realized PnL, Sortino, and Sharpe are taken from the audited Luz Capital reproduction; the corresponding live AlphaFund-vs-SPY breakdown for Mark III is forwarded from the formal paper’s Appendix J and the live-trading reliability/capacity audit lives in the formal paper’s Appendix K.

| Quantity             | Mark I                                 | Mark II                           | Mark III                   |
|----------------------|--|-----------------------------------|----------------------------|
| Model class          | long-only US equities                  | long/short US equities            | long/short US equities     |
| Model size           | ~10M params                            | ~25M params                       | ~56M params                |
| Training hardware    | 1xconsumer GPU                         | 4xA100                            | 8XH100                     |
| Bar grain            | daily                                  | ~30-minute                        | ~1-hour                    |
| Universe             | ~100 liquid US single-names            | ~400 US single-names              | ~500–800 US single-names   |
| Cycle horizon        | end-of-day rebalance                   | intraday + overnight              | intraday + overnight       |
| Window               | ~8 months                              | 10.5 months                       | end-Oct 2025 – May 2026    |
| Realized PnL         | underperformed                         | +34.5%                            | live (formal paper App. J) |
| Sortino              | < 1.0                                  | 3.06                              | live                       |
| Sharpe               | —                                      | 2.44                              | live                       |
| Beta to SPY          | —                                      | ~0.4                              | live                       |
| Turnover (per cycle) | ~3x                                    | ~10x                              | live                       |
| Cause of transition  | retired (bid–ask spread mis-modelling) | regime-shift detection (Aug 2024) | live (operator-managed)    |

**Table 5:** Deployment hyperparameters across Mark I, Mark II, and Mark III. Pre-Mark III numbers are reconstructed from the deployment diary; the Mark II PnL, Sortino, and Sharpe come from the audited Luz Capital reproduction (referenced by the body table). The Mark III row is the deployment under which the live trading record is currently being accumulated; the formal paper’s Appendix J carries the per-account AlphaFund-vs-SPY breakdown.

**What’s in  $\hat{\mu}_t$  and  $\hat{\phi}_t$ .** The headline edge  $\hat{\mu}_t$  in (21) is the Mark III EWM’s per-trade forecast conditional on the current book and filtration. The friction surface  $\hat{\phi}_t$  aggregates the four components catalogued in subsection F.1 — impact, exchange/clearing/regulatory fees, financing, and adversar-

ial response — into a per-trade scalar drag. Both are estimated against the ~\$400M of executed trades the Mark III deployment has cleared; the resulting numerical surfaces are proprietary, in line with the redacted rows of the investments parameter table.

## I Improvement Certificate

### I.1 Certificate of Monotone Improvement

**Definition 42** (Certificate of monotone improvement). The **certificate of monotone improvement** is the thresholded form of the held-out t-RSI statistic that gates each capital commit.  $\text{Cert}_t = 1$  when, for every active channel  $c$ , the local Fisher information of the EWM  $\mathcal{I}(c, \widehat{\mathcal{W}}_t)$  has reached its evaluator-specific readiness floor  $\varepsilon_c$  and the held-out horizon t-RSI clears the Sharpe-margin threshold  $\zeta = \delta/\mathcal{U}_t$ . The certificate fires channel-by-channel because the proper-scoring rule of Equation 38 accumulates evidence per channel history; it ratchets the corporate loop because every fired commit adds a new row that tightens both the EWM posterior and the t-RSI denominator on subsequent cycles.

$$\text{Cert}_t = \mathbf{1}\left(\forall_{c \in \text{active}} \left(\mathcal{I}(c, \widehat{\mathcal{W}}_t) \geq \varepsilon_c \wedge \text{t-RSI}(t, H) \geq \zeta\right)\right) \quad (71)$$

**Example.** Suppose a candidate Mark III refit arrives at cycle  $t$ . The sensors row’s Fisher information  $\mathcal{I}(\mathbf{S}, \widehat{\mathcal{W}}_t)$  has cleared  $\varepsilon_{\mathbf{S}}$  (the multi-seed scaling sweep is identified), and so have the actuators and parameters rows. The held-out three-month t-RSI from subsection E.2 stands at 1.45 against a margin threshold  $\zeta = 1.0$ . Both clauses hold, so  $\text{Cert}_t = 1$  and the controller commits the refit. If the refit had instead pushed t-RSI to 0.6, the second clause would fail and the controller would reject the commit.

## References

- [1] Daron Acemoglu. The simple macroeconomics of AI. *Economic Policy*, 40(121):13–58, 2025. doi: 10.1093/epolic/eiae042. Originally circulated as NBER Working Paper 32487, May 2024.
- [2] Daron Acemoglu and Pascual Restrepo. Artificial intelligence, automation and work. Working Paper 24196, National Bureau of Economic Research, January 2018.
- [3] Armen Aghajanyan, Lili Yu, Alexis Conneau, Wei-Ning Hsu, Karen Hambardzumyan, Susan Zhang, Stephen Roller, Naman Goyal, Omer Levy, and Luke Zettlemoyer. Scaling laws for generative mixed-modal language models. *arXiv preprint arXiv:2301.03728*, 2023.
- [4] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2001.
- [5] Thierry Ané and Hélyette Geman. Order flow, transaction clock, and normality of asset returns. *The Journal of Finance*, 55(5):2259–2284, 2000.
- [6] Abdul Fatir Ansari, Lorenzo Stella, Caner Turkmen, Xiyuan Zhang, Pedro Mercado, Huibin Shen, Oleksandr Shchur, Syama Sundar Rangapuram, Sebastian Pineda Arango, Shubham Kapoor, Jasper Zschiegner, Danielle C. Maddix, Hao Wang, Michael W. Mahoney, Kari Torkkola, Andrew Gordon Wilson, Michael Bohlke-Schneider, and Yuyang Wang. Chronos: Learning the language of time series. *arXiv preprint arXiv:2403.07815*, 2024.

- [7] Karl Johan Åström. Optimal control of Markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10:174–205, 1965. doi: 10.1016/0022-247X(65)90154-X.
- [8] David Autor and Anna Salomons. Is automation labor share–displacing? Productivity growth, employment, and the labor share. *Brookings Papers on Economic Activity*, 2018(1):1–87, 2018.
- [9] David H. Bailey and Marcos López de Prado. The deflated Sharpe ratio: Correcting for selection bias, backtest overfitting, and non-normality. *Journal of Portfolio Management*, 40(5):94–107, 2014.
- [10] Dimitri P. Bertsekas. Model predictive control and reinforcement learning: A unified framework based on dynamic programming. *IFAC-PapersOnLine*, 2024. Presented at the 2024 IFAC World Congress.
- [11] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Academic Press, 1978.
- [12] Fischer Black and Robert Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43, 1992. doi: 10.2469/faj.v48.n5.28.
- [13] Jean-Philippe Bouchaud, Julius Bonart, Jonathan Donier, and Martin Gould. *Trades, Quotes and Prices: Financial Markets Under the Microscope*. Cambridge University Press, Cambridge, 2018. doi: 10.1017/9781316659335.
- [14] Stephen Boyd and Lieven Vandenbergh. *Convex Optimization*. Cambridge University Press, 2004. doi: 10.1017/CBO9780511804441.
- [15] Stephen Boyd, Enzo Busseti, Steven Diamond, Ronald N. Kahn, Kwangmoo Koh, Peter Nys-trup, and Jan Speth. Multi-period trading via convex optimization. *Foundations and Trends in Optimization*, 3(1):1–76, 2017. doi: 10.1561/24000000023.
- [16] Frederick P. Brooks. *The Mythical Man-Month: Essays on Software Engineering*. Addison-Wesley, Reading, MA, 1975. Source of Brooks’s Law.
- [17] Alexei Chekhlov, Stanislav Uryasev, and Michael Zabarankin. Drawdown measure in portfolio optimization. *International Journal of Theoretical and Applied Finance*, 8(1):13–58, 2005. doi: 10.1142/S0219024905002767.
- [18] Joseph Chen, Harrison Hong, Ming Huang, and Jeffrey D. Kubik. Does fund size erode mutual fund performance? the role of liquidity and organization. *American Economic Review*, 94(5):1276–1302, 2004. doi: 10.1257/0002828043052277. Empirical evidence that mutual fund performance erodes with size, driven primarily by liquidity-related trading frictions.
- [19] Peter K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, 1973.
- [20] R. H. Coase. The nature of the firm. *Economica*, 4(16):386–405, 1937. doi: 10.1111/j.1468-0335.1937.tb00002.x.
- [21] Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991. doi: 10.1111/j.1467-9965.1991.tb00002.x.

- [22] Tom Davidson, Rose Hadshar, and William MacAskill. Three types of intelligence explosion. Technical report, Forethought Foundation, 3 2025. URL <https://www.forethought.org/research/three-types-of-intelligence-explosion>.
- [23] Marc Peter Deisenroth and Carl Edward Rasmussen. PILCO: A model-based and data-efficient approach to policy search. In *Proceedings of the 28th International Conference on Machine Learning*, pages 465–472, 2011.
- [24] Delaware Court of Chancery. ebay Domestic Holdings, Inc. v. Newmark, 16 A.3d 1, 2010.
- [25] Edward W. Felten, Manav Raj, and Robert Seamans. Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217, 2021.
- [26] Thierry Foucault, Marco Pagano, and Ailsa Röell. *Market Liquidity: Theory, Evidence, and Policy*. Oxford University Press, New York, 2013.
- [27] Andrea Frazzini, Ronen Israel, and Tobias J. Moskowitz. Trading costs. *SSRN Electronic Journal*, 2018. doi: 10.2139/ssrn.3229719. AQR analysis of \$1 trillion of live institutional trades; estimates capacity and per-trade impact as a function of fund size, and quantifies the turnover compression that large funds adopt to keep per-trade Q/ADV in the low-impact regime.
- [28] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.
- [29] Donald Goldfarb and Garud Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003. doi: 10.1287/moor.28.1.1.14260.
- [30] Richard C. Grinold and Ronald N. Kahn. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. McGraw-Hill, New York, 2 edition, 2000.
- [31] David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018.
- [32] Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. In *International Conference on Learning Representations*, 2020.
- [33] Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023.
- [34] Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023. DreamerV3.
- [35] Larry Harris. *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press, Oxford, 2003.
- [36] Tom Henighan, Jared Kaplan, Mor Katz, Mark Chen, Christopher Hesse, Jacob Jackson, Heewoo Jun, Tom B. Brown, Prafulla Dhariwal, Scott Gray, Chris Hallacy, Benjamin Mann, Alec Radford, Aditya Ramesh, Nick Ryder, Daniel M. Ziegler, John Schulman, Dario Amodei, and Sam McCandlish. Scaling laws for autoregressive generative modeling. *arXiv preprint arXiv:2010.14701*, 2020. Cross-domain scaling laws for images, video, multimodal, and math.

- [37] Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katie Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Jack W. Rae, Oriol Vinyals, and Laurent Sifre. Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*, 2022.
- [38] C. S. Holling. Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4:1–23, 1973. doi: 10.1146/annurev.es.04.110173.000245.
- [39] Johannes Jakubik, Sujit Roy, C. E. Phillips, Paolo Fraccaro, Denys Godwin, Bianca Zadrozny, Daniela Szwarzman, Carlos Gomes, Gabby Nyirjesy, Blair Edwards, Daiki Kimura, Naomi Simumba, Linsong Chu, S. Karthik Mukkavilli, Devyani Lambhate, Kamal Das, Ranjini Bangalore, Dario Oliveira, Michal Muszynski, Kumar Ankur, Muthukumaran Ramasubramanian, Iksha Gurung, Sam Khallaghi, Hanxi (Steve) Li, Michael Cecil, Maryam Ahmadi, Fatemeh Kordi, Hamed Alemohammad, Manil Maskey, Raghu Ganti, Kommy Weldemariam, and Rahul Ramachandran. Foundation models for generalist geospatial artificial intelligence. arXiv preprint arXiv:2310.18660, 2023. Prithvi; geospatial foundation-model scaling from IBM/NASA.
- [40] Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*, 2020.
- [41] Andrej Karpathy. autoresearch: AI agents running research on single-GPU training automatically, March 2026. URL <https://github.com/karpathy/autoresearch>. GitHub repository. Accessed: 2026-03-15.
- [42] J. L. Kelly. A new interpretation of information rate. *Bell System Technical Journal*, 35(4): 917–926, 1956. doi: 10.1002/j.1538-7305.1956.tb03809.x.
- [43] Daniel Kokotajlo, Eli Lifland, Thomas Larsen, and Romeo Dean. AI 2027. <https://ai-2027.com/>, April 2025. AI Futures Project scenario forecast.
- [44] Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, 1985. doi: 10.2307/1913210.
- [45] Yann LeCun. A path towards autonomous machine intelligence. *OpenReview preprint*, 2022. Version 0.9.2, 2022-06-27.
- [46] Charles-Albert Lehalle and Sophie Laruelle. *Market Microstructure in Practice*. World Scientific, 2nd edition, 2018.
- [47] Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, Allan dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Salvatore Candido, and Alexander Rives. Evolutionary-scale prediction of atomic-level protein structure. *Science*, 379(6637):1123–1130, 2023. ESM-2; documents biology-domain scaling laws across 8M–15B parameters.
- [48] Marcos López de Prado. *Advances in Financial Machine Learning*. Wiley, Hoboken, NJ, 2018.
- [49] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952. doi: 10.1111/j.1540-6261.1952.tb01525.x.

- [50] John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford University Press, 1995.
- [51] Robert C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 51(3):247–257, 1969. doi: 10.2307/1926560.
- [52] METR. Measuring autonomous AI R&D capabilities, 2025. URL <https://metr.org/blog/2025-03-19-measuring-ai-rd-capabilities/>. Accessed: 2026-04-01.
- [53] Attilio Meucci. The Black–Litterman approach: Original model and extensions. *Encyclopedia of Quantitative Finance*, 2010. doi: 10.1002/9780470061602.eqf14009.
- [54] Michigan Supreme Court. Dodge v. Ford Motor Co., 204 Mich. 459, 170 N.W. 668, 1919.
- [55] Paul Milgrom and John Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58(6):1255–1277, 1990.
- [56] Niklas Muennighoff, Alexander M. Rush, Boaz Barak, Teven Le Scao, Aleksandra Piktus, Nouamane Tazi, Sampo Pyysalo, Thomas Wolf, and Colin Raffel. Scaling data-constrained language models. *Advances in Neural Information Processing Systems*, 36, 2023. arXiv:2305.16264.
- [57] Heejeong Nam, Quentin Le Lidec, Lucas Maes, Yann LeCun, and Randall Balestriero. Causal-JEPA: Learning world models through object-level latent interventions. *arXiv preprint arXiv:2602.11389*, 2026.
- [58] Richard R. Nelson and Sidney G. Winter. *An Evolutionary Theory of Economic Change*. Harvard University Press, 1982.
- [59] Maureen O’Hara. *Market Microstructure Theory*. Blackwell Publishers, Cambridge, MA, 1995.
- [60] Ľuboš Pástor, Robert F. Stambaugh, and Lucian A. Taylor. Scale and skill in active management. *Journal of Financial Economics*, 116(1):23–45, 2015. doi: 10.1016/j.jfineco.2014.11.008. Documents decreasing returns to scale at the industry and fund level; reports the empirical relationship between fund AUM and portfolio turnover, with turnover declining as AUM grows.
- [61] Edith T. Penrose. *The Theory of the Growth of the Firm*. Basil Blackwell, Oxford, 1959.
- [62] Warren B. Powell. *Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions*. John Wiley & Sons, 2022.
- [63] Simo Särkkä. *Bayesian Filtering and Smoothing*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2013. doi: 10.1017/CBO9781139344203.
- [64] Yuki Sato and Kiyoshi Kanazawa. Universal validity of the square-root law of market impact at trader and individual-stock resolution. *arXiv preprint*, 2025. Tokyo Stock Exchange order-flow study; replicates the  $\sqrt{Q}/ADV$  market-impact law down to individual traders and individual stocks.
- [65] Jürgen Schmidhuber. Gödel machines: Fully self-referential optimal universal self-improvers. In Ben Goertzel and Cassio Pennachin, editors, *Artificial General Intelligence*, pages 199–226. Springer, 2007. doi: 10.1007/978-3-540-68677-4\_7.

- [66] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, and David Silver. Mastering Atari, Go, Chess and Shogi by planning with a learned model. *Nature*, 588:604–609, 2020. doi: 10.1038/s41586-020-03051-4.
- [67] Jaime Sevilla, Tamay Besiroglu, Ben Cottier, Josh You, Edu Roldán, Pablo Villalobos, and Ege Erdil. Can AI scaling continue through 2030?, 2024. URL <https://epoch.ai/blog/can-ai-scaling-continue-through-2030/>. Accessed: 2026-03-13.
- [68] William F. Sharpe. The arithmetic of active management. *Financial Analysts Journal*, 47(1): 7–9, 1991. doi: 10.2469/faj.v47.n1.7.
- [69] George Soros. *The Alchemy of Finance: Reading the Mind of the Market*. Simon and Schuster, New York, 1987.
- [70] Anirudh Subramanyam, Yuxin Chen, and Robert L. Grossman. Scaling laws revisited: Modeling the role of data quality in language model pretraining. *arXiv preprint arXiv:2510.03313*, 2026.
- [71] Richard S. Sutton. Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bulletin*, 2(4):160–163, 1991. doi: 10.1145/122344.122377.
- [72] Richard S. Sutton. The bitter lesson. <http://www.incompleteideas.net/IncIdeas/BitterLesson.html>, 2019.
- [73] Alex Tamkin, Kunal Handa, Avash Shrestha, Anna Goldie, Sean McCleish, Samuel R. Bowman, Deep Ganguli, and Esin Durmus. Clio: Privacy-preserving insights into real-world AI use. Anthropic Economic Index technical report; <https://www.anthropic.com/research/cli>, 2024. AEI V4 dataset release.
- [74] Edward O. Thorp. The Kelly criterion in blackjack, sports betting, and the stock market. In S. A. Zenios and W. T. Ziemba, editors, *Handbook of Asset and Liability Management, Volume 1: Theory and Methodology*, pages 385–428. North-Holland, 2006.
- [75] Pei Wang, Xiang Liu, and Quinn Dougherty. Conceptions of artificial intelligence and singularity. *Information*, 9(4):79, 2018. doi: 10.3390/info9040079. Widely cited formal treatment of recursive self-improvement definitions.
- [76] Mike West and Jeff Harrison. *Bayesian Forecasting and Dynamic Models*. Springer Series in Statistics. Springer, 2 edition, 1997. doi: 10.1007/b98971.
- [77] Roman V. Yampolskiy. From seed AI to technological singularity via recursively self-improving software. *arXiv preprint arXiv:1502.06512*, 2015.
- [78] Joshua You, David Owen, David Porter, and Tom Wilson. Scaling intelligence: The exponential growth of AI’s power needs. Technical Report 3002033669, EPRI, 2025. URL <https://www.epri.com/research/products/000000003002033669>.
- [79] Eliezer Yudkowsky. Creating friendly AI 1.0: The analysis and design of benevolent goal architectures. <https://intelligence.org/files/CFAI.pdf>, 2001. Introduces the “seed AI” concept.

- [80] Xiaohua Zhai, Alexander Kolesnikov, Neil Houlsby, and Lucas Beyer. Scaling vision transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12104–12113, 2022.