AI Fragility: Reading the Cycle in Its Own Filings

A reproducible data brief

Michael Richardson

2026-06-24

The 2026 AI-compute cycle is usually argued on valuation. The more decision-relevant question is **structural fragility** — and it is legible in the data these companies file themselves. This brief reads six independent indicators directly from the filings, then sets a market signal against a ground-truth signal in a single divergence gauge. Every figure is computed live from the source data, in-process and reproducibly; the code sits foldable beside each result.

The dataset

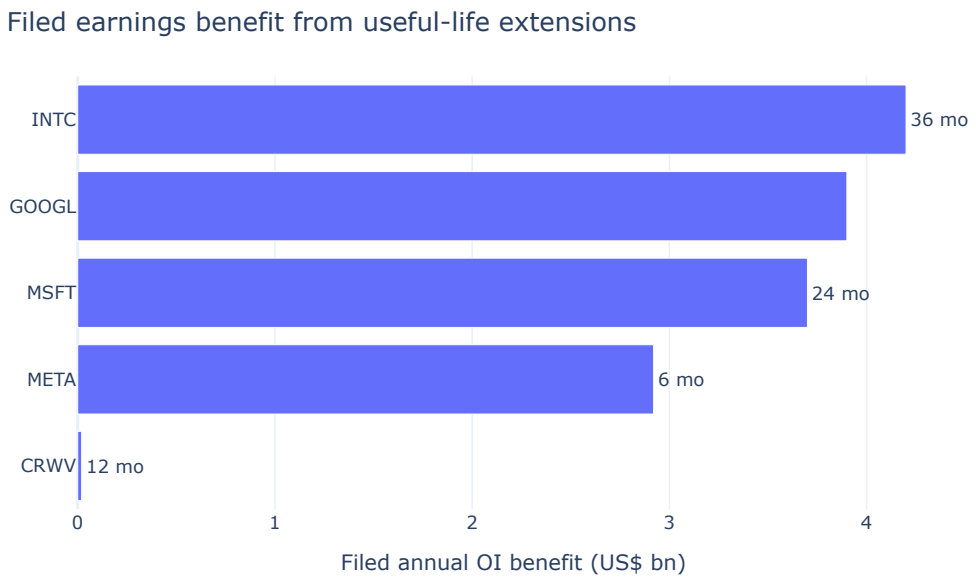
Six filing-sourced indicator tables, queried live through DuckDB:

source	rows
capex_demand.csv	11
depreciation.csv	11
financing_edges.csv	31
ground_truth.csv	4
insider.csv	22
soxx_daily.csv	257

Depreciation Integrity

The first indicator asks a narrow accounting question with a wide reach: has a firm extended the useful life of its depreciable assets, and does that extension convert into reported income without a matching dollar of cash? Where the move is real and AI-linked, the benefit (§2.1, `delta_dna = ppe_depreciable × (1/life_old − 1/life_new)`) flatters operating income on paper alone. The methodology's one hard rule sits underneath it: a life *shortened* scores zero regardless of size.

firm	move	life yrs	ext months	stated benefit $B	delta $B	D&A	annual D&A $B	score floor
AMZN	shorten	6.0 → 5.0	—	—	—		41.86	0
INTC	extend	5.0 → 8.0	36	4.2	8.094		9.951	—
GOOGL	extend	? → 6.0	—	3.9	—		21.14	—
MSFT	extend	4.0 → 6.0	24	3.7	—		22	—
META	extend	5.0 → 5.5	6	2.92	—		18	—
CRWV	extend	5.0 → 6.0	12	0.02	1.019		2.454	—
ORCL	extend	5.0 → 6.0	12	—	1.451		3.867	—



The direction of travel is uniform: every firm that touched a useful life lengthened it, and four did so while running the largest AI-capex programs on record. Amazon is the control. It moved the same lever the other way — six years to five — and absorbed a $1.4B charge against income, which is why it scores zero here despite carrying the heaviest depreciation line ($41.86B) in the set. The signal is not the size of depreciation; it is the choice to make it smaller while everyone's assets are aging faster.

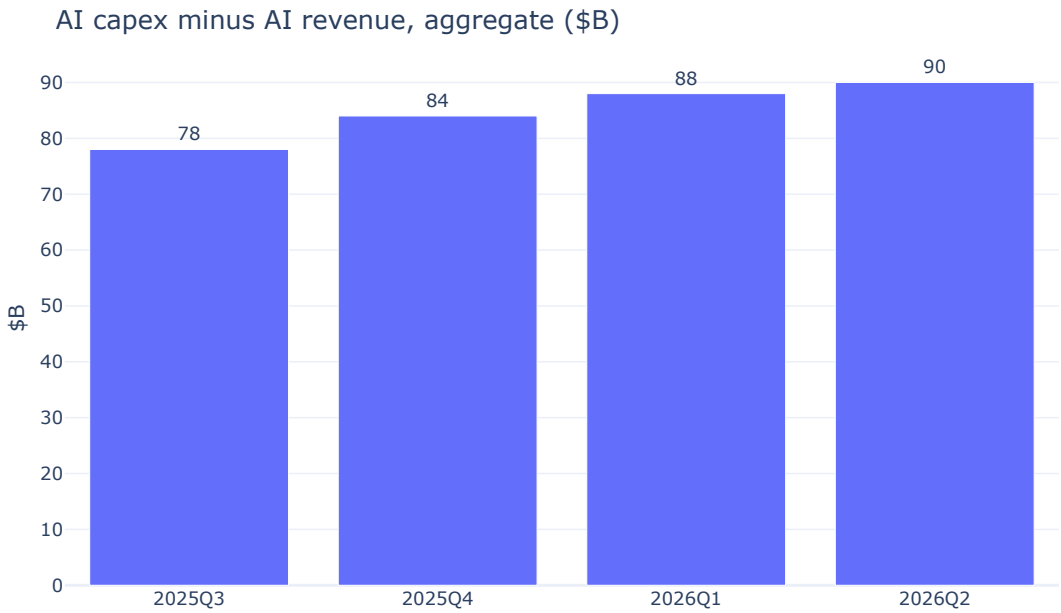
Capex vs Demand Gap

This indicator asks whether AI capital spending is outrunning the demand that would justify it. It pairs two readings: how much faster capex is growing than the revenue it is meant to serve, and whether full-segment revenue clears a deliberately generous break-even hurdle. The hurdle is the methodology's break-even factor — required revenue per dollar of capex, set so the firm gets credit for *all* segment revenue, not just AI-specific lines.

The factor follows §2.2 exactly: `(CoC + 1/L) / m` with `CoC = 10%`, `L = 6yr`, `m = 30%`, giving **0.889** of revenue required per dollar of capex per year. A firm fails when FY2025 segment revenue falls below capex × 0.889.

firm	capex $B	segment rev $B	break-even rev $B	headroom $B	head-room %	capex/rev growth x	break-even
GOOGL	91.45	58.705	81.3	−22.6	−28	2.07	FAIL
AMZN	128.3	128.725	114	14.7	13	3.25	pass
MSFT	64.55	106.26	57.4	48.9	85	—	pass
META	69.69	200.97	61.9	139	224	3.95	pass
AMD	1	16.635	0.9	15.7	1771	—	pass

The capex gap — AI capital outlay net of AI revenue, aggregated across the hyperscalers — is the same picture viewed at the system level. It does not narrow.



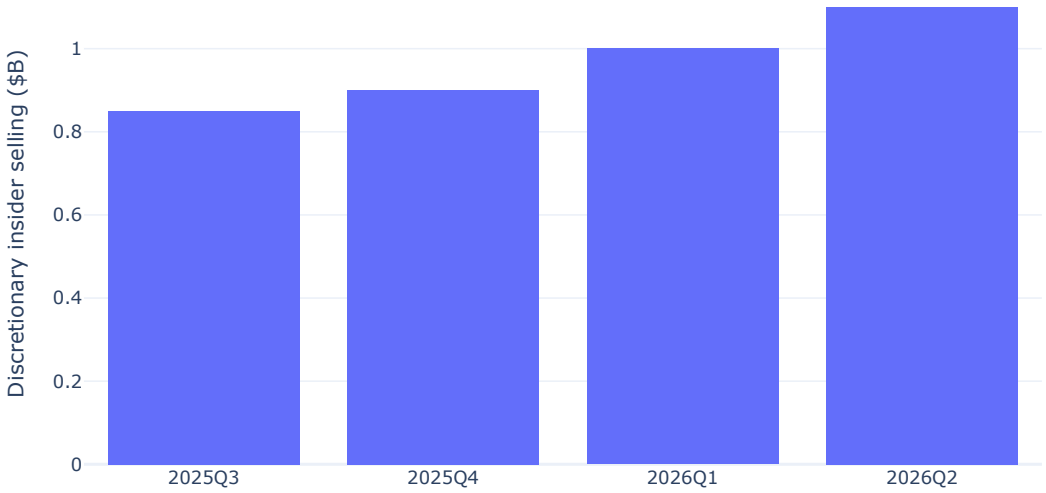
One firm fails the break-even test on full segment revenue: Alphabet, where Google Cloud's $58.7B sits $22.6B below the $81.3B the capex requires — a 28% shortfall. Capex is also growing roughly 2–4× faster than the revenue lines it funds across the cohort (GOOGL 2.07×, AMZN 3.25×, META 3.95×), even where the level test still clears. The aggregate gap widens from $78B to $90B over four quarters; spending is being committed ahead of the demand, and the test is constructed to flatter the firms, not to indict them.

Insider Selling Intensity

Indicator I3 separates two kinds of insider selling that look identical on a tape and mean opposite things. Rule 10b5-1 plan sales — pre-committed, scheduled, blind to subsequent information — score 20–40 regardless of dollar size. The signal is *discretionary* code-S activity: a sale an officer chose to make, in a window when that officer held material non-public information, with no 10b5-1 footnote on the Form 4. The table below splits each issuer's window total along that line.

ticker	score	discretionary $B	plan 10b5-1 $B	largest dollar seller	largest seller basis
NVDA	46	0.93	1.57	Stevens Mark A (Director)	discretionary (no plan detected)
AVGO	57	0.496	0.752	Samueli Henry (Director Co-Founder)	10b5-1
AMD	37	0.016	0.294	Su Lisa T (Chair President & CEO)	10b5-1 (adopted 2025-09-09)

The three compute leaders divide cleanly. NVDA carries $0.93B of discretionary code-S — director Mark Stevens alone at $802M, with no detected plan — against $1.57B run through confirmed 10b5-1 plans, including CEO Huang's $1.05B (under 1% of his stake). AVGO shows $0.50B discretionary spread across the entire C-suite — CEO Tan, the CLO, the CFO, and two more officers, none with a detected plan — while its largest single ticket, co-founder Samueli's $749M, is plan-based and therefore discounted. AMD is the quiet one: $16M discretionary, $294M planned. Note the column ordering matters more than the dollar leader. AVGO's and AMD's largest sellers by dollar (Samueli, Su) are both 10b5-1; the discretionary cluster, not the headline number, is what scores.



Discretionary selling is not a one-quarter event. The universe-level Form 4 total rises every quarter across the window, from $0.85B in 2025Q3 to $1.10B in 2026Q2 — a 29% increase while the same names were guiding investors toward accelerating AI demand. The pattern is what the indicator is built to catch: insiders choosing to reduce exposure, in size, in the windows when they know most.

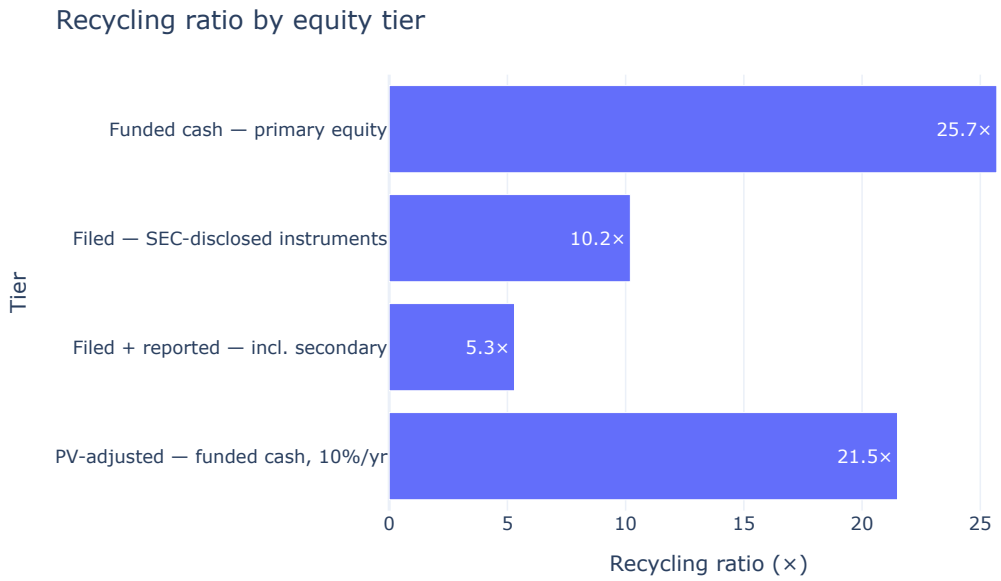
Circular Financing

The financing graph of the AI-compute complex is a directed multigraph over twelve principals, with four edge types — `invests`, `buys_compute`, `supplies`, `marks_up`. The structure it encodes is a loop: an investor funds a lab, the lab commits to buy compute from the investor's cloud, that cloud revenue underwrites the investor's capex, and the capex buys the investor's own chips through the lab it funded. The recycling ratio measures the loop's leverage — compute committed out of the core labs (OpenAI, Anthropic, xAI) divided by equity put in — across three provenance tiers. Per METHODOLOGY.md §2.4.

Tier	Equity in ($B)	Compute out ($B)	Recycling ratio (×)
Funded cash — primary equity	21	539.5	25.7
Filed — SEC-disclosed instruments	53	539.5	10.2

Tier	Equity in ($B)	Compute out ($B)	Recycling ratio (×)
Filed + reported — incl. secondary	101	539.5	5.3
PV-adjusted — funded cash, 10%/yr	21	450.8	21.5

The same dollar of disclosed equity supports roughly 26 dollars of committed compute back to the funder's own cloud on a funded-cash basis, easing to 5 dollars only when every reported secondary round is admitted as equity. Provenance, not arithmetic, is what moves the nominal ratio. Present-valued at 10% over each commitment's disclosed horizon, the funded-cash ratio is about 21× — a conservative upper bound, since the two Microsoft commitments carry no disclosed end-date and are left undiscounted; discounting them over a typical multi-year cloud term brings it nearer 18×. Stock or flow, discounted or not, the loop turns far above any arm's-length benchmark.



Provider	Commitments ($B)	Share (%)
Microsoft	280	51.9
Amazon	238	44.1
xAI	15	2.8
CoreWeave	6.5	1.2

Of the labs' committed compute — the same $540B universe as the recycling ratio — Microsoft and Amazon receive 96 percent; the loop has two destinations. On the filing-grade (PRIMARY-dollar) subset the top-two share is 98 percent; counting the media-only xAI–Google compute legs, not yet dollar-quantified in filings, would lower it further. Mark-to-model gains booked on those same customer stakes total $16.8B — Microsoft +$4.5B, Amazon +$12.3B — earnings recognized on the appreciation of the firms one funds. Six directed cycles run through the cash-flow subgraph, and the largest single commitment, Nvidia's $6.3B backstop to CoreWeave, surfaced only in a September 2025 8-K (accession 0001769628), absent from the March 2025 IPO prospectus that first sold the relationship.

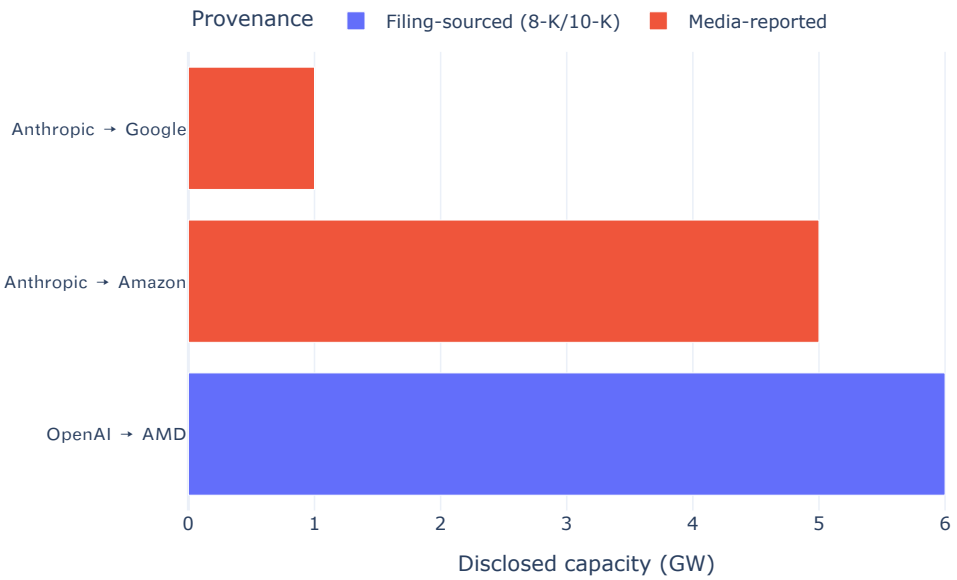
Energy and Diminishing Returns

Indicator I5 asks whether physical limits — power, cooling, chip economics — are beginning to bind on AI capability gains. It is the thinnest-data indicator in the framework and carries the lowest composite weight (0.10): the firm-level cost-per-capability curve is largely proprietary, and we will not present estimate as measurement. What the filings *do* record, unambiguously, is the scale of the power being committed. The clean signal here is the appearance of gigawatt-scale electrical-capacity figures inside the same compute-purchase agreements that drive Indicator I4 — the point at which the build stops being denominated in dollars and starts being denominated in power.

The methodology requires that any I5 score of 60 or above rest on at least one filing-sourced item, not media alone. The table below isolates every financing edge that discloses a gigawatt-scale figure, tagged by provenance.

Power commitment	Disclosed capacity	Provenance	Window
OpenAI → AMD	6 GW	PRIMARY (8-K/10-K)	through 2030
Anthropic → Amazon	5 GW	REPORTED (media)	10-year commitment signed Apr 2026
Anthropic → Google	1 GW	REPORTED (media)	Oct 2025+

Three edges carry an explicit gigawatt figure — 12 GW of committed capacity in aggregate. Exactly one is filing-sourced: OpenAI's 6 GW Instinct commitment to AMD, disclosed in AMD's 8-K EX-99.1 (filed 2025-10-06). The other two — Anthropic's 5 GW with Amazon and over 1 GW with Google — are media-attested and not yet in a filing.



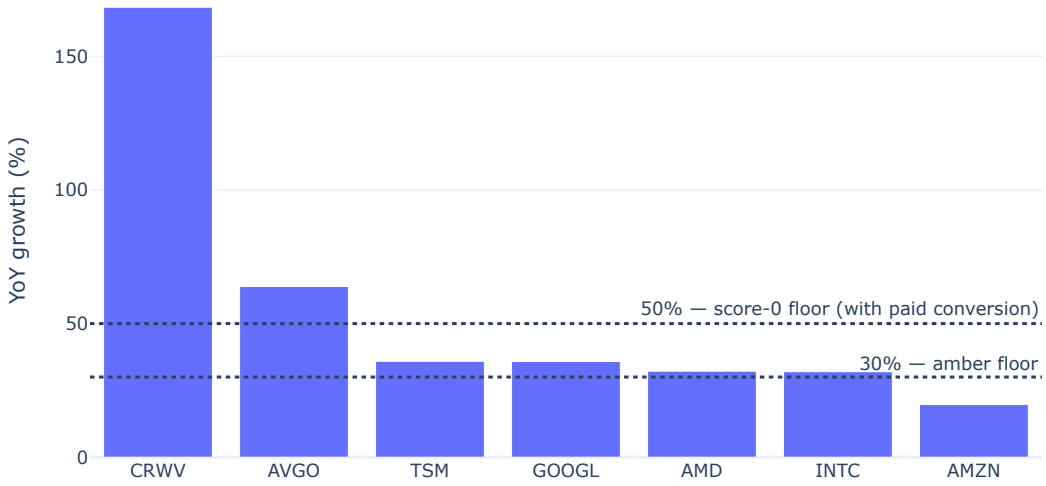
One filing carries a gigawatt-scale power figure; the rest of the disclosed capacity is currently media-attested. By the methodology's own rule, that single filing-sourced item is the floor under any elevated I5 read — the indicator is directionally supportive, not independently load-bearing, and it is flagged as such. The honest reading is that power has begun to appear in the filings as a unit of account, but the cost-per-capability curve that would make I5 stand on its own is deferred to Phase 2.

Organic End-User Demand

Indicator 6 asks whether reported AI revenue reflects genuine paid adoption by independent end-users, or is instead recycled through the same ecosystem that funds the build (the I4 counterparties) or rebranded from existing product lines. The methodology anchors the test on the MIT NANDA finding that roughly 95% of enterprise GenAI pilots show no measurable P&L impact (Fortune, August 2025):

headline revenue growth in the 30%–50%+ band scores well only when paired with demonstrated paid retention and pilot-to-production conversion above 50%; growth sourced from ecosystem participants scores worse, not better. The two views below read directly from the filings-grounded CSVs — segment revenue growth, and the counterparty composition of the clearest recycled-demand case.

FY2025 AI/cloud revenue growth vs the I6 demand bands



Firm	AI/cloud rev FY24 (B) | $AI/cloudrevFY25$(B)	YoY growth (%)	
CRWV	1.9	5.1	168.4
AVGO	12.2	20	63.9
TSM	90.08	122.42	35.9
GOOGL	43.23	58.705	35.8
AMD	12.579	16.635	32.2
INTC	12.817	16.919	32
AMZN	107.556	128.725	19.7

Counterparty (I4 node)	Committed / disclosed share	Window
Meta	14.2B	through Dec 2031
Microsoft	67% of FY2025 revenue	FY2025

Counterparty (I4 node)	Committed / disclosed share	Window
Nvidia	320M	through 2024
Nvidia	6.3B (initial backstop obligation)	through Apr 13 2032
OpenAI	6.5B	through May 2031

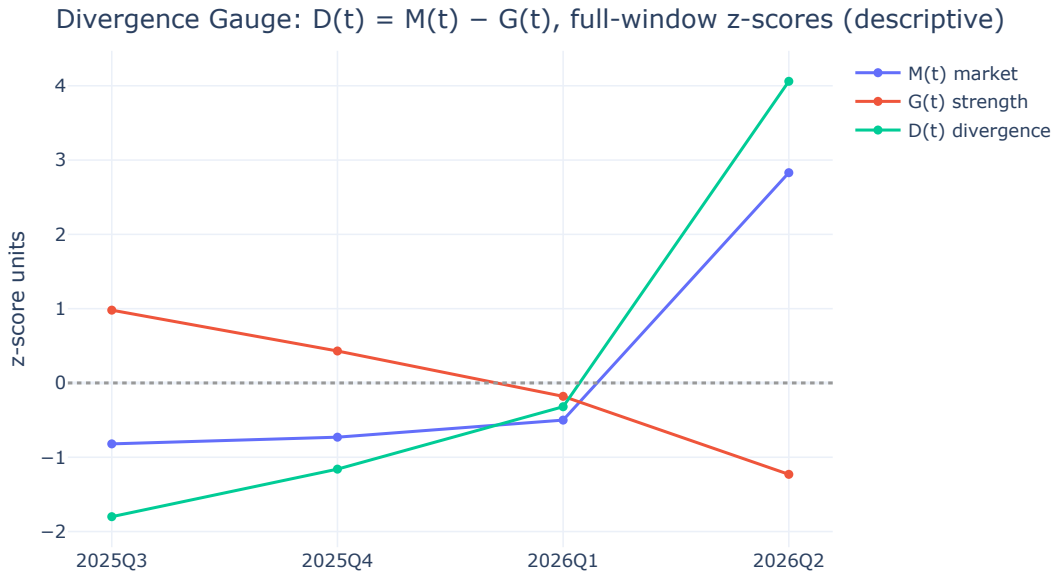
Revenue growth alone clears the methodology's headline band for most of the complex — CoreWeave at 168%, Broadcom at 64%, and four firms clustered at 32%–36% — but the indicator scores the *source* of that growth, not its rate. CoreWeave is the limiting case: 67% of its FY2025 revenue is a single I4 counterparty (Microsoft, "Customer A" in its 10-K), with the remainder committed by OpenAI, Meta, and Nvidia — every named buyer is an investor in, or a lab funded by, the same circular structure. That is growth from ecosystem participants rather than demonstrated independent end-user retention, which is exactly the 80-band the rubric reserves, and it is consistent with the NANDA anchor: an "AI revenue" label growing fastest where the demand is most recycled, not where paid pilot-to-production conversion is most proven.

Divergence Gauge

The Divergence Gauge `D(t) = M(t) − G(t)` sets a market signal built from SOXX price behaviour against a ground-truth strength signal built from the deterioration series. It widens when semiconductor momentum and overextension climb while the underlying fundamentals erode — the gap between what the tape says and what the filings say. The market term `M(t)` is the equal-weight mean of three full-window z-scored components (63-day momentum, price-to-SMA50 overextension, 20-day annualized instability); the ground-truth term `G(t) = −mean(z)` over `ai_layoff_share`, `insider_disc_bn`, and `capex_gap_bn` (`depr_phantom_bn` excluded — constant, zero variance). Per METHODOLOGY.md §4.

Quarter	M(t) market	G(t) strength	D(t) divergence	z: AI layoff	z: insider sell	z: capex gap
2025Q3	−0.82	0.98	−1.8	−0.6	−1.01	−1.32
2025Q4	−0.73	0.43	−1.16	−0.54	−0.56	−0.19
2026Q1	−0.5	−0.18	−0.32	−0.36	0.34	0.57

Quarter	M(t) market	G(t) strength	D(t) divergence	z: AI layoff	z: insider sell	z: capex gap
2026Q2	2.83	−1.23	4.06	1.49	1.24	0.94



Through 2025Q1 the two signals track close to each other and `D(t)` sits below zero: price had not yet detached from fundamentals. In 2026Q2 the gap inverts hard — `M(t)` jumps to +2.83 as SOXX closes at 639.45 (63-day momentum +88.0%, instability +0.74 annualized) while `G(t)` falls to −1.23, dragged by the AI-layoff share and discretionary insider selling both reaching their window highs. `D(t)` widens from −1.80 to +4.06, a +5.86 swing, the strongest single-quarter divergence in the series.

Method and limitations

This brief is reproducible by construction: re-running it re-derives every figure from the source data. Each indicator is computed only from filing-sourced inputs; where a value cannot be sourced cleanly it is shown blank rather than imputed. Two Phase-1 simplifications are stated plainly. The divergence gauge standardizes its components over the full window — descriptive, not real-time: it carries look-ahead bias and is not a tradeable signal, and an expanding-window version is deferred. The gauge also weights its three market components equally; empirical calibration is future

work. Indicator I5 (energy) rests on the thinnest data in the set and is weighted accordingly — directionally supportive, not independently load-bearing.