

UOIF Reflection: Proof-Driven Recompute of $\Psi(x)$ with Confidence and Live Verification

UOIF Working Note

August 08, 2025

1 Model and Proof Logic

$$\Psi(x) = \underbrace{[\alpha S(x) + (1 - \alpha) N(x)]}_{\text{hybrid linearity}} \cdot \underbrace{\exp(-[\lambda_1 R_a + \lambda_2 R_v])}_{\text{exponential penalty}} \cdot \underbrace{P(H | E, \beta)}_{\text{Bayesian posterior}}, \quad \lambda_1=0.85, \lambda_2=0.15, S(x)=0.60.$$

Proof logic: (i) $O(\alpha) = \alpha S(x) + (1 - \alpha)N(x)$ is affine; $\partial O/\partial \alpha = S(x) - N(x) < 0$ if $N(x) > S(x)$ (monotone benefit to external). (ii) With $R_a, R_v \geq 0$, Penalty $\in (0, 1]$ ensures boundedness and regularization. (iii) $P(H | E, \beta) = \min\{\beta P(H | E), 1\}$ encodes uplift while capping certainty.

2 RK4 Baselines and Metrics (Physics-Guided)

Runge–Kutta 4 (RK4) for $\dot{y} = f(t, y)$:

$$y_{n+1} = y_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad k_1 = f(t_n, y_n), \quad k_2 = f(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1), \dots$$

Local error $\mathcal{O}(h^5)$, global $\mathcal{O}(h^4)$; trajectories serve as symbolic baselines. Metrics: RMSE = $\sqrt{\frac{1}{T} \sum \|\hat{y}_t - y_t\|_2^2}$, $R^2 = 1 - \frac{\sum \|\hat{y}_t - y_t\|_2^2}{\sum \|y_t - \bar{y}\|_2^2}$.

3 IMO 2025 Results (Canonical live)

Sources: official results pages + DeepMind 2025 + Evan 2025 + AoPS [?, ?, ?, ?, ?, ?, ?]. Confidence: 0.98. Settings: $N=0.97$, $\alpha \in [0.12, 0.15]$, $\beta=1.15$; $O(\alpha) = 0.97 - 0.37\alpha$. Confidence: 0.96. Penalty (conservative mixed state): $R_a=0.15, R_v=0.05 \Rightarrow \text{Penalty} = \exp(-0.135) = 0.8737$. Confidence: 0.85. Posterior (conservative): $P(H | E, \beta) \approx 0.913$. Confidence: 0.85. Value (conservative): $\alpha=0.12 \Rightarrow \Psi(x) \approx 0.9256 \cdot 0.8737 \cdot 0.913 \approx \mathbf{0.738}$; $\alpha=0.15 \Rightarrow \mathbf{0.726}$. Label: Empirically Grounded. Confidence: 0.90. Penalty (eased for canonical results): $R_a=0.12, R_v=0.04 \Rightarrow \text{Penalty} = \exp(-0.108) = 0.8977$. Confidence: 0.85. Posterior (capped): $P(H | E, \beta) = 1.0$. Confidence: 0.88. Value (eased): $\alpha=0.12 \Rightarrow \mathbf{0.831}$; $\alpha=0.15 \Rightarrow \mathbf{0.821}$. Label: Primitive/Empirically Grounded (results primitives). Confidence: 0.90. Sensitivity: $\partial \Psi(x)/\partial \alpha < 0$. Confidence: 0.92.

4 IMO 2025 Problems (Pending canonical)

Sources: Evan 2025 + AoPS + DeepMind 2025 (expert) [?, ?]. Confidence: 0.88. Settings: $\alpha \in [0.15, 0.20]$, $N \in [0.88, 0.90]$, $\beta=1.05$. Penalty: $R_a=0.25, R_v=0.10 \Rightarrow \text{Penalty} \approx 0.7965$. Posterior: $P(H | E, \beta) = 0.945$. Value (midpoint $\alpha=0.17, N=0.89$): $\Psi(x) \approx 0.633$; range **0.60–0.65**. Label: Interpretive/Contextual. Confidence: 0.85.

5 IMO 2024 (DeepMind P1/P2/P4)

Sources: DeepMind 2024 pages + Evan + archives [?, ?]. Confidence: 0.90. Settings: $N=0.96$, $\alpha \in [0.10, 0.15]$, $\beta=1.05$; $O(\alpha) = 0.96 - 0.36\alpha$. Penalty: $\text{Penalty} = \exp(-0.0925) \approx 0.9117$. Posterior: $P(H \mid E, \beta) = 0.945$. Value: $\alpha=0.10 \Rightarrow \mathbf{0.796}$; $\alpha=0.15 \Rightarrow \mathbf{0.781}$. Label: Primitive/Empirically Grounded. Confidence: 0.88.

6 Keystone Reflections

$\Psi(x)$ functions as an evidential synthesizer: it (i) privileges canonical artifacts, (ii) regularizes overconfidence via exponential damping, and (iii) calibrates expert uplift through a capped posterior. This enables transparent, monotone promotions as official artifacts appear—supporting AI-driven mathematics where certified interpretive proofs coexist with canonical primitives.

7 Verification (Real-Time: Aug 08, 2025)

Official 2025 **results** pages are live (year info, country, individual, statistics), while the **problems** page remains pending; this matches prior exchanges and our gating.

References