

Breaking Plausible Deniability: A Consensus-Based Definition of "Actionable Dark Fleet" Behaviors for Subsea Cable Protection

Koichiro Hayashi ¹*

*** Presenting Author**

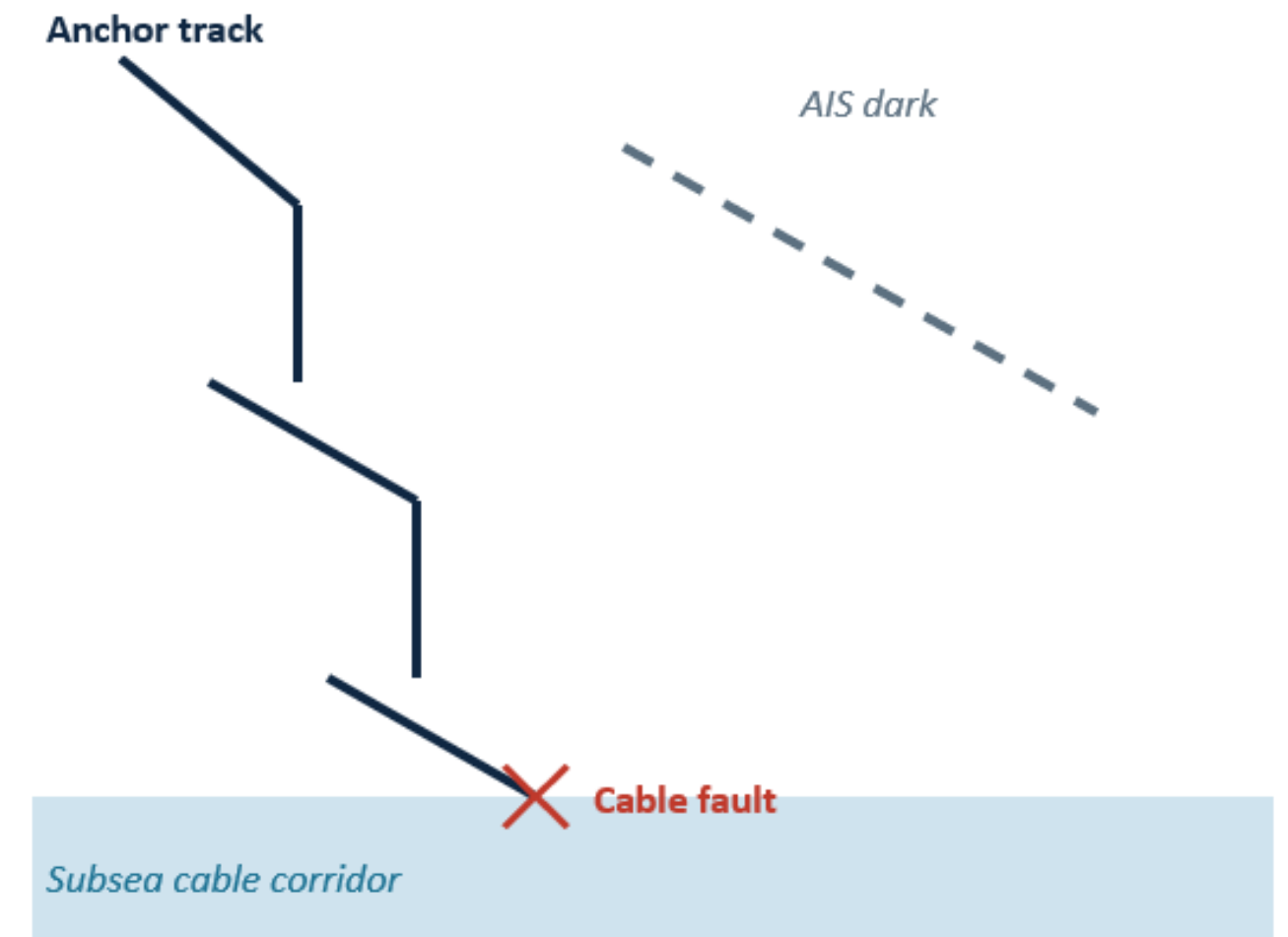
¹ NYK Line

A Vessel That Looks Guilty

- Loiters for hours over a cable corridor — no port call, no congestion, no weather reason
- Weak-oversight flag; opaque ownership; repeated AIS gaps
- An anchor track inconsistent with random drift
- A cable fails within the same window
- ***Analysts argue:*** too many signals overlap to read as ordinary navigation

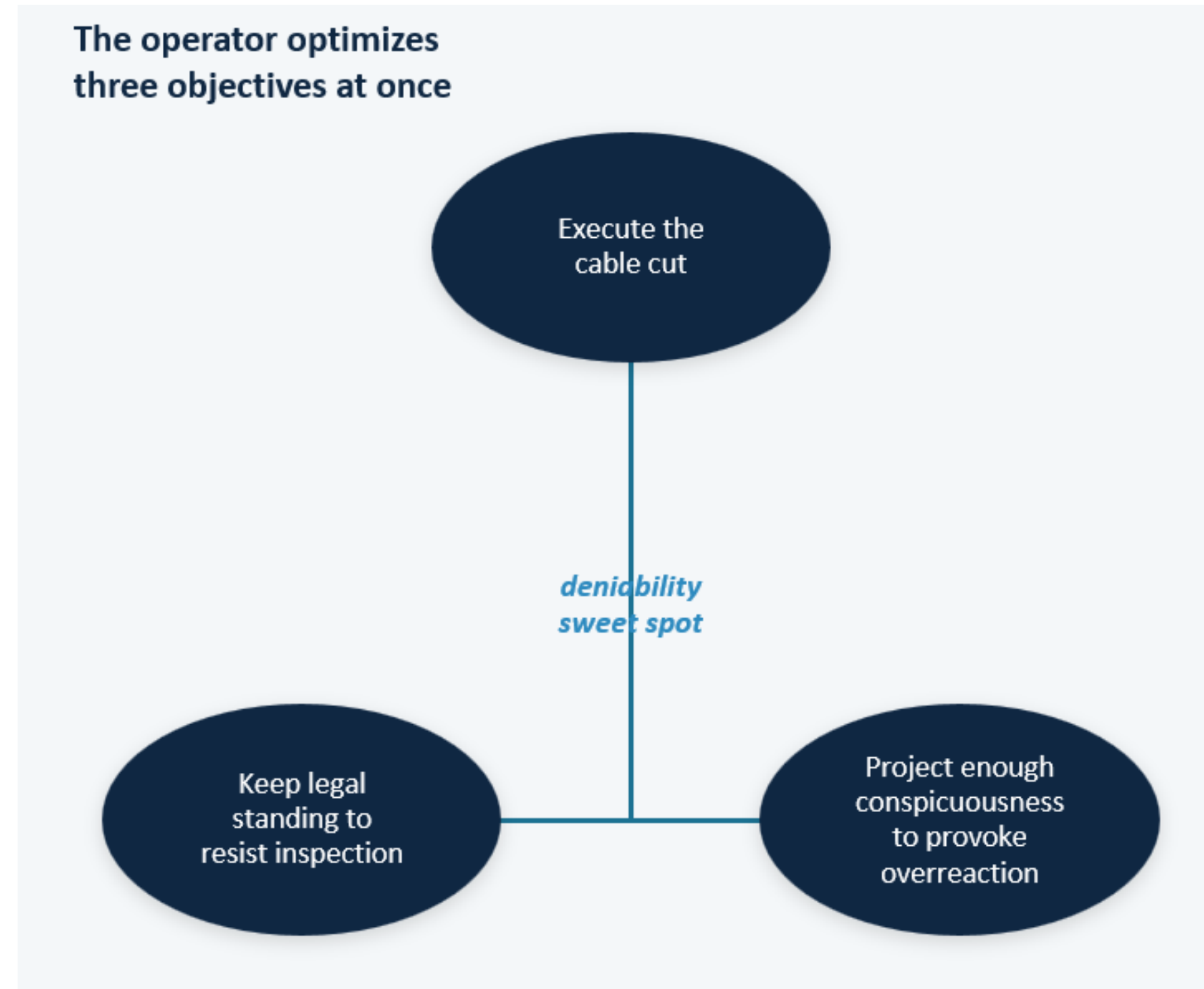
Yet intent cannot be proven — the ship is, in effect,
unprovably guilty

STYLIZED COMPOSITE — NOT A SPECIFIC CASE



Plausibility Serves a Goal Beyond the Cable Cut

- Under UNCLOS, inspection on the high seas needs the flag state's consent - direct enforcement is hard by default
- So opacity is not only a shield - the same opacity makes the ship look suspicious, on purpose, before any cut
- This plausibility is engineered not to raise the chance of cutting a cable - but to serve a goal beyond the cut



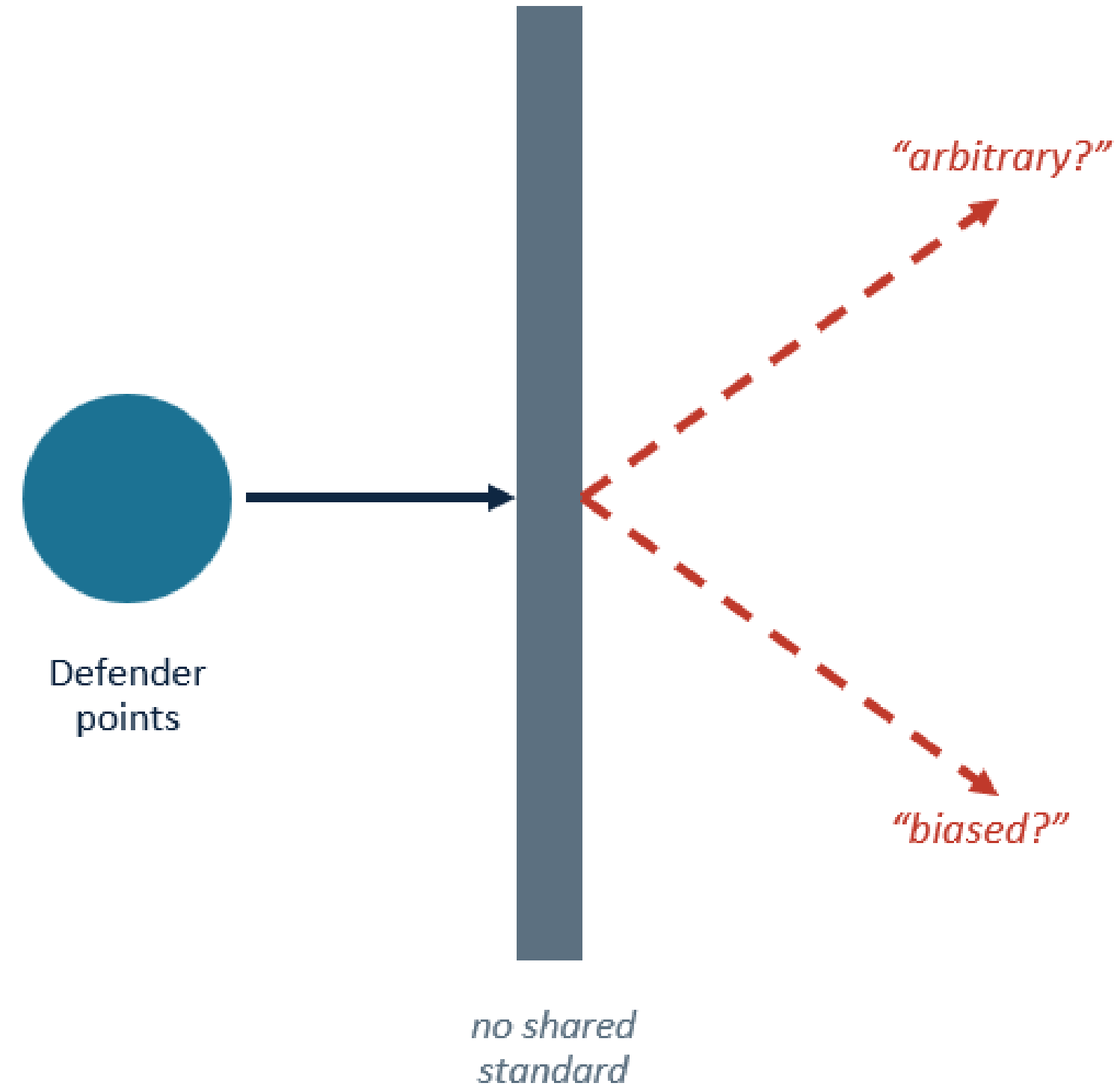
Calculated Conspicuousness

- Classic sanctions-evasion dark fleet: wants to **hide** — “don’t get caught”
- Sabotage vessel: wants to **be seen** — but untouchable
- Same observable attributes... opposite intent about visibility
- The suspicious look is not incidental — it is a **chosen signal**

	Sanctions-evasion dark fleet	Sabotage vessel
Goal of visibility	Stay hidden	Be seen, stay untouchable
Shared attributes	Opacity, AIS gaps, old tonnage	<i>Same</i>
Strategic intent	Avoid detection	Provoke + deny

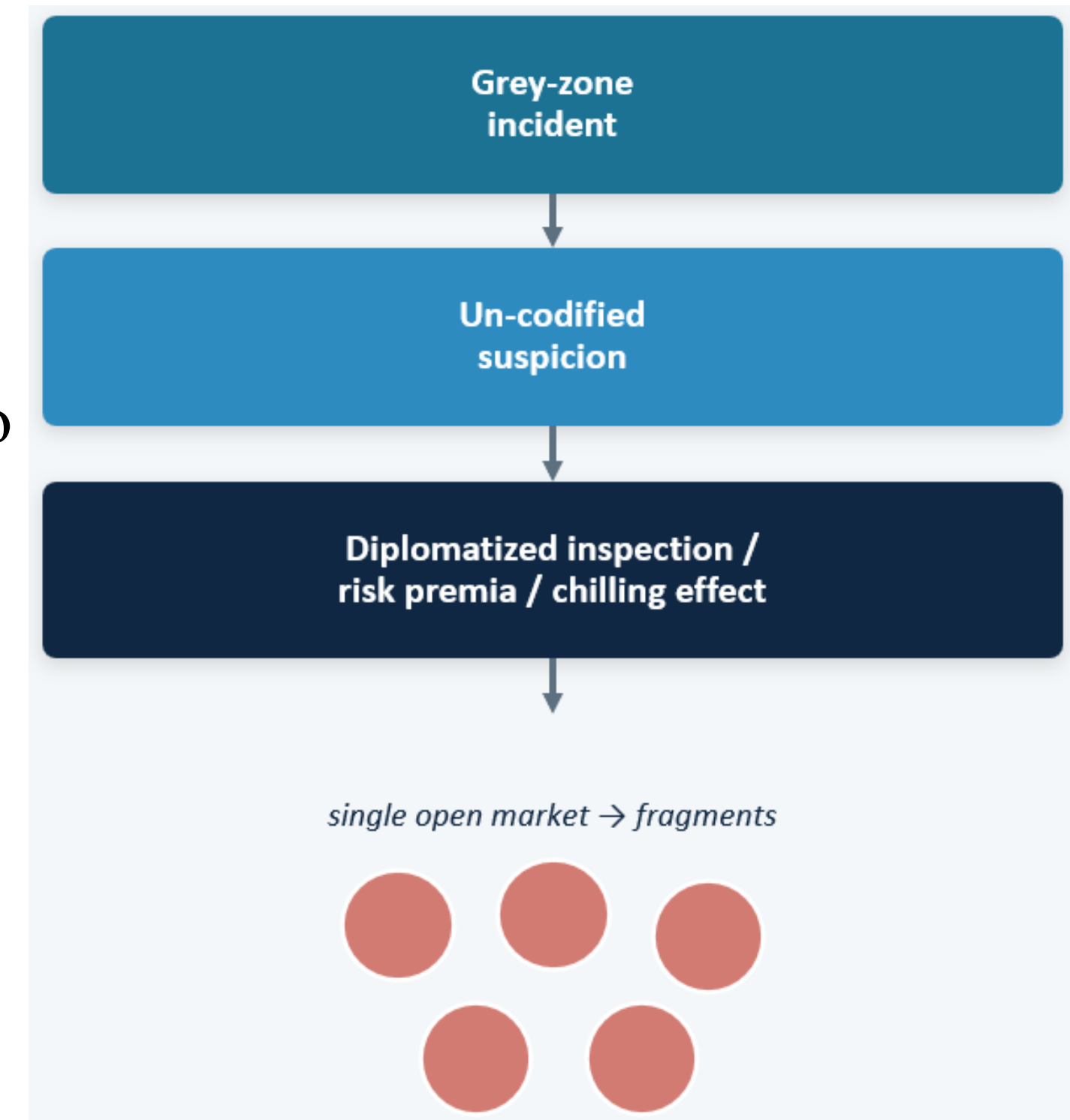
The Attacker's Optimal Point: Suspicion That Stays Un-Codified

- The damage is not only the cut cable — it is also the erosion of an open, rules-based order
- This works only while suspicion stays **informal and contestable**
- No shared standard → every accusation is deflected as “arbitrary,” “biased,” “selective”
- Targeted states retaliate under the same pretext → freedom of navigation hollows out
- The optimal point is not the absence of suspicion — it is the **non-codification** of suspicion



The Real Target Is the Market, Not the Cable

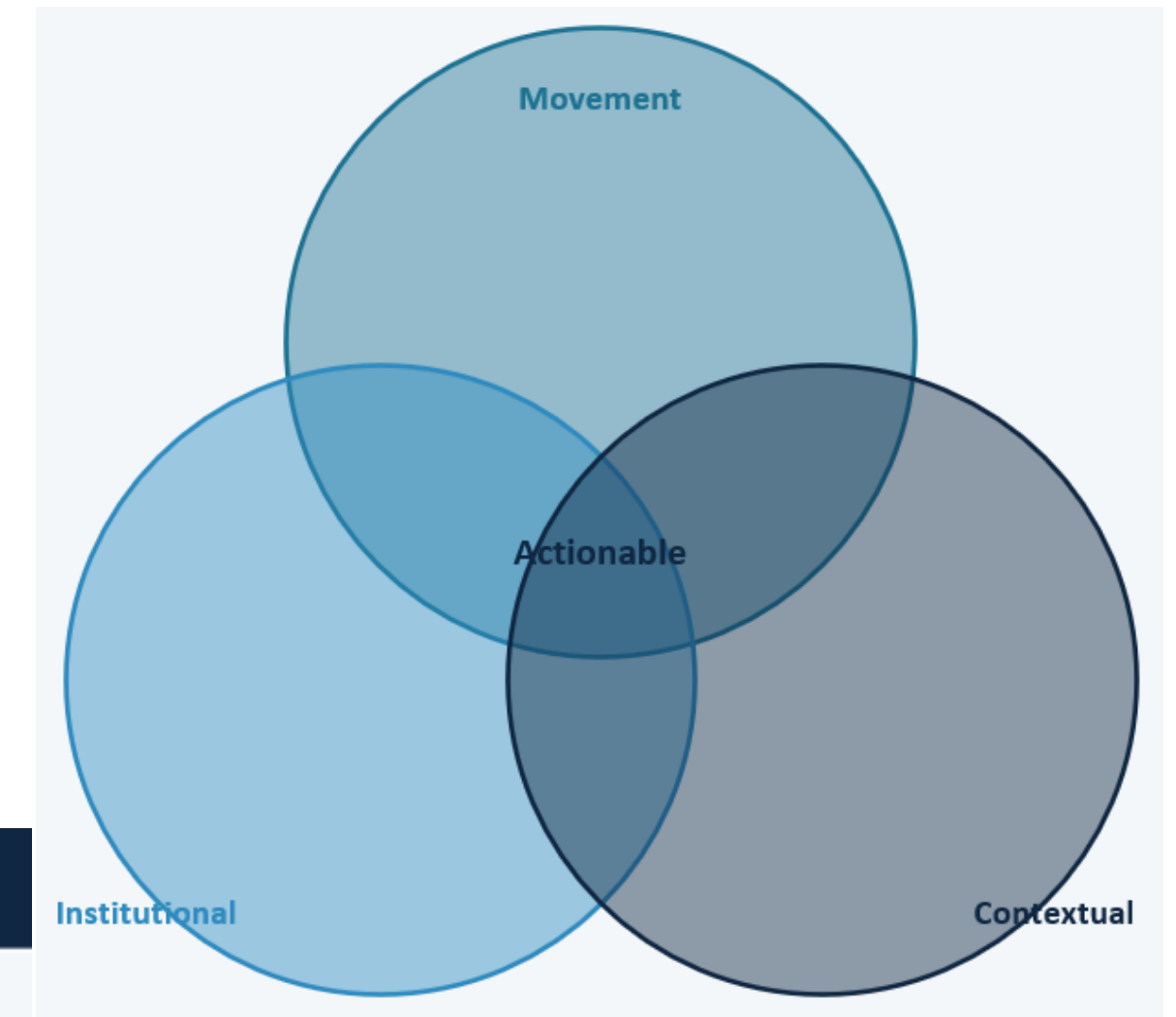
- **Harm cascades well beyond the sabotage itself:**
 - Inspection procedures become diplomatized
 - Risk premia and an “**ambiguity tax**” rise
 - Shipping markets **fragment** along security lines
- Ultimate objective: reconfigure the single open market into something they can manipulate
- This is a **maritime-economics** problem — not only a security one
- The ambiguity tax is borne by honest operators, via chilling effects on legitimate trade



Naming the Suspicion: An Actionable Dark Fleet Index (ADFI)

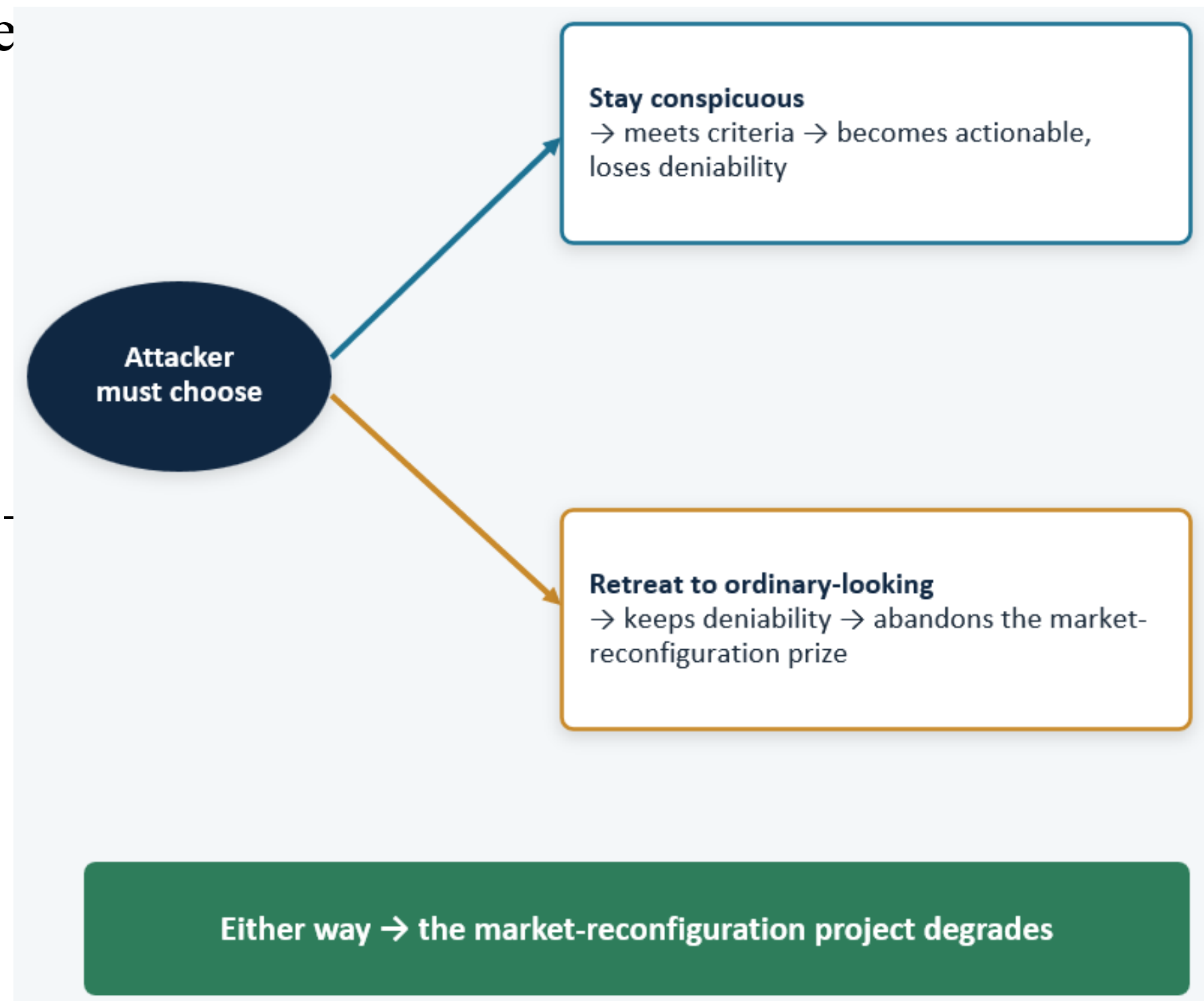
- Not a single indicator — *an AIS gap alone proves nothing*
- A **multi-signal, transparent** trigger — independent signals must overlap
- **Procedurally constrained:** logging, flag-state notification, third-party review
- **Economically proportionate:** high-consequence contexts only (near critical corridors)

Signal Category	Examples
Movement	Loitering, anomalous anchoring, AIS dark periods
Institutional	Flag quality (PSC lists), P&I insurance, ownership opacity
Contextual	Proximity to cable corridors, timing of reported outages



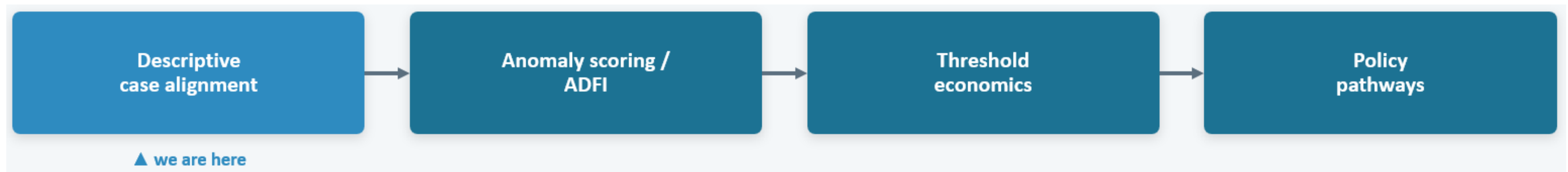
Breaking Deniability: The Attacker's Mirror Dilemma

- A shared, objective standard removes the gap the attacker lives in
- **The attacker is forced to choose — and either branch degrades the project:**
- The cable-cutting risk itself does *not* fall — an ordinary-looking ship can still drag an anchor
- But the strategic fruit — manipulating the order — is **denied**



A Research Agenda — and Its Hard Problems

- Build ADFI: expert-informed + supervised + unsupervised weights (avoid overfitting)
- Align AIS features with documented incident timelines
- Benchmark against benign explanations: weather, congestion, mechanical, legitimate anchoring
- Evaluate thresholds economically: deterrence vs. delays, false positives, fragmentation



OPEN PROBLEMS (work-in-progress)

- The standard can be attacked as “Biased” → robustness as a focal point is a design requirement, not an assumption
- Adversarial adaptation — but here, evading the index means abandoning the goal
- Labeling is scarce; attribution is hard

What I Can Tell You — What We Build Together

- Even without changing UNCLOS, a diffused objective definition **shrinks the grey-zone operating space**
- It lowers coordination costs among willing states and stakeholders
- It strips away the “arbitrariness” card — the attacker’s **last line of deniability**
- ***Open question:*** who designs the standard, and how does it stay credible to those who might be on the other side?

