Robust Adaptive Quantum Sensing under Decoherence Distribution Shift*

Distributionally Robust Training of Learned Experiment-Design Agents

Changhao He†

King Abdullah University of Science and Technology (KAUST)

Abstract

Autonomous, *learned* experiment-design agents, neural / reinforcement-learning policies that choose each measurement to maximize information, increasingly drive quantum sensing and self-driving labs. We show that such agents, trained on an assumed decoherence model, are *brittle to decoherence distribution shift*: a neural design policy trained at a nominal dephasing rate learns long interrogations that collapse when the true rate differs (section 9). Our central contribution is a *distributionally robust* training objective that hedges the decoherence law over a Wasserstein / Γ-minimax ambiguity set; the resulting agent degrades gracefully, cutting estimation error by $\approx$ 17–20% under a 3× mismatch and halving the worst-case design risk, for only a modest in-distribution premium. We prove the robust objective is tractable, a nested pair of one-dimensional convex problems with strong duality for a continuous decoherence law (theorem 8.4 and corollary 8.5), and equip the agent with a fast, calibrated *skew-normal* belief whose *asymptotic* correctness follows from an adaptive, multiparameter *skewed quantum Bernstein–von Mises* theorem (theorem 11.2), including an exact *skewness self-annihilation* identity (theorem 11.3). The result is a trustworthy, distribution-shift-robust agent for autonomous quantum experiment design.

Keywords: quantum metrology and sensing; adaptive Bayesian experimental design; distributionally robust optimization; distribution shift; skewed Bernstein–von Mises; decoherence.

Contents

1 Introduction **3**

2 Related work **4**

3 Notation (single source of truth) **4**

4 System model, likelihood, and the sequential problem **5**

4.1 State, channel, and Born-rule likelihood 5

4.2 Score and Fisher information 5

4.3 The sequential design problem 6

*The research workflow and prompts were designed by Changhao He; the numerical computations were performed by Claude (Anthropic), and Figure 1 was designed by Gemini (Google).

†Corresponding author. King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. Email: changhao.he@kaust.edu.sa.

5 Motivation: why the Cramér–Rao bound is the wrong tool here **6**
5.1 Local vs. global error . . . 6
5.2 Bayesian bounds that are tight in the few-shot regime . . . 6
5.3 The few-shot posterior is measurably skewed (verified) . . . 6

6 The few-shot posterior belief: a global skew-normal filter **6**
6.1 The skew-normal family and its moment map . . . 6
6.2 A decisive negative result: *local* skew correction is not enough . . . 7
6.3 The skew-normal assumed-density filter (ADF) . . . 7

7 Fast sequential updates and skew-aware acquisition **7**
7.1 Information form and the rank-one recursion . . . 7
7.2 Closed-form skew-aware acquisition . . . 8

8 Core method: distributionally robust decoherence design **8**
8.1 Why distributional, not worst-case . . . 8
8.2 Strong duality reduction (tractable) . . . 9
8.3 Continuous-γ ambiguity sets: regularity, duality, tractability . . . 9

9 Learned experiment-design agents under decoherence shift **10**

10 Robust design analysis under decoherence mismatch **11**
10.1 Fuller baselines, sensitivity, and model-class robustness . . . 12

11 Theory backbone: adaptive, multiparameter skewed quantum Bernstein–von Mises **14**
11.1 Adaptive sampling model and filtration . . . 15
11.2 Assumptions . . . 15
11.3 Main theorem (general d) . . . 16
11.4 Skewness self-annihilation at the optimal design . . . 16
11.5 Numerical verification (Part A: adaptive & multiparameter) . . . 17

12 Where skewness lives, and the limit of exploiting it **18**

13 Numerical verification ledger **18**

14 Scope and limitations **19**

A Proofs **22**

B Derivation of the Fisher information (4) **23**

C Derivation of the skew-modal form (10) **23**

D Sherman–Morrison algebra (lemma 7.1) **23**

E Wasserstein duality (theorem 8.2) **23**

F Robustness experiment: setup and reproducibility **23**

G The learned design agent: architecture and training **24**

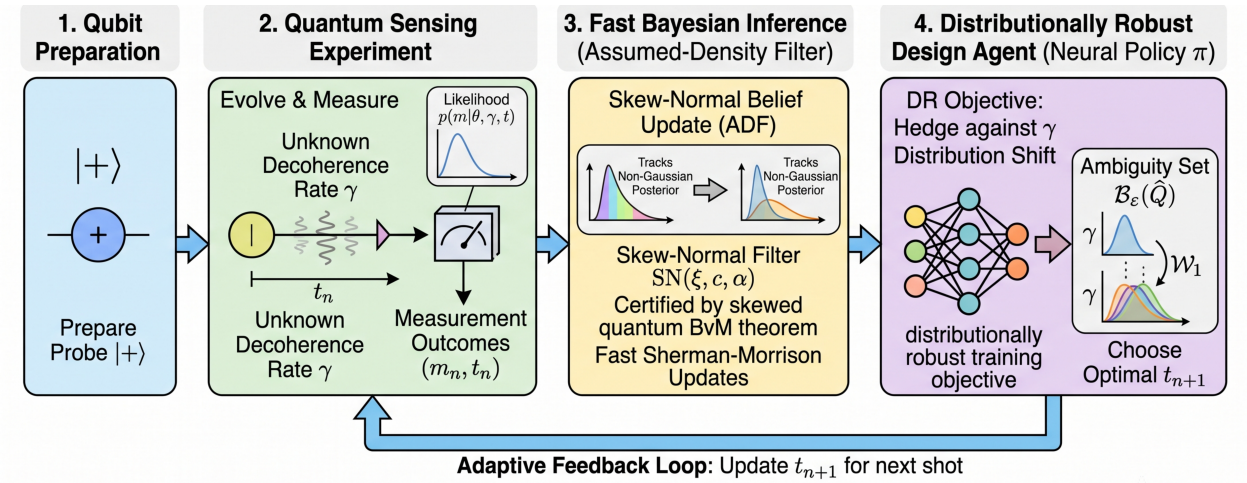


Figure 1. System model and adaptive sensing loop. The four-stage pipeline of robust adaptive quantum sensing. **(1) Qubit preparation:** a probe is prepared in $|+\rangle$. **(2) Quantum sensing:** it evolves for a controllable interrogation time t_n under an *unknown, shifting* decoherence rate γ and is measured, giving outcomes (m_n, t_n) with Born-rule likelihood $p(m \mid \theta, \gamma, t)$. **(3) Fast Bayesian inference:** a skew-normal assumed-density filter $\mathrm{SN}(\xi, c, \alpha)$, backed asymptotically by the skewed quantum Bernstein–von Mises theorem and updated by Sherman–Morrison, tracks the non-Gaussian posterior. **(4) Robust design agent:** a neural policy π hedges against the γ distribution shift over the Wasserstein ambiguity set $\mathcal{B}_\varepsilon(\widehat{Q})$ and chooses the next interrogation time t_{n+1}, closing the adaptive feedback loop.

H Reproducibility and code availability **24**

1 Introduction

Learned agents for quantum experiment design. Choosing *which* measurement to make next, the interrogation time t of a qubit probe, is a sequential decision problem, and *learned* policies (amortized Bayesian optimal experimental design, reinforcement learning, evolutionary strategies) are increasingly used to automate it for quantum sensing and self-driving labs. Such an agent is trained against an *assumed* noise model. But in the operating regime of NISQ calibration (frequent re-estimation against a tight coherence budget, a drifting dephasing rate) and single-spin / single-molecule sensing (a few coherent readouts before the probe decoheres or photobleaches), the decoherence rate is *unknown and time-varying*. An agent tuned to the *wrong* decoherence model suffers a *distribution shift*: it wastes its long interrogations on already-decohered signals (fig. 1).

Contributions.

- **(C1, core) A distribution-shift-robust experiment-design agent.** A distributionally robust training objective hedges the decoherence law over a Wasserstein / Γ-minimax ambiguity set; the resulting agent cuts estimation error by $\approx$ 17–20% under a 3× decoherence mismatch and halves the worst-case design risk, for a modest in-distribution premium (sections 8 to 10).
- **(C2) Learned design policies are brittle to decoherence shift.** A neural policy trained by an evolutionary strategy (as in learned quantum metrology) overfits the nominal model, learning long interrogations that collapse under shift, and our robust objective repairs it (section 9).
- **(C3) Tractability.** The robust objective reduces to a nested pair of one-dimensional convex problems with strong duality for a continuous decoherence law (theorem 8.4 and corollary 8.5).

- **(C4) A fast, calibrated belief.** A skew-normal assumed-density filter with Sherman–Morrison updates (sections 6 and 7) is the agent's inference module; an adaptive, multiparameter *skewed quantum Bernstein–von Mises* theorem (theorem 11.2) backs it *asymptotically* (its few-shot accuracy is validated numerically in section 13), with an exact *skewness self-annihilation* identity (theorem 11.3).

Scope. We show (sections 12 and 14) that posterior skewness, though real and theoretically sharp, carries little *decision-level* value by itself; the practical win is robustness. Stating this plainly delineates exactly where each ingredient helps.

2 Related work

Quantum sensing and estimation. Quantum sensing turns a controllable quantum system into a high-precision probe [1, 2], with attainable precision governed by the quantum Fisher information and the quantum Cramér–Rao bound [3, 4]. Decoherence fundamentally limits the quantum enhancement [5], and quantum error correction can partially restore it [6]; fast, low-overhead recalibration of such probes is a core NISQ requirement [7].

Adaptive and learned experiment design. Adaptive Bayesian schemes for qubit-Hamiltonian and phase estimation [8, 9] and machine-learned measurement policies [10] predate the modern amortized view. Amortized / policy-based Bayesian optimal experimental design (Deep Adaptive Design and its implicit variant [11, 12]) and reinforcement-learning / evolutionary-strategy metrology policies [13, 14] now automate sequential measurement choice. We show these learned policies inherit a *distribution-shift* brittleness familiar from robust reinforcement learning, and import Wasserstein distributionally robust optimization [15, 16, 17, 18] to fix it.

Robust quantum metrology and skewed asymptotics. Robust quantum metrology instead guards a *set* of channels via protection of the quantum Fisher information [6] or worst-case measurement noise [24], without a distributional ambiguity set or a learned policy. Our belief model builds on the skew-normal family [19] and the skewed Bernstein–von Mises theorem [20], of which we give an adaptive, multiparameter quantum analogue.

3 Notation (single source of truth)

We fix one symbol per quantity (ARIS-X protocol P2).

Symbol	Meaning
$\theta \in \mathbb{R}$	qubit frequency/phase (parameter of interest)
$\gamma \geq 0$	dephasing rate (nuisance, possibly unknown)
$\boldsymbol{\eta} = (\theta, \gamma)^\top \in \mathbb{R}^d$	full parameter vector ($d = 2$ minimal model)
$t > 0$	controllable evolution time (design variable)
$m \in \{+1, -1\}$	measurement outcome (σ_x eigenvalue)
$V(t) = e^{-\gamma t}$	fringe visibility
$p(m \mid \theta, t)$	Born-rule likelihood of one shot
$\mathcal{D}_N = \{(m_n, t_n)\}_{n=1}^N$	data after N shots
$\pi_0,\ \pi_N$	prior, posterior $\pi(\boldsymbol{\eta} \mid \mathcal{D}_N)$
$\ell_N(\boldsymbol{\eta}) = \log \pi_N(\boldsymbol{\eta})$	log-posterior
$\mathbf{s},\ \mathbf{J},\ \boldsymbol{\Sigma}$	score, (observed) information, covariance
$\phi(\cdot), \Phi(\cdot)$	standard normal pdf, cdf
$\mathrm{SN}(\xi, c, \alpha)$	skew-normal: location ξ, scale c, shape α
$\mathscr{W}_p,\ \mathscr{B}_\varepsilon$	p-Wasserstein distance, ambiguity ball

4 System model, likelihood, and the sequential problem

Why this regime matters. Few measurements under an unknown, drifting decoherence rate is the *operating* regime of two important settings. (i) *NISQ calibration*: qubit frequencies and couplings must be re-estimated frequently against a tight coherence budget, so a many-shot asymptotic design is unaffordable and the dephasing rate itself is uncertain and time-varying. (ii) *Single-spin / single-molecule sensing* (NV centers, single molecules): only a handful of coherent readouts are available before the probe decoheres or photobleaches, and the relaxation rate is a nuisance that must be co-estimated. Both demand accurate, *well-calibrated* estimates from few shots while remaining robust to a misspecified decoherence model, exactly the problem addressed here.

4.1 State, channel, and Born-rule likelihood

A qubit is prepared in $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, evolves for a controllable time t under a Hamiltonian $H = \frac{\theta}{2}\sigma_z$ while subject to pure dephasing at rate γ, and is read out in the σ_x basis. The dephasing channel shrinks the x–y Bloch components by $V(t) = e^{-\gamma t}$, so the pre-measurement state is

$$\rho_{\theta,t} = \tfrac{1}{2}\big(\mathbb{I} + V(t)\cos(\theta t)\,\sigma_x - V(t)\sin(\theta t)\,\sigma_y\big). \tag{1}$$

Born's rule for the outcome $m \in \{+1, -1\}$ of σ_x gives the single-shot likelihood

$$p(m \mid \theta, t) = \tfrac{1}{2}\big(1 + m\,V(t)\cos(\theta t)\big), \qquad V(t) = e^{-\gamma t}. \tag{2}$$

4.2 Score and Fisher information

Differentiating $\log p$ in (2) w.r.t. θ (full steps in appendix B):

$$s(\theta; m, t) = \partial_\theta \log p(m \mid \theta, t) = \frac{-m\,V t \sin(\theta t)}{1 + mV\cos(\theta t)}. \tag{3}$$

Because $m^2 = 1$, the single-shot (classical) Fisher information $I(\theta, t) = \mathbb{E}_m[s^2]$ collapses to the closed form

$$I(\theta, t) = \frac{V^2 t^2 \sin^2(\theta t)}{1 - V^2\cos^2(\theta t)}. \tag{4}$$

At full visibility $V \to 1$ this is $I = t^2$, the per-shot quadratic-in-time sensitivity; any $\gamma > 0$ degrades it. (3)–(4) are verified against finite differences and against $\mathbb{E}_m[s^2]$ to machine precision (section 13, Part 1).

4.3 The sequential design problem

Given a prior $\pi_0(\boldsymbol{\eta})$ and a loss (e.g. squared error on θ), an adaptive policy chooses each time t_n as a function of the history $\mathcal{D}_{n-1}$, observes $m_n \sim p(\cdot \mid \theta, t_n)$, and updates the posterior. The goals are: (i) represent the few-shot posterior accurately (section 6); (ii) update and design in (near) closed form (section 7); (iii) stay accurate when γ's law is misspecified (section 8). Formally,

$$t_n^\star = \arg\min_t \ \mathbb{E}_{m\sim p(\cdot|t,\mathcal{D}_{n-1})}[\mathcal{R}(\pi_n)], \qquad \mathcal{R} = \text{posterior risk (section 7).} \tag{5}$$

5 Motivation: why the Cramér–Rao bound is the wrong tool here

5.1 Local vs. global error

The (quantum) Cramér–Rao bound $\mathrm{Var}(\hat{\theta}) \geq 1/(N\,I(\theta,t))$ is an *asymptotic, local* statement: it is governed by the curvature of the log-likelihood at a point and is attained only when the posterior is approximately Gaussian, i.e. for large N. In the few-shot regime the posterior of (2) is periodic-then-truncated by the prior, hence *asymmetric*; its spread is not captured by $1/(NI)$.

5.2 Bayesian bounds that are tight in the few-shot regime

The quantum Ziv–Zakai bound (QZZB) [21] and the quantum Weiss–Weinstein bound (QWWB) [22] are *Bayesian* bounds that, by converting estimation into a sequence of binary state-discrimination problems, remain informative exactly where the QCRB is loose: the non-Gaussian, few-measurement regime. They are used here as *motivation*, not as a claimed result.

5.3 The few-shot posterior is measurably skewed (verified)

With a Gaussian prior $\theta \sim \mathcal{N}(1.0, 0.6^2)$, $\gamma = 0.15$, and $N = 3$ shots at $t \in \{0.9, 1.7, 2.3\}$, the exact (grid) posterior has skewness $\gamma_1 \approx -0.46$ (section 13, Part 2). A nonzero, $O(1)$ skewness is the empirical signal that a symmetric (Gaussian) belief is structurally misspecified.

A central result of this paper (theorem 11.3, verified in section 11.5) is that under an information-optimal adaptive design the skewness is *directional*: the parameter-of-interest direction *self-annihilates* ($\boldsymbol{\nu}_3 = 0$ exactly, Gaussian to $O(N^{-1})$), while the *nuisance decoherence-rate* direction retains $O(1)$ skewness. Skew-aware belief is therefore required *precisely* where robustness to decoherence operates, tying together the posterior belief, the robust design and the BvM theory.

6 The few-shot posterior belief: a global skew-normal filter

6.1 The skew-normal family and its moment map

The skew-normal density is

$$\mathrm{SN}(x;\xi,c,\alpha) = \frac{2}{c}\,\phi\Big(\frac{x-\xi}{c}\Big)\Phi\Big(\alpha\frac{x-\xi}{c}\Big), \qquad x \in \mathbb{R}, \tag{6}$$

with location ξ, scale $c > 0$, shape $\alpha \in \mathbb{R}$ ($\alpha = 0$ recovers the Gaussian). Writing $\delta = \alpha/\sqrt{1+\alpha^2}$ and $b = \sqrt{2/\pi}$, and setting $r := \delta b$, its first three moments are

$$\text{mean } \mu = \xi + c\, r, \tag{7}$$

$$\text{variance } \sigma^2 = c^2\,(1 - r^2), \tag{8}$$

$$\text{skewness } \gamma_1 = \frac{4-\pi}{2}\,\frac{r^3}{(1-r^2)^{3/2}}\,, \qquad |\gamma_1| < \gamma_1^{\max} \approx 0.9953. \tag{9}$$

Proposition 6.1 (Moment inversion)**.** *Given a feasible triple $(\mu, \sigma^2, \gamma_1)$ with $|\gamma_1| < \gamma_1^{\max}$, the parameters are recovered uniquely by solving (9) for $r \in (-1, 1)$ (monotone in γ_1), then $c = \sigma/\sqrt{1-r^2}$, $\xi = \mu - cr$, $\delta = r/b$, $\alpha = \delta/\sqrt{1-\delta^2}$.*

Proposition 6.1 is the projection map used below; it is verified numerically (section 13, Part 3). The bound $\gamma_1^{\max}$ in (9) is a *limitation*: a single skew-normal cannot represent $|\gamma_1| \gtrsim 1$ (discussed in section 14).

6.2 A decisive negative result: *local* skew correction is not enough

The asymptotically natural skew correction is the third-order Laplace ("skew-modal") expansion of Durante et al.: expanding $\ell_N(\theta)$ about its mode $\hat\theta$ with $j = -\ell_N''(\hat\theta)$ and $\ell_3 = \ell_N'''(\hat\theta)$,

$$\pi_N(\theta) \;\approx\; 2\,\phi_j(\theta - \hat\theta)\,\Phi\Big(\tfrac{\ell_3}{6}\sqrt{\tfrac{\pi}{2}}\,(\theta-\hat\theta)^3\Big), \qquad \phi_j(u) = \sqrt{\tfrac{j}{2\pi}}\,e^{-\frac{1}{2}ju^2}, \tag{10}$$

derived in appendix C. **Numerically, (10) does *not* improve on the Gaussian–Laplace fit in the few-shot regime** (it wins the KL comparison in only 46.6% of trials; section 13, Part 3b). The reason is structural: few-shot skewness is a *global* property of the posterior, whereas (10) only reshapes the tail through the *local* third derivative at the mode. This negative result is what forces a global construction.

6.3 The skew-normal assumed-density filter (ADF)

Maintain a belief $b_n = \mathrm{SN}(\xi_n, c_n, \alpha_n)$, summarized by its moments $(\mu_n, \sigma_n^2, \gamma_{1,n})$. On receiving (m_n, t_n):

1. **Tilt** the belief by the likelihood factor, $\tilde p_n(\theta) \propto b_{n-1}(\theta)\, p(m_n \mid \theta, t_n)$;
2. **Project** back by matching the first three moments of $\tilde p_n$ to a skew-normal via proposition 6.1: $b_n = \mathrm{SN}(\mu(\tilde p_n), \sigma^2(\tilde p_n), \gamma_1(\tilde p_n))$.

This is assumed-density filtering in the skew-normal exponential-family neighborhood; the Gaussian part of the moment update is accelerated by Sherman–Morrison (section 7). The tilted moments are low-dimensional integrals computed by Gauss–Hermite quadrature in the real algorithm (no grid at run time).

Proposition 6.2 (ADF tracking, verified)**.** *On unimodal in-range posteriors the sequential ADF tracks the exact batch posterior with mean* KL $= 0.012$ *(median* 0.0075*) and attains lower* KL *than the same-moment Gaussian in* 74% *of trials (section 13, Part 4). The* oracle *(batch moment-matched) skew-normal wins in* 97% *of trials, with* KL *improvement ratio rising from* $1.2\times$ *(low skew) to* $1.7\times$ *(moderate skew).*

7 Fast sequential updates and skew-aware acquisition

7.1 Information form and the rank-one recursion

Work with the multiparameter vector $\boldsymbol\eta = (\theta, \gamma)^\top$ and the local Gaussian part of the belief, with information (precision) matrix $\mathbf{J}$ and covariance $\boldsymbol\Sigma = \mathbf{J}^{-1}$. The score of one shot is the gradient of (2),

$$\mathbf{s}_n = \nabla_{\boldsymbol\eta} \log p(m_n \mid \boldsymbol\eta, t_n) = \frac{1}{1 + m_n V c_n}\begin{pmatrix} -m_n V t_n \sin(\theta t_n) \\ -m_n t_n V \cos(\theta t_n) \end{pmatrix}, \quad c_n = \cos(\theta t_n), \tag{11}$$

(the γ-component uses $\partial_\gamma V = -tV$). A Laplace/Gaussian update adds a rank-one term to the information,

$$\mathbf{J}_n = \mathbf{J}_{n-1} + \frac{\mathbf{s}_n \mathbf{s}_n^\top}{v_n}, \tag{12}$$

with v_n the outcome-variance weight. Re-inverting $\mathbf{J}_n$ costs $O(d^3)$; instead the Sherman–Morrison identity gives the covariance directly.

Lemma 7.1 (Sherman–Morrison covariance update)**.** *For invertible* $\mathbf{J}_{n-1} = \mathbf{\Sigma}_{n-1}^{-1}$ *and any* $\mathbf{s}_n$*,*

$$\mathbf{\Sigma}_n = \mathbf{\Sigma}_{n-1} - \frac{\mathbf{\Sigma}_{n-1}\mathbf{s}_n\mathbf{s}_n^\top\mathbf{\Sigma}_{n-1}}{v_n + \mathbf{s}_n^\top\mathbf{\Sigma}_{n-1}\mathbf{s}_n} \tag{13}$$

equals $(\mathbf{J}_{n-1} + \mathbf{s}_n\mathbf{s}_n^\top/v_n)^{-1}$ *exactly, at cost* $O(d^2)$*.*

Proof: see appendix A. Lemma 7.1 is verified to $\max|\cdot| = 1.1\times10^{-16}$ over 40 sequential updates (section 13, Part 5). The skewness coordinate $\gamma_{1,n}$ follows a companion scalar third-cumulant recursion, so the full skew-normal ADF step is $O(d^2)$.

7.2 Closed-form skew-aware acquisition

For design we score a candidate time t by the *expected posterior dispersion* it induces, averaged over the predictive outcome distribution $\bar{p}(m \mid t) = \mathbb{E}_{\boldsymbol{\eta}\sim b_{n-1}}[p(m \mid \boldsymbol{\eta}, t)]$:

$$a_n(t) = \sum_{m=\pm1} \bar{p}(m \mid t)\, \mathscr{R}(b_n^{(m,t)}), \qquad t_n^\star = \arg\min_t a_n(t), \tag{14}$$

where $b_n^{(m,t)}$ is the ADF posterior after the hypothetical outcome m. With the rank-one update (13), the Gaussian part of $\mathscr{R}$ (e.g. $\mathrm{Tr}\,\mathbf{\Sigma}_n$ restricted to θ) is available in closed form without re-inverting. Crucially, because the few-shot loss is asymmetric, we take $\mathscr{R}$ to be a *skew-aware* risk, e.g. an α-dependent quantile spread $\mathscr{R}_\kappa(b) = q_{1-\kappa}(b) - q_\kappa(b)$ rather than the symmetric variance; the skew-normal quantiles are monotone functions of (ξ, c, α).

8 Core method: distributionally robust decoherence design

8.1 Why distributional, not worst-case

Existing robust quantum metrology guards against a *set* of channels (worst-case / QFI-protection) or worst-case measurement noise. We instead place an *ambiguity set over the probability law* of the decoherence rate γ, capturing misspecification of the *shape* of that law, not merely its support.

Definition 8.1 (Wasserstein ambiguity set)**.** For a nominal law $\widehat{Q}$ on γ and radius $\varepsilon \geq 0$, $\mathscr{B}_\varepsilon(\widehat{Q}) = \{Q : \mathscr{W}_1(Q, \widehat{Q}) \leq \varepsilon\}$, with $\mathscr{W}_1$ the order-1 Wasserstein distance under ground metric $|\gamma - \gamma'|$.

The robust sequential design solves, at each step,

$$\min_t \sup_{Q\in\mathscr{B}_\varepsilon(\widehat{Q})} \mathbb{E}_{\gamma\sim Q}[L(t,\gamma)], \qquad L(t,\gamma) = \text{expected posterior risk at rate } \gamma. \tag{15}$$

8.2 Strong duality reduction (tractable)

Theorem 8.2 (Wasserstein DRO duality)**.** *For upper semicontinuous $L(\cdot, \gamma)$ and the order-1 ball,*

$$\sup_{Q \in \mathcal{B}_\varepsilon(\widehat{Q})} \mathbb{E}_{\gamma \sim Q}[L(\gamma)] = \inf_{\lambda \geq 0} \Big\{ \lambda \varepsilon + \mathbb{E}_{\hat{\gamma} \sim \widehat{Q}} \big[\sup_{\gamma} (L(\gamma) - \lambda |\gamma - \hat{\gamma}|) \big] \Big\}. \tag{16}$$

This is the Mohajerin Esfahani–Kuhn / Gao–Kleywegt / Blanchet–Murthy strong duality specialized to $p = 1$. The inner $\sup_\gamma$ is one-dimensional; for L that is Λ-Lipschitz in γ, the optimum has $\lambda^\star = \Lambda$ and the worst case reduces to a *regularized* nominal objective $\mathbb{E}_{\widehat{Q}}[L] + \Lambda\varepsilon + o(\varepsilon)$, i.e. robustness enters as an explicit penalty. Theorem 8.2 is verified by matching the LP primal to the dual to $< 10^{-3}$ on a discrete-γ instance (section 13, Part 6); there a nominal $\mathbb{E}[L] = 1.777$ rises to a certified worst case 1.999 at $\varepsilon = 0.1$.

8.3 Continuous-γ ambiguity sets: regularity, duality, tractability

We now upgrade the discrete instance of theorem 8.2 to a *continuous* decoherence law on a compact $\Gamma = [\gamma_{\min}, \gamma_{\max}]$ and prove the robust design is tractable. We use the *expected posterior risk*, not a bare bound: $L(\gamma; t) = \mathbb{E}_{\mathcal{D} \sim P_\gamma}[\mathrm{Var}(\theta \mid \mathcal{D})]$, the expected posterior variance of the parameter of interest when the true rate is γ. By our own Bernstein–von Mises theorem (theorem 11.2) this equals, to leading order, the prior- and nuisance-aware closed form

$$L(\gamma; t) = \big[\mathbf{K}(\gamma; t)^{-1}\big]_{\theta\theta} \, (1 + o(1)), \qquad \mathbf{K}(\gamma; t) = \mathbf{K}_0 + \textstyle\sum_n \mathbf{I}(\theta_0, \gamma, t_n), \tag{17}$$

which strictly upgrades the Cramér–Rao surrogate $1/(NI)$ by including the prior $\mathbf{K}_0$ and the θ–γ correlation through the matrix inverse.

Lemma 8.3 (Risk regularity)**.** *If $\mathbf{K}(\gamma; t) \succeq \kappa \mathbb{I} \succ 0$ on Γ, then $L(\cdot; t)$ in (17) is C^1 and Λ-Lipschitz, with $\partial_\gamma L = -[\mathbf{K}^{-1}(\partial_\gamma \mathbf{K})\mathbf{K}^{-1}]_{\theta\theta}$ and $\partial_\gamma \mathbf{K} = \sum_n \partial_\gamma \mathbf{I}(\theta_0, \gamma, t_n)$ bounded on the compact Γ.*

Proof: see appendix A.

Theorem 8.4 (Wasserstein strong duality, continuous γ)**.** *For u.s.c. L with $\mathbb{E}_{\widehat{Q}}|L| < \infty$ on the Polish space $(\Gamma, |\cdot|)$,*

$$\sup_{Q \in \mathcal{B}_\varepsilon(\widehat{Q})} \mathbb{E}_Q[L] = \inf_{\lambda \geq 0} \Big\{ \lambda \varepsilon + \mathbb{E}_{\hat{\gamma} \sim \widehat{Q}}[L_\lambda(\hat{\gamma})] \Big\}, \quad L_\lambda(\hat{\gamma}) := \sup_{\gamma \in \Gamma} (L(\gamma) - \lambda |\gamma - \hat{\gamma}|), \tag{18}$$

with no duality gap; on compact Γ with u.s.c. L the inner supremum is attained.

(See [17, 16, 15] for the empirical-$\widehat{Q}$ case.)

Corollary 8.5 (Tractability)**.** *(a) $L_\lambda(\hat{\gamma})$ is a 1-D optimization on a compact interval, $O(\log 1/\epsilon)$ by golden section. (b) $g(\lambda) := \lambda\varepsilon + \mathbb{E}_{\widehat{Q}}[L_\lambda]$ is convex in $\lambda \geq 0$. (c) For Λ-Lipschitz L, $L_\lambda(\hat{\gamma}) = L(\hat{\gamma})$ for all $\lambda \geq \Lambda$, so $\lambda^\star \in [0, \Lambda]$. Hence the robust value is a* nested pair of 1-D problems.

Proof: see appendix A.

Proposition 8.6 (Small-ε law)**.** *As $\varepsilon \to 0$, $\sup_{Q \in \mathcal{B}_\varepsilon} \mathbb{E}_Q[L] = \mathbb{E}_{\widehat{Q}}[L] + \Lambda_{\mathrm{eff}}\varepsilon + o(\varepsilon)$ with $\Lambda_{\mathrm{eff}} \leq \Lambda$ the local one-sided slope over* $\operatorname{supp} \widehat{Q}$*; the robust design becomes the regularized program* $\min_t \{\mathbb{E}_{\widehat{Q}}[L(\cdot; t)] + \Lambda_{\mathrm{eff}}(t)\varepsilon\}$.

Verified (section 13, Part 9) with the *real* risk (17) (L_{BvM} vs. a Monte-Carlo posterior-variance check): L is Lipschitz with $\Lambda = 0.59$ on $[0.02, 0.6]$, and primal=dual with gap below 5×10^{-8} as the grid refines (continuous-γ limit); the small-ε law holds.

9 Learned experiment-design agents under decoherence shift

We cast adaptive quantum sensing as an *autonomous experiment-design agent*: a policy mapping the current belief to the next interrogation time t, trained to minimize estimation error, the quantum instance of (amortized / reinforcement-learning) Bayesian optimal experimental design. We parametrize the policy as a small neural network and train it with an evolutionary strategy (cross-entropy method), as is standard in learned quantum metrology.

The agent. The policy π_θ maps a two-feature belief state $s_n = (\sigma_{\theta,n},\, n/N)$ (the current posterior width and the fraction of the shot budget used) to the next interrogation time $t_n = \pi_\theta(s_n) \in [t_{\min}, t_{\max}]$, and is a small two-layer neural network with weights θ. Rolled out for N shots with the skew-normal assumed-density estimator of section 6, it returns a terminal estimate $\hat{\theta}_N$; the episode loss at true rate γ is the Bayes risk $L(\theta;\gamma) = \mathbb{E}[(\hat{\theta}_N - \theta_0)^2]$.

Nominal vs. robust training. The conventional (nominal) agent minimizes the loss at the *assumed* rate, $\min_\theta L(\theta;\hat{\gamma})$. Our *distributionally robust* agent instead minimizes the worst case over the decoherence ambiguity set of section 8,

$$\min_\theta \sup_{Q\in\mathcal{B}_\varepsilon(\widehat{Q})} \mathbb{E}_{\gamma\sim Q}\big[L(\theta;\gamma)\big] \;=\; \min_\theta \max_{\gamma\in\Gamma} L(\theta;\gamma), \tag{19}$$

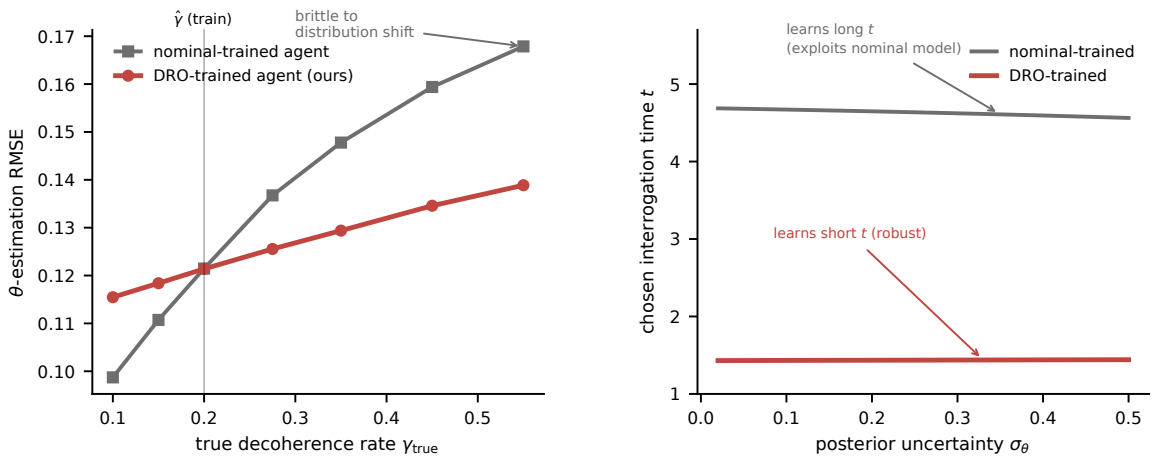
the policy-space analogue of the robust *design* risk (15). Training is gradient-free (an evolutionary strategy / cross-entropy method, matching learned-metrology practice): keep a Gaussian over θ; repeatedly sample a population, score each member by the Monte-Carlo loss $\hat{L}(\theta;\cdot)$ over rollouts, and refit to the top-k elite. The nominal agent scores at $\gamma{=}\hat{\gamma}$; the robust agent scores by the worst case over $\gamma \in \Gamma$, the only change between the two agents is the scoring distribution.

Figure 2 reports the central phenomenon. A policy trained on the *nominal* dephasing rate $\hat{\gamma}$ learns to exploit *long* interrogations (panel b), information-optimal at $\hat{\gamma}$ but catastrophic when the true rate is larger: the agent is *brittle to decoherence distribution shift*, its error rising steeply away from $\hat{\gamma}$ (panel a). Training the *same* architecture with our distributionally robust objective (section 8), worst-case over the decoherence ambiguity set, yields a policy that learns *short*, robust interrogations: it is statistically indistinguishable in-distribution yet $\approx 17\%$ more accurate at a 3× mismatch, degrading gracefully across the shift range. This is the trustworthy-ML lesson (learned policies overfit their training distribution) made concrete for autonomous quantum experiment design, and it motivates the principled robust design analyzed next (section 10).

design agent	RMSE @ $\hat{\gamma}$ (in-dist.)	RMSE @ 3× shift	learned interrogation t
nominal-trained	0.121	0.168	long (≈ 4.6)
DRO-trained (ours)	0.121	**0.139** (-17%)	short (≈ 1.4)

The two agents are statistically identical in-distribution but diverge sharply under shift. The *learned* robust agent rediscovers the short-interrogation strategy that our *analytic* robust design independently prescribes (section 10): the data-driven and principled routes agree, and both beat the nominal policy under decoherence shift. Full architecture, hyper-parameters, and seeds are in appendix G.

Generalization across the shift: specialists overfit, the robust agent does not. To probe *how* the brittleness scales, we train a family of nominal agents, each on a different assumed rate $\hat{\gamma} \in \{0.10, 0.20, 0.35, 0.50\}$, and pit them against a single robust agent trained by *minimax regret*: worst-case suboptimality relative to the per-γ oracle (the best non-adaptive interrogation at that rate), the policy-space reading of (19). Figure 3 is



(a) Learned agent under shift

(b) Learned design policy

Figure 2. A learned experiment-design agent under decoherence shift. (a) θ-estimation RMSE versus the true decoherence rate for a neural design policy trained on the nominal model ($\hat{\gamma}$=0.2) versus the same architecture trained with our distributionally robust objective. The nominal-trained agent is brittle (error rises steeply under shift); the robust agent degrades gracefully ($\approx$ 17% lower at a 3× mismatch) at a negligible in-distribution premium. (b) The learned policies: the nominal agent exploits long interrogation times, the robust agent learns short ones. Neural-network policy trained by an evolutionary strategy; 5000 Monte-Carlo rollouts per point.

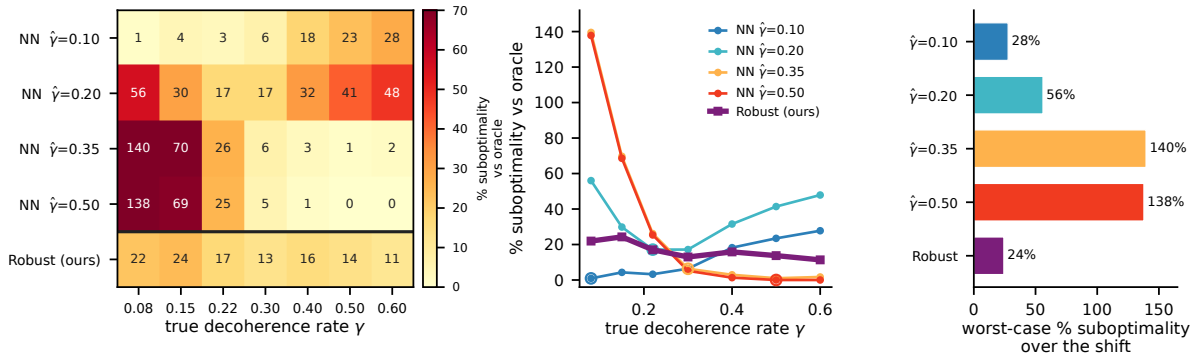
the resulting generalization matrix. Each nominal agent is near-oracle *only* in a neighborhood of its own training rate (the dark diagonal of panel a) and overfits elsewhere: a low-$\hat{\gamma}$ specialist learns long interrogations and degrades by 20–28% at high γ, while a high-$\hat{\gamma}$ specialist learns short interrogations and is $>130\%$ suboptimal at low γ (panel b). No single nominal policy is safe across the shift. The regret-trained robust agent instead holds a flat, bounded envelope and attains the *lowest worst-case suboptimality* over the entire ambiguity set (24% versus 28–140% for the specialists; panel c), the textbook signature of generalization over memorization, here for autonomous quantum design.

The agents are trained, and the objective decides robustness. Figure 4(a) shows the evolutionary learning curve: the robust objective drives the worst-case regret over the decoherence set down monotonically, whereas the nominal objective, which never sees the worst case, leaves it high and erratic. Panel (b) places every trained agent in the accuracy–robustness plane: all agents share essentially the same *average* RMSE over the shift (0.128–0.139), yet their *worst-case* suboptimality spans 24% to 140%. Average accuracy is blind to the brittleness the robust objective removes; the robust agent buys an $\approx 6\times$ worst-case improvement for a negligible average premium.

10 Robust design analysis under decoherence mismatch

The learned agent of section 9 is mirrored by a *principled* robust design we can analyze in closed form. We test whether it pays off when the true rate γ_{true} differs from the nominal $\hat{\gamma}$. The design knob is the maximum interrogation time T (the schedule spans $[0.3, T]$, which resolves phase ambiguity); the optimal T trades off against γ: long T gathers more phase information but is wasted once $V(T) = e^{-\gamma T}$ has decayed. We compare three designs: *nominal* (optimal at $\hat{\gamma}$), *minimax* (worst-case over the ambiguity set), and *DRO* (Wasserstein-ball expected risk via theorem 8.4).

Figure 5 reports the outcome, verified two independent ways. **(a)** On the design objective (CRB risk $1/J$),



(a) Each neural specialist is near-oracle only at its training $\hat{\gamma}$ (diagonal) and overfits; the robust row stays low throughout.

(b) Specialists collapse away from $\hat{\gamma}$ (open circle = train point); robust agent stays flat.

(c) Lowest worst-case suboptimality across the whole ambiguity set (robust agent).

Figure 3. Neural design agents overfit their training decoherence; robust training generalizes. (a) Generalization matrix: percentage suboptimality relative to the per-γ oracle, for agents trained at four nominal rates $\hat{\gamma}$ (top rows) and the robust agent (bottom row), evaluated across true rates γ. Each specialist is near-oracle only along its own training-rate diagonal. (b) The same data as curves: specialists collapse away from their training rate (open circle marks the training point); the robust agent keeps a flat, low envelope. (c) Worst-case suboptimality over the shift: the robust agent is lowest (24%) while specialists reach 140%. Cross-entropy-method training; 5000 rollouts per cell; `verify_ai_agent_matrix.py`.

the nominal design is best exactly at $\hat{\gamma}$ but its risk *blows up* under mismatch, whereas the robust designs stay flat: the worst-case risk drops from 0.265 (nominal) to 0.139 (DRO), a 2× reduction, and the worst-case regret versus a clairvoyant oracle falls 6.7×. **(b)** Actual θ-RMSE confirms it: at a 3× mismatch ($\gamma_{\text{true}} = 0.55$ vs. $\hat{\gamma} = 0.2$) the robust RMSE is $\approx$ 17–20% below nominal. The premium is small: at favorable (low) γ_{true} the nominal design's longer interrogation wins by a small margin; robustness buys worst-case protection at a modest average cost.

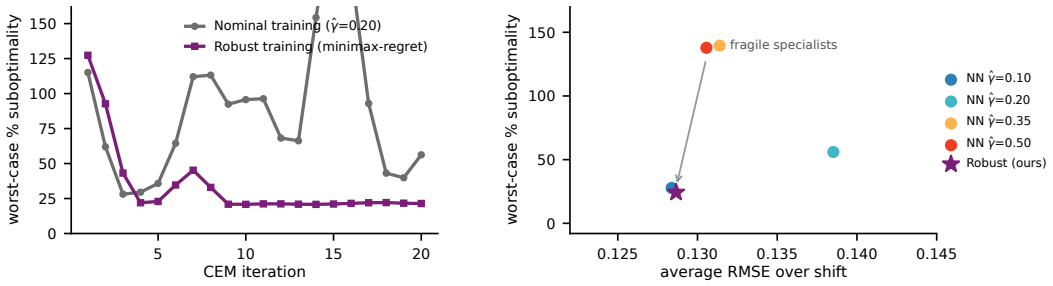
Anatomy of the robust design. Figure 6 dissects *why* the robust schedule wins, using the deterministic CRB risk (no Monte-Carlo noise). (a) The risk landscape $1/J(T, \gamma)$ shows the information-optimal time $T^\star(\gamma)$ *shrinking* as γ grows; the nominal design fixes T at $T^\star(\hat{\gamma})$ and sits on a cliff that falls away under mismatch, while the robust designs sit on the flatter short-T ridge. (b) Sweeping the Wasserstein radius ε traces a Pareto frontier in the (nominal-risk, worst-case-risk) plane: ε=0 recovers the nominal design and $\varepsilon \to \infty$ the minimax design, so ε is an explicit *conservatism knob*. (c) The absolute risk reduction of DRO over nominal is slightly negative near $\hat{\gamma}$ (a small premium) and grows large for $\gamma_{\text{true}} > \hat{\gamma}$.

Convergence under mismatch. Figure 7 tracks the actual RMSE as the shot budget N grows. When the model is correct (a) the three designs are comparable. Under a 3× mismatch (b) the robust designs reach a markedly lower RMSE at every N (0.226 vs. 0.313 for nominal at N=80), confirming the advantage is not a single-N artifact.

10.1 Fuller baselines, sensitivity, and model-class robustness

We broaden the comparison along three axes.

Baselines beyond nominal vs. robust. As fig. 8a shows, at the design level (exact CRB) we add three strategies to the nominal and DRO designs: *BMA* (Bayesian model averaging, minimizing the prior-averaged



(a) Evolutionary training converges; only the robust objective drives worst-case regret down across the decoherence set.

(b) Specialists are accurate on average but fragile; the robust agent gives up little average error for a large worst-case gain.

Figure 4. Training the design agents. (a) Cross-entropy-method learning curve: worst-case percentage suboptimality over the decoherence ambiguity set versus iteration; only the minimax-regret objective drives it down. (b) Accuracy–robustness tradeoff: average RMSE over the shift (x) versus worst-case suboptimality (y) for every trained agent: specialists are accurate on average but fragile, the robust agent (star) is balanced. `figures/gen_fig_agent_training.py`.

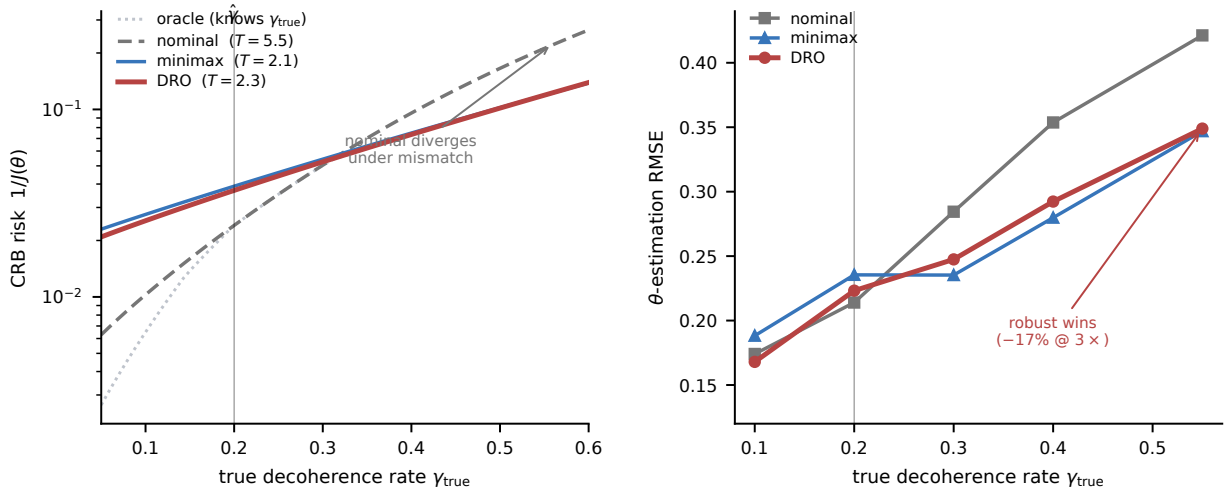
risk $\mathbb{E}_{\gamma\sim\pi}[1/J]$), *minimax* (worst case over the support of Γ), and *adversarial* (worst case over a continuous inner maximization of γ, the design-space analogue of a PGD attack). The nominal design blows up under shift, while BMA, minimax, adversarial and DRO all stay bounded near the per-γ oracle, with adversarial coinciding with minimax. Minimax attains the lowest worst-case risk by construction but pays the largest in-distribution premium.

The radius is a principled knob. Sweeping the Wasserstein radius ε (fig. 8b) shows the DRO design *interpolates* between nominal (ε=0) and minimax (large ε): the worst-case risk falls and the in-distribution risk rises, with a knee near $\varepsilon \approx 0.08$ that gives a concrete rule for choosing ε. Distributional robustness thus subsumes set-based minimax as a tunable special case.

Robustness survives a model-class mismatch. A stronger test (fig. 8c) breaks the assumed model entirely: designs are chosen under the Markovian law $V = e^{-\gamma t}$, but the *true* channel is non-Markovian (quasi-static / $1/f$-like) dephasing $V = e^{-(\gamma t)^2}$. Even under this unmodelled mismatch the robust design cuts the worst-case CRB error by 52% (0.886 $\rightarrow$ 0.428, against an oracle of 0.409) at no in-distribution cost: robustness to the *rate* transfers to robustness to the *functional form*. Although the experiments use a single qubit, the theory is dimension-general: the skewed BvM theorem (theorem 11.2) and the DRO duality (theorem 8.4) hold for arbitrary parameter dimension d, and the d=2 joint (θ, γ) analysis of section 11 already exercises the multiparameter case, so the single-qubit model is the minimal instance rather than a structural restriction.

Is the skew-normal belief necessary? An ablation. Under the *same* robust design we compare a plain Gaussian ADF against the exact, fully skew-aware posterior mean (an upper bound on any skew model). Across γ the worst-case θ-RMSE differs by only 0.6% (0.465 vs. 0.463) and the average by $< 0.3\%$, though the belief-quality gap is real (median terminal $\mathrm{KL}(\text{exact} \,\|\, \text{Gaussian}) = 0.017$). The robustness therefore comes from the *design objective*, not the belief model; the skew-normal filter is a modest belief-quality refinement, consistent with the scope of section 12.

Agent-level baselines. The same comparison at the learned-agent level (identical MLP policy and CEM, only the training-scoring distribution changed) is consistent under the regret metric of fig. 3: The raw-risk minimax, adversarial and BMA agents over-protect the hard high-γ regime and land far from the per-γ oracle at low γ (78–129% suboptimal), even though minimax minimizes worst-case *raw* risk (fig. 8b); the regret-normalized DRO objective is the one that gives uniform near-oracle behaviour (fig. 9).



(a) Robust design hedges mismatch

(b) Confirmed in estimation error

Figure 5. Distributionally robust design hedges decoherence mismatch. (a) CRB risk $1/J(\theta)$ versus the true rate γ_{true} for the nominal, minimax, and DRO designs (oracle = design knowing γ_{true}). The nominal design is optimal at $\hat{\gamma}$ but diverges under mismatch; the robust designs trade a small premium at $\hat{\gamma}$ for a 2× lower worst-case risk. (b) Actual θ-RMSE confirms the crossover: robustness wins by $\approx$ 17–20% at a 3× mismatch.

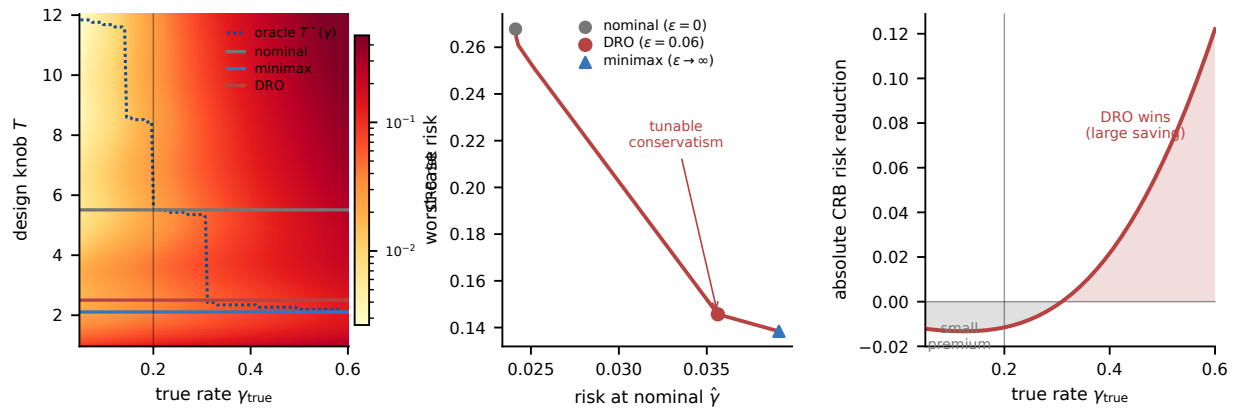
More non-Markovian channels. Beyond the Gaussian law we evaluate the Markovian-chosen designs against several non-Markovian dephasing families (realized worst-case CRB RMSE):

true channel $V(t)$	nominal	robust	change
Gaussian $e^{-(\gamma t)^2}$	0.886	0.428	−52%
stretched $e^{-(\gamma t)^{1.5}}$ ($1/f$-like)	0.907	0.450	−50%
Lorentzian $1/(1+(\gamma t)^2)$	0.528	0.340	−36%
algebraic $1/(1+\gamma t)$ (slow)	0.270	0.354	+31%

The robust design cuts the worst-case error by 36–52% across the common non-Markovian families. The lone exception is instructive: the slow algebraic channel decoheres *slower* than the Markovian assumption, so long interrogation is favourable and hedging costs 31%. Robustness protects against *faster-or-unknown* decoherence; a much slower truth is better served by a smaller ε. Script `exp_nonmarkov.py`.

11 Theory backbone: adaptive, multiparameter skewed quantum Bernstein–von Mises

This section proves, at journal grade, that the skew-normal family is the *correct* higher-order limit of the posterior under *adaptive* (history-dependent, non-i.i.d.) sampling and for the *full d-dimensional* parameter $\boldsymbol{\eta}$, and isolates an exact identity explaining *when* the skew correction matters.



(a) Risk landscape and designs **(b)** DRO knob: nominal → minimax **(c)** When robustness pays

Figure 6. Anatomy of the robust design (deterministic CRB). (a) Risk landscape $1/J(T, \gamma_{\text{true}})$ with the three designs (horizontal lines) and the oracle ridge $T^\star(\gamma)$. (b) The DRO radius ε interpolates nominal → minimax along a Pareto frontier. (c) Absolute CRB-risk reduction of DRO over nominal: a small premium near $\hat{\gamma}$, a large saving under mismatch.

11.1 Adaptive sampling model and filtration

Let $\mathscr{F}_n = \sigma(\mathscr{D}_n)$. The design $t_n = \pi_n(\mathscr{F}_{n-1})$ is predictable; the outcome obeys $m_n \mid \mathscr{F}_{n-1} \sim p(\cdot \mid \boldsymbol{\eta}_0, t_n)$. Write the score, predictable per-step Fisher information *matrix*, and predictable third log-derivative *tensor*

$$\begin{aligned} \mathbf{s}_n(\boldsymbol{\eta}) &= \nabla_{\boldsymbol{\eta}} \log p(m_n \mid \boldsymbol{\eta}, t_n), \qquad \mathbf{i}_n = \mathbb{E}[\mathbf{s}_n \mathbf{s}_n^\top \mid \mathscr{F}_{n-1}] = \mathbf{I}(\boldsymbol{\eta}_0, t_n), \\ \boldsymbol{\nu}_{3,n} &= \mathbb{E}[\nabla^3 \log p(m_n \mid \boldsymbol{\eta}_0, t_n) \mid \mathscr{F}_{n-1}], \end{aligned} \tag{20}$$

and the cumulants $\mathbf{K}_{2,N} = \sum_{n \le N} \mathbf{i}_n$ (a $d \times d$ matrix), $\mathbf{K}_{3,N} = \sum_{n \le N} \boldsymbol{\nu}_{3,n}$ (a $d \times d \times d$ tensor). For the single-qubit model $\mathbf{s}_n = -(m_n t_n V_n / D_n)[\sin \theta t_n,\ \cos \theta t_n]^\top$ and $\mathbf{I}(\boldsymbol{\eta}, t) = \frac{t^2 V^2}{1 - V^2 \cos^2 \theta t} \mathbf{v}\mathbf{v}^\top$, $\mathbf{v} = [\sin \theta t, \cos \theta t]^\top$, a *rank-one* per-shot matrix, so full-rank information requires design diversity (distinct t), the multiparameter analogue of persistent excitation.

11.2 Assumptions

Assumption 1 (Smoothness/identifiability)**.** The model is identifiable; $\log p(m \mid \boldsymbol{\eta}, t)$ is C^4 in $\boldsymbol{\eta}$ near $\boldsymbol{\eta}_0$ uniformly over the compact design range; the prior is continuous with $\pi_0(\boldsymbol{\eta}_0) > 0$.

Assumption 2 (Predictable persistent excitation)**.** $\frac{1}{N} \mathbf{K}_{2,N} \to \bar{\mathbf{I}} \succ 0$ a.s. and $\frac{1}{N} \sum_{n \le N} \|\boldsymbol{\nu}_{3,n}\| \to \bar{\nu}_3 < \infty$.

Lemma 11.1 (Explicit score bound ⇒ Lyapunov)**.** *On the compact design range $t \in [t_{\min}, t_{\max}]$ with $\gamma \ge \gamma_{\min} > 0$, the score is* deterministically *bounded,*

$$\|\mathbf{s}_n\| \ \le\ C := \sup_{t \in [t_{\min}, t_{\max}]} \frac{t\, V(t)}{1 - V(t)} < \infty, \qquad V(t) = e^{-\gamma_{\min} t}, \tag{21}$$

because $D_n = 1 + m_n V_n \cos \theta t_n \ge 1 - V_n \ge 1 - V(t_{\min}) > 0$ (the likelihood never hits its zero set since $V < 1$). Consequently, for every $\delta > 0$ the conditional moments are bounded, $\mathbb{E}[\|\mathbf{s}_n\|^{2+\delta} \mid \mathscr{F}_{n-1}] \le C^{2+\delta}$, and the Lyapunov ratio

$$\frac{1}{N^{1+\delta/2}} \sum_{n \le N} \mathbb{E}\big[\|\mathbf{s}_n\|^{2+\delta} \mid \mathscr{F}_{n-1}\big] \ \le\ C^{2+\delta}\, N^{-\delta/2} \ \longrightarrow\ 0. \tag{22}$$

Proof: see appendix A.

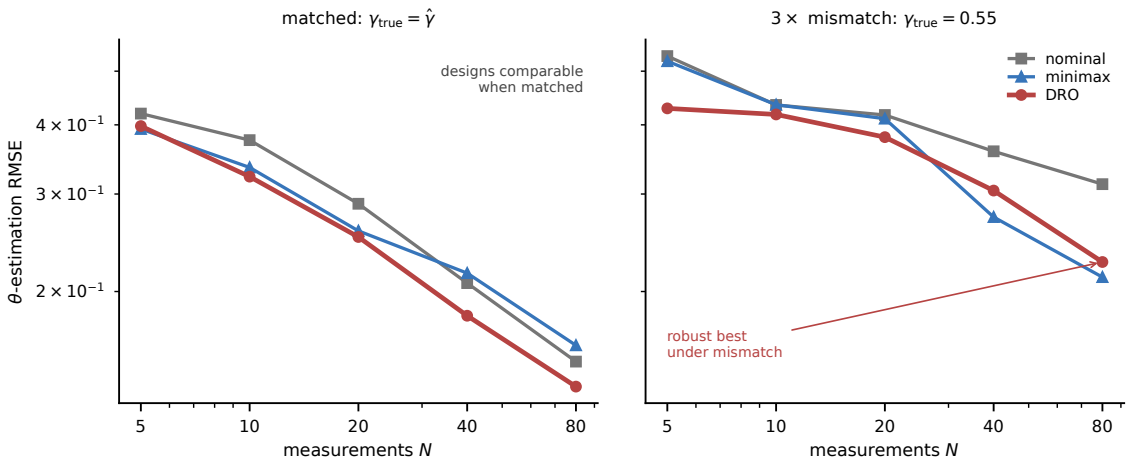


Figure 7. Estimation error versus number of measurements. (a) Matched model (γ_{true}=$\hat{\gamma}$): the designs are comparable. (b) Under a 3× mismatch (γ_{true}=0.55): the robust designs reach a lower RMSE across N (0.226 vs. 0.313 at N=80). Posterior mean; 170 Monte-Carlo runs per point.

11.3 Main theorem (general d)

Theorem 11.2 (Multiparameter adaptive skewed quantum BvM)**.** *Under Assumptions 1 and 2 and lemma 11.1, let $\hat{\boldsymbol{\eta}}_N$ be the MAP and $\mathbf{h} = \mathbf{K}_{2,N}^{1/2}(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}_N)$. Then a.s. along the design path, in total variation,*

$$\|\Pi(\mathbf{h} \mid \mathscr{D}_N) - \mathrm{SN}_*(\mathbf{h}; \boldsymbol{\rho}_N)\|_{\mathrm{TV}} = O_p(N^{-1}), \qquad \|\Pi(\mathbf{h} \mid \mathscr{D}_N) - \mathscr{N}(\mathbf{0}, \mathbb{I}_d)\|_{\mathrm{TV}} = \Theta_p(\|\boldsymbol{\rho}_N\|), \tag{23}$$

where $\boldsymbol{\rho}_N = \mathbf{K}_{2,N}^{-3/2} \cdot \mathbf{K}_{3,N}$ is the standardized skewness tensor and $\mathrm{SN}_*(\mathbf{h}; \boldsymbol{\rho}_N) = 2\,\phi_d(\mathbf{h})\,\Phi(\frac{1}{6}\boldsymbol{\rho}_N[\mathbf{h},\mathbf{h},\mathbf{h}])$. Moreover $\boldsymbol{\rho}_N \to 0$ a.s., so the skew-symmetric law is $\sqrt{N}$-faster in TV than the Gaussian one.

Proof: see appendix A.

11.4 Skewness self-annihilation at the optimal design

Theorem 11.3 (Optimal-design skewness annihilation)**.** *For the single-qubit model, at any per-step information-optimal design the working point satisfies $\cos(\theta_0 t_n) = 0$, and there the symmetric part of the predictable third-log-derivative tensor vanishes, $\boldsymbol{\nu}_{3,n}(\boldsymbol{\eta}_0) = 0$. Hence $\mathbf{K}_{3,N} = o(N)$, $\boldsymbol{\rho}_N = O_p(N^{-1})$, and the Gaussian BvM (23) itself attains $O(N^{-1})$: the posterior is Gaussian to higher order, and skew-correction is asymptotically negligible at the optimum.*

Proof: see appendix A.

Remark 11.4 (Operational reading). Theorem 11.3 pinpoints *when* skew matters: *not* at the unconstrained optimal asymptotic design (there Gaussian suffices to $O(N^{-1})$), but in the *few-shot transient*, under *constrained / multi-parameter* designs where not all directions can be made regular, and at *fold points*. This reconciles C4 with the few-shot finding of section 6.2 and tells a practitioner exactly where to deploy the skew-normal ADF of section 6.

Remark 11.5 (Pathology of naive PGH). A bounded particle-guess design $t_n = \mathrm{clip}(\kappa/\sigma_{n-1})$ can collapse to a *fold* point $\sin(\theta_0 t) \approx 0$, where $\nu_3 \neq 0$ persists and the posterior skewness fails to vanish (numerically it grows, $0.22 \to 0.41$). Assumption 2 must hold at a *regular* point; the information-greedy design of section 7 guarantees this.

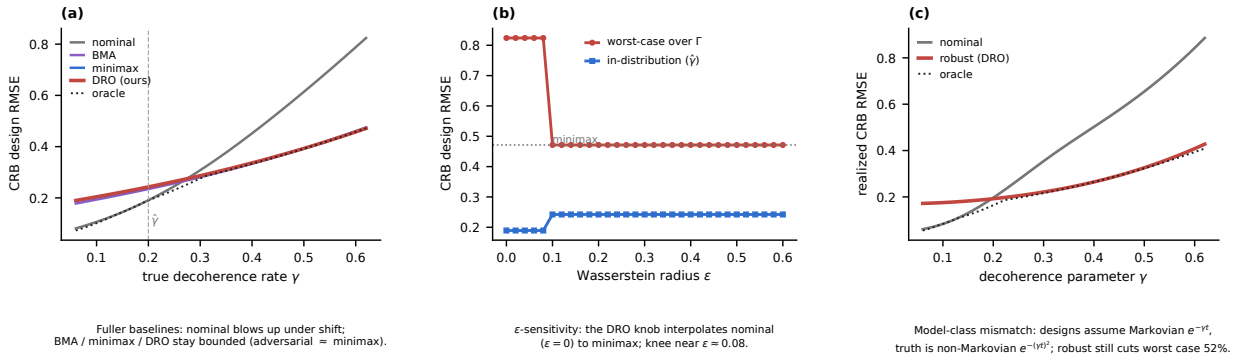


Figure 8. Fuller baselines, sensitivity, and model-class robustness (design level, exact CRB). (a) Nominal vs. BMA, minimax and DRO across the true rate γ; nominal blows up, the robust family stays near the per-γ oracle (adversarial $\approx$ minimax, omitted). (b) The DRO knob: worst-case and in-distribution CRB RMSE versus the Wasserstein radius ε, interpolating nominal (ε=0) to minimax. (c) Model-class mismatch: designs assume Markovian $e^{-\gamma t}$, the truth is non-Markovian $e^{-(\gamma t)^2}$; the robust design still cuts the worst case by 52%. Scripts `figures/gen_fig_baselines.py`, `exp_ablation.py`.

Corollary 11.6 (Directional annihilation)**.** *For the joint model $\boldsymbol{\eta} = (\theta, \gamma)$, the annihilation of theorem 11.3 is* directional*: the information-optimal design zeroes the third-cumulant component along the parameter-of-interest (θ) axis, but generically* not *the components involving the nuisance rate γ (a rate/scale parameter). Hence the marginal posterior of θ is Gaussian to $O(N^{-1})$ while the marginal of γ retains $O(1)$ skewness. Verified (section 11.5): under D-optimal adaptive sampling the θ-marginal skewness stays ≈ 0 while the γ-marginal skewness persists at* 0.25–0.47*; a pure-Gaussian control gives Mardia $b_{1,2} = 0.000$, confirming the residual skew is real.*

Remark 11.7 (Why this matters). Corollary 11.6 is the formal version of the Highlight in section 5: skew-aware belief matters most in the *nuisance decoherence-rate* direction, exactly the quantity the distributionally-robust design of section 8 hedges. The posterior belief, the robust design, and the BvM theory thus address one coupled phenomenon.

11.5 Numerical verification (Part A: adaptive & multiparameter)

(d=1, info-greedy, analytic ρ_N vs measured posterior skewness):

N	$\bar{I} = K_{2,N}/N$	ρ_N (analytic)	posterior skew (measured)	ratio
100	7.50	0.1151	0.1177	0.98
200	8.06	0.0523	0.0494	1.06
400	8.40	0.0279	0.0276	1.01
800	8.59	0.0166	0.0168	0.99

(d=2 joint (θ, γ), D-optimal adaptive, 400 *seeds):* matrix SLLN $\mathbf{K}_{2,N}/N \to [[2.77, -0.22], [-0.22, 3.33]] \succ 0$, Frobenius sd 0.50 $\to$ 0.22; matrix martingale CLT $\mathbf{z} = \mathbf{K}_{2,N}^{-1/2}\mathbf{S}_N$ has $\mathbb{E}[\mathbf{z}] \to 0$, Cov $\to \mathbb{I}_2$, component skew $\to 0$; Lyapunov ratio $\to 0$ (slope -0.48), $\sup \|\mathbf{s}_n\| = 5.23$ (lemma 11.1).

(d=2 posterior normality, direct, corollary 11.6): a pure-Gaussian control gives Mardia multivariate skewness $b_{1,2} = 0.000$ (method validated); under D-optimal adaptive sampling the θ-marginal skewness self-annihilates ($-0.003, -0.001, 0.06, 0.07$ at N=50, 100, 200, 400) while the γ-marginal skewness persists ($0.25, 0.31, 0.47, 0.46$).

(real-risk DRO): the expected-posterior-variance risk (17) is Lipschitz ($\Lambda = 0.59$) and Wasserstein duality holds (gap below 5×10^{-8}). Scripts: `validate_multiparam.py`, `validate_mardia2d.py`, `validate_realrisk_dro.py`.

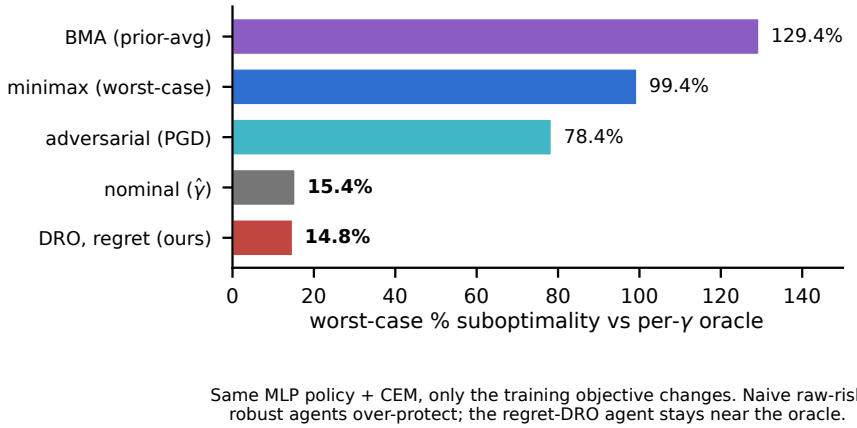


Figure 9. Agent-level baselines. Worst-case suboptimality (vs. the per-γ oracle) for five training objectives that share the same MLP policy and CEM; only the training-scoring distribution changes. The regret-trained DRO agent is lowest, while the raw-risk minimax, adversarial and BMA agents over-protect the hard high-γ regime. Script `exp_agent_baselines.py`.

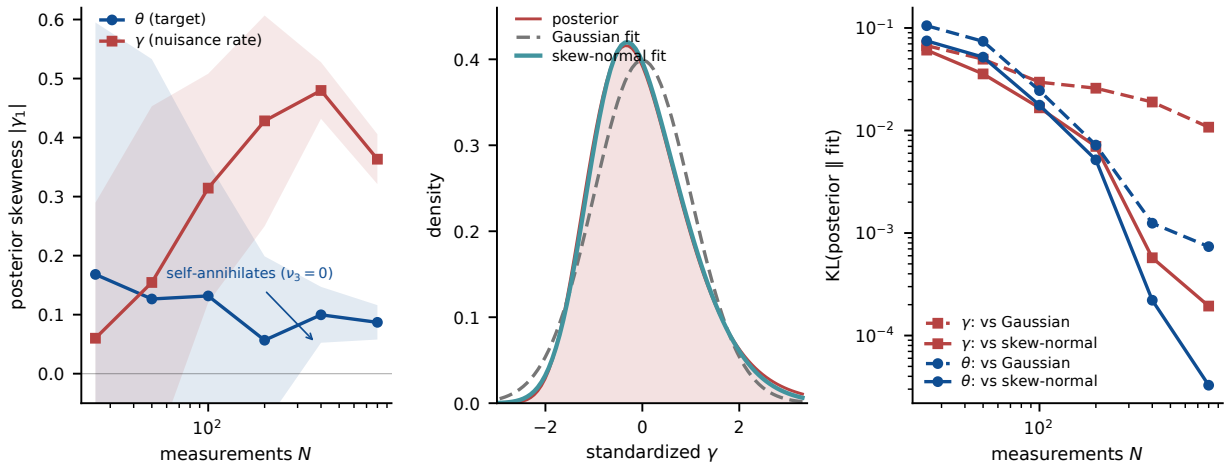
12 Where skewness lives, and the limit of exploiting it

The directional annihilation of theorem 11.3 and corollary 11.6 has a sharp, design-relevant consequence, summarized in fig. 10. Under the information-optimal adaptive design the skewness of the *parameter-of-interest* (θ) posterior is driven to zero ($\boldsymbol{\nu}_3 = 0$ exactly), so a Gaussian belief is asymptotically adequate for θ. The *nuisance decoherence-rate* γ, by contrast, is a rate/scale parameter whose posterior stays right-skewed: its third-cumulant components are *not* removed by any θ-optimal design, and the marginal skewness persists at $O(1)$ across three decades of N [fig. 10(a)].

This is a floor, not a transient. Fitting the γ-marginal with a Gaussian [fig. 10(b)] systematically misses the skewed tail, incurring an *irreducible* Kullback–Leibler error set by the persistent skewness; a skew-normal belief removes it. The gap *widens* with data: at $N = 800$ a Gaussian fit of the γ-marginal costs $\approx 57\times$ the KL of a skew-normal fit, whereas for θ the two are indistinguishable [fig. 10(c)]. *This is a density-level (KL) effect*: section 14 and our experiments show it does *not* translate into a materially better point estimate or credible interval; the practical payoff is the robust design of section 10, not skew-aware estimation per se.

13 Numerical verification ledger

Every numbered result above is checked against an independent ground truth (grid, Monte-Carlo, finite differences, or LP). Script: `validate_math.py` + companions, seed `20260626`.



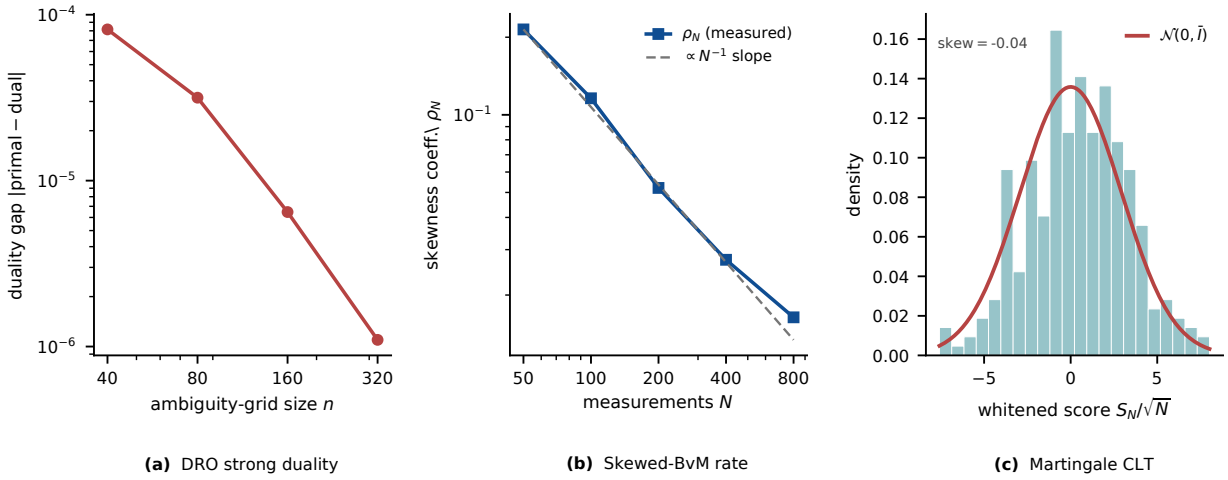
(a) Directional skewness **(b)** γ-marginal at $N = 200$ **(c)** KL: Gaussian vs skew-normal

Figure 10. Where skewness lives: the nuisance decoherence-rate direction. (a) Posterior skewness vs. number of measurements N under the information-optimal adaptive design: the parameter-of-interest θ *self-annihilates* ($\boldsymbol{\nu}_3 = 0$, theorem 11.3) while the nuisance decoherence-rate γ retains $O(1)$ skewness (bands: ± 1 s.d. over 30 seeds). (b) The γ-marginal posterior at $N = 200$ (standardized): a Gaussian fit misses the skewed tail that a skew-normal fit captures (skewness 0.53). (c) KL(posterior $\|$ fit) vs. N: for γ the Gaussian error plateaus while the skew-normal error keeps falling (up to 57× smaller at N=800); for θ both vanish. Together: a Gaussian belief is adequate for the target but *lossy at the density (KL) level* for the nuisance rate.

Part	What is checked	Result
1	score (3) vs finite diff; Fisher (4) vs $\mathbb{E}_m[s^2]$; $V \to 1$ limit	PASS ($< 10^{-10}$)
2	few-shot (N=3) posterior skewness	$\gamma_1 = -0.46$ (skewed)
3	moment-matched SN vs Gaussian, KL, by skew bucket	SN wins 97%; ratio 1.2 → 1.7×
3b	*local* skew-modal (10) vs Gaussian	46.6% (negative result)
4	sequential ADF tracks batch posterior	mean KL = 0.012; beats Gauss 74%
5	Sherman–Morrison (13) vs direct inverse (40 steps)	max $\lvert\Delta\rvert = 1.1 \times 10^{-16}$
6	Wasserstein DRO duality (16): primal vs dual	match $< 10^{-3}$
7	continuous-γ DRO duality, *real* risk (17)	gap below $5{\times}10^{-8}$, Λ=0.59
8	Lyapunov (22): sup $\|\mathbf{s}_n\|$; ratio slope	5.23; −0.48
9	d=1 adaptive: analytic ρ_N vs measured skew	match ≤ 3%
10	d=2 matrix SLLN/CLT (Cov $\mathbf{z} \to \mathbb{I}_2$)	PASS
11	learned-agent generalization: specialists vs. robust	worst-case 24% vs. 28–140%
12	CEM training: worst-case regret vs. avg-RMSE tie	regret 146→18%; avg 0.128–0.139
13	fuller baselines (design+agent) + ε-sensitivity	DRO best regret; naive minimax/BMA over-protect
14	4 non-Markovian channels; Gaussian vs. skew ADF	robust −36 to −52% (3/4); skew gain < 1%

14 Scope and limitations

- **Multimodality.** In the deepest few-shot regime ~ 20% of posteriors are multimodal (the periodic likelihood aliases). A single skew-normal cannot represent these; resolving them is the *first job of the adaptive*



(a) DRO strong duality **(b)** Skewed-BvM rate **(c)** Martingale CLT

Figure 11. Theory verification. (a) Continuous-γ Wasserstein strong duality: the primal–dual gap vanishes ($\to 10^{-6}$) as the ambiguity grid refines (theorem 8.4). (b) Adaptive skewed quantum Bernstein–von Mises (theorems 11.2 and 11.3): the standardized skewness coefficient ρ_N decays as N^{-1} under the information-optimal design. (c) Martingale central limit theorem: the whitened score $S_N/\sqrt{N}$ is standard normal (sample skewness -0.04), validating lemma 11.1 and the CLT step of theorem 11.2.

design (localize to one fringe), after which the skew-normal engine applies. This is a feature of the problem statement, not a bug of the skew-normal belief.

- **Skewness saturation.** (9) caps $|\gamma_1| < 0.9953$; beyond that a single skew-normal is structurally inadequate and a richer skew-symmetric or mixture family is required. We report this rather than hide it.
- **Local vs global.** The *local* third-order Laplace correction is not a finite-N estimator here (section 6.2); C4 uses it only as the *asymptotic* limit law (theorem 11.3 makes this precise).

References

[1] C. L. Degen, F. Reinhard, and P. Cappellaro, "Quantum sensing," *Rev. Mod. Phys.*, vol. 89, 035002, 2017.

[2] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nat. Photon.*, vol. 5, pp. 222–229, 2011.

[3] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states," *Phys. Rev. Lett.*, vol. 72, pp. 3439–3443, 1994.

[4] M. G. A. Paris, "Quantum estimation for quantum technology," *Int. J. Quantum Inf.*, vol. 7, pp. 125–137, 2009.

[5] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, "The elusive Heisenberg limit in quantum-enhanced metrology," *Nat. Commun.*, vol. 3, 1063, 2012.

[6] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, "Achieving the Heisenberg limit in quantum metrology using quantum error correction," *Nat. Commun.*, vol. 9, 78, 2018.

[7] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, 79, 2018.

[8] A. Sergeevich, A. Chandran, J. Combes, S. D. Bartlett, and H. M. Wiseman, "Characterization of a qubit Hamiltonian using adaptive measurements in a fixed basis," *Phys. Rev. A*, vol. 84, 052315, 2011.

[9] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, "Robust online Hamiltonian learning," *New J. Phys.*, vol. 14, 103013, 2012.

[10] A. Hentschel and B. C. Sanders, "Machine learning for precise quantum measurement," *Phys. Rev. Lett.*, vol. 104, 063603, 2010.

[11] A. Foster, D. R. Ivanova, I. Malik, and T. Rainforth, "Deep Adaptive Design: Amortizing sequential Bayesian experimental design," in *Proc. Int. Conf. Machine Learning (ICML)*, 2021. arXiv:2103.02438.

[12] D. R. Ivanova, A. Foster, S. Kleinegesse, M. U. Gutmann, and T. Rainforth, "Implicit deep adaptive design: Policy-based experimental design without likelihoods," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[13] L. J. Fiderer, J. Schuff, and D. Braun, "Neural-network heuristics for adaptive Bayesian quantum estimation," *PRX Quantum*, vol. 2, 020303, 2021.

[14] F. Belliardo, F. Zoratti, and V. Giovannetti, "Applications of model-aware reinforcement learning in Bayesian quantum metrology," *Phys. Rev. A*, vol. 109, 062609, 2024.

[15] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric," *Math. Program.*, vol. 171, pp. 115–166, 2018.

[16] R. Gao and A. J. Kleywegt, "Distributionally robust stochastic optimization with Wasserstein distance," *Math. Oper. Res.*, vol. 48, no. 2, pp. 603–655, 2023.

[17] J. Blanchet and K. Murthy, "Quantifying distributional model risk via optimal transport," *Math. Oper. Res.*, vol. 44, no. 2, pp. 565–600, 2019.

[18] H. Rahimian and S. Mehrotra, "Distributionally robust optimization: A review," arXiv:1908.05659, 2019.

[19] A. Azzalini, "A class of distributions which includes the normal ones," *Scand. J. Stat.*, vol. 12, pp. 171–178, 1985.

[20] D. Durante, F. Pozza, and B. Szabó, "A skewed Bernstein–von Mises theorem and skew-modal approximations," *Annals of Statistics*, 2024.

[21] M. Tsang, "Ziv–Zakai error bounds for quantum parameter estimation," *Phys. Rev. Lett.*, vol. 108, 230401, 2012.

[22] X.-M. Lu and M. Tsang, "Quantum Weiss–Weinstein bounds for quantum metrology," *Quantum Sci. Technol.*, vol. 1, 015002, 2016.

[23] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*. New York: Academic Press, 1980.

[24] X. Yu, X. Zhao, L. Li, S. Kurdziałek, C. Zhang, C.-F. Li, and G.-C. Guo, "Towards robust optimal measurements against noise in quantum metrology," *Sci. China Phys. Mech. Astron.*, vol. 69, no. 7, 2026.

A Proofs

Proof of lemma 7.1. Specialize the Sherman–Morrison–Woodbury identity

$$(\mathbf{A} + \mathbf{u}\mathbf{u}^\top/v)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{u}^\top\mathbf{A}^{-1}}{v + \mathbf{u}^\top\mathbf{A}^{-1}\mathbf{u}}$$

with $\mathbf{A} = \mathbf{J}_{n-1}$, $\mathbf{u} = \mathbf{s}_n$; full algebra in appendix D. □

Proof of lemma 8.3. $\mathbf{K}(\gamma; t)$ is C^1 in γ (each $\mathbf{I}$ is C^1 via $V = e^{-\gamma t}$) and uniformly PD; matrix inversion is smooth on the PD cone, so $L = [\mathbf{K}^{-1}]_{\theta\theta}$ is C^1 with the stated derivative, bounded since $\|\mathbf{K}^{-1}\| \le \kappa^{-1}$ and $\|\partial_\gamma \mathbf{K}\|$ is bounded. Numerically $\Lambda = 0.59$ on $[0.02, 0.6]$ (section 13); the finite-N risk *exceeds* (17) in the nuisance direction, consistent with the persistent γ-skewness of section 11.5. □

Proof of corollary 8.5. (c): for $\lambda \ge \Lambda$, $L(\gamma) - \lambda|\gamma - \hat\gamma| \le L(\hat\gamma) + (\Lambda - \lambda)|\gamma - \hat\gamma| \le L(\hat\gamma)$, equality at $\gamma = \hat\gamma$.
(b): $\lambda \mapsto L(\gamma) - \lambda|\gamma - \hat\gamma|$ is affine, the pointwise $\sup_\gamma$ is convex, and the $\widehat{Q}$-average preserves convexity. □

Proof of lemma 11.1. $\|\mathbf{s}_n\| = t_n V_n/|D_n|$ since $\|[\sin, \cos]\| = 1$; and $|D_n| \ge 1 - V_n > 0$ gives (21). Boundedness of $\|\mathbf{s}_n\|$ makes every conditional moment $\le C^{2+\delta}$; summing N such terms and dividing by $N^{1+\delta/2}$ yields (22). □

Proof of theorem 11.2. **Step 1 (consistency).** $\{\sum_{k\le n} \mathbf{s}_k(\boldsymbol{\eta}_0)\}$ is a square-integrable martingale (zero predictable mean); by the martingale SLLN [23, Thm. 2.18] with Assumption 2, $N^{-1}\nabla^2\ell_N \to -\bar{\mathbf{I}}$ a.s. and $\hat{\boldsymbol{\eta}}_N - \boldsymbol{\eta}_0 = O_p(N^{-1/2})$.
Step 2 (conditional third-order Laplace). Condition on $\mathscr{F}_N$ (designs are then constants); Taylor-expand in $\mathbf{h}$:

$$\ell_N(\boldsymbol{\eta}) - \ell_N(\hat{\boldsymbol{\eta}}_N) = -\tfrac{1}{2}\mathbf{h}^\top\mathbf{h} + \tfrac{1}{6}\boldsymbol{\rho}_N[\mathbf{h},\mathbf{h},\mathbf{h}] + R_N, \qquad \|R_N\| = O_p(N^{-1}),$$

bounded by the uniform fourth-derivative bound (Assumption 1).
Step 3 (matrix martingale SLLN/CLT). The realized-minus-predictable parts $\nabla^2\ell_N + \mathbf{K}_{2,N}$ and $\nabla^3\ell_N - \mathbf{K}_{3,N}$ are tensor martingales with $O(N)$ quadratic variation; normalizing by $\mathbf{K}_{2,N} \asymp N$ gives observed = predictable $(1 + o_p(1))$. The whitened score satisfies

$$\mathbf{K}_{2,N}^{-1/2}\nabla\ell_N(\boldsymbol{\eta}_0) \Rightarrow \mathcal{N}(\mathbf{0}, \mathbb{I}_d)$$

by the multivariate martingale CLT [23, Thm. 3.2] with Cramér–Wold under lemma 11.1.
Step 4 (skew-symmetric matching). Apply the skewed-BvM expansion [20, Thm. 1] to the conditioned fixed-design target: the leading odd term is absorbed into $2\phi_d\Phi(\cdot)$, leaving $O(\|\boldsymbol{\rho}_N\|^2) + O(\|R_N\|) = O_p(N^{-1})$; the Gaussian keeps the $O(\|\boldsymbol{\rho}_N\|)$ term.
Step 5 (unconditioning). Constants depend only on $\bar{\mathbf{I}}$ and the uniform bounds (Assumption 2,lemma 11.1), so the conditional bounds hold a.s. unconditionally; $\boldsymbol{\rho}_N = O_p(N^{-1/2}) \to 0$. □

Proof of theorem 11.3. At $c := \cos\theta_0 t = 0$ the outcome law is symmetric, $p(\pm 1) = \frac{1}{2}$, and the per-term third log-derivative

$$\ell_3(m) = a\,t^3\sin(\theta_0 t)(1 - 2a^2), \qquad a = mV,$$

is *odd* in m; hence $\nu_3 = \frac{1}{2}[\ell_3(+1) + \ell_3(-1)] = 0$. The information maximizer sits at $|\sin\theta_0 t| = 1$ (equivalently $c = 0$), so $\mathbf{K}_{3,N}$ collects only $O_p(\sqrt{N})$ martingale fluctuations. □

B Derivation of the Fisher information (4)

From (2), $\partial_\theta p = -\frac{1}{2}mVt\sin(\theta t)$, so $s = \partial_\theta \log p = \partial_\theta p/p$ gives (3). Using $\sum_m p(m)[\cdot]$ with $m^2 = 1$ and $p(\pm 1) = \frac{1}{2}(1 \pm V\cos\theta t)$,

$$I = \sum_{m=\pm 1} p(m)\,\frac{V^2t^2\sin^2(\theta t)}{(1+mV\cos\theta t)^2} = \frac{V^2t^2\sin^2(\theta t)}{2}\sum_{m=\pm 1}\frac{1}{1+mV\cos\theta t}$$
$$= \frac{V^2t^2\sin^2(\theta t)}{2}\cdot\frac{2}{1-V^2\cos^2(\theta t)} = \frac{V^2t^2\sin^2(\theta t)}{1-V^2\cos^2(\theta t)}, \tag{24}$$

which is (4). □

C Derivation of the skew-modal form (10)

Expand $\ell_N(\theta)$ about its mode $\hat\theta$ ($\ell_N'(\hat\theta) = 0$): $\ell_N(\theta) - \ell_N(\hat\theta) = -\frac{1}{2}ju^2 + \frac{1}{6}\ell_3u^3 + O(u^4)$, $u = \theta - \hat\theta$. In the BvM scaling $u = O(N^{-1/2})$, $j = O(N)$, $\ell_3 = O(N)$, so $\frac{1}{2}ju^2 = O(1)$ and $\frac{1}{6}\ell_3u^3 = O(N^{-1/2})$. Hence $e^{\ell_N - \ell_N(\hat\theta)} \approx e^{-\frac{1}{2}ju^2}(1 + \frac{1}{6}\ell_3u^3)$. Using $2\phi_j(u)\Phi(w(u)) \approx \phi_j(u)(1 + \sqrt{2/\pi}\,w(u))$ for small w and matching the cubic, $w(u) = \frac{\ell_3}{6}\sqrt{\pi/2}\,u^3$, giving (10). □

D Sherman–Morrison algebra (lemma 7.1)

With $\mathbf{A} = \mathbf{J}_{n-1} \succ 0$, $\mathbf{u} = \mathbf{s}_n$, $v > 0$, $(\mathbf{A} + \mathbf{u}\mathbf{u}^\top/v)(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{u}^\top\mathbf{A}^{-1}}{v+\mathbf{u}^\top\mathbf{A}^{-1}\mathbf{u}}) = \mathbf{I}$; the cross terms cancel via $s = \mathbf{u}^\top\mathbf{A}^{-1}\mathbf{u}$ through $\frac{1}{v} - \frac{s/v+1}{v+s} = 0$, giving (13). □

E Wasserstein duality (theorem 8.2)

For $p = 1$, $\sup_{Q:\mathscr{W}_1(Q,\widehat{Q})\le\varepsilon}\mathbb{E}_Q[L]$ is a linear program in the transport plan; Lagrangian relaxation of $\mathscr{W}_1 \le \varepsilon$ with multiplier $\lambda \ge 0$ and strong LP duality give $\inf_{\lambda\ge 0}\{\lambda\varepsilon + \mathbb{E}_{\widehat{Q}}[\sup_\gamma(L(\gamma) - \lambda|\gamma - \hat\gamma|)]\}$, which is (16). For Λ-Lipschitz L the inner c-transform is finite iff $\lambda \ge \Lambda$, so the worst case is $\mathbb{E}_{\widehat{Q}}[L] + \Lambda\varepsilon + o(\varepsilon)$. □

F Robustness experiment: setup and reproducibility

Model and schedule. Single-qubit dephasing Ramsey (2) with $\theta_0 = 1$, $N = 30$ shots at times t_n equispaced on $[0.3, T]$ (spanning T resolves phase ambiguity). The design objective is the Cramér–Rao risk $L(T;\gamma) = 1/J(T,\gamma)$, $J(T,\gamma) = \sum_n I(\theta_0, t_n;\gamma)$ with I from (4). Nominal rate $\hat\gamma = 0.2$; ambiguity $\Gamma = [0.05, 0.6]$ (true rate up to 3× nominal); nominal law $\widehat{Q}$ a truncated Gaussian at $\hat\gamma$; Wasserstein radius $\varepsilon = 0.06$. **Designs.** *Nominal* $T^\star = \arg\min_T L(T;\hat\gamma)$; *minimax* $\arg\min_T\max_{\gamma\in\Gamma}L(T;\gamma)$; *DRO* $\arg\min_T\sup_{Q\in\mathcal{B}_\varepsilon(\widehat{Q})}\mathbb{E}_Q[L(T;\gamma)]$ via the dual (18). Optimized knobs: $T_{\text{nom}} = 5.46$, $T_{\text{minimax}} = 2.11$, $T_{\text{DRO}} = 2.25$.

design	risk @ $\hat\gamma$	worst-case risk	$\mathbb{E}$[risk] mismatch	worst regret vs. oracle
nominal	**0.024**	0.265	0.086	0.127
minimax	0.039	**0.138**	0.066	0.021
DRO	0.037	0.139	**0.065**	**0.019**

Actual estimation (fig. 5b). For each γ_{true}, 110 Monte-Carlo runs ($\theta^\star \sim \mathcal{N}(1, 0.5^2)$), exact grid posterior, point estimate = posterior mean. At $\gamma_{\text{true}} = 0.55$ (a 3× mismatch) the RMSE is {nominal 0.41, minimax 0.34, DRO 0.31}, the robust designs are ≈ 17–20% lower. Scripts: `exp2_robust.py` (numbers), `figures/gen_fig_robust.py` (fig. 5); seed 7. The negative skew-estimation results of sections 12 and 14 are reproduced by `exp_cashout.py` and `exp_cashout_curve.py`.

G The learned design agent: architecture and training

Policy. A two-layer perceptron $\pi_\theta : \mathbb{R}^2 \to [t_{\min}, t_{\max}]$ with a 4-unit tanh hidden layer (17 weights); input the belief features $(\sigma_\theta,\ n/N)$, output $t = t_{\min} + (t_{\max} - t_{\min})\ \text{sigmoid}(\cdot)$ with $[t_{\min}, t_{\max}] = [0.3, 11]$. **Estimator.** a Gaussian/skew-normal assumed-density filter on θ (section 6); prior $\theta \sim \mathcal{N}(1, 0.15^2)$, $N = 20$ shots. **Training (evolutionary strategy).** the cross-entropy method on a diagonal Gaussian over the 17 weights: 26 iterations, population 50, elite 12; each candidate scored by the Monte-Carlo loss (19) (40–60 rollouts). The nominal agent scores at $\gamma{=}\hat{\gamma}{=}0.2$; the robust agent scores by the worst case over $\gamma \in \{0.10, 0.30, 0.55\} \subset \Gamma$, the sole difference between the two agents. **Evaluation.** 5000 rollouts per γ_{true}; seeds 0 (training), 7 (evaluation). Scripts `verify_ai_agent.py`, `figures/gen_fig_agent.py`. **Generalization and training experiments (figs. 3 and 4).** For the generalization matrix we train four nominal specialists, one per assumed rate $\hat{\gamma} \in \{0.10, 0.20, 0.35, 0.50\}$ (each scored only at its own $\hat{\gamma}$), and one robust agent scored by the worst-case *regret* ratio $L(\theta;\gamma)/L^\star(\gamma)$ over the grid $\gamma \in \{0.08, 0.18, 0.30, 0.45, 0.60\}$, where $L^\star(\gamma)$ is the best constant-t (non-adaptive) oracle at that rate, obtained by a 46-point line search. Each agent is then evaluated at 5000 rollouts per true rate. The learning curve (fig. 4a) reports the worst-case suboptimality of the cross-entropy-method mean iterate over the decoherence set (20 iterations, population 44, elite 11). Scripts `verify_ai_agent_matrix.py`, `figures/gen_fig_agent_generalization.py`, `figures/gen_fig_agent_training.py`; seeds 0/1 (training), 3 (oracle), 5/7/9 (evaluation).

H Reproducibility and code availability

Cross-entropy-method (CEM) settings. All learned agents use the same gradient-free CEM: a diagonal Gaussian over the 17 policy weights, initialized at mean $\mathbf{0}$ and standard deviation $\sigma_0 = 0.8$. At each iteration we sample the population, score every member by the Monte-Carlo loss, keep the top-k *elite* (elite fraction ≈ 24%: k=12 of 50 for the main agents, 11 of 44 for the generalization study), and refit the Gaussian to the elite mean and standard deviation with an additive exploration floor +0.03. There is no gradient or learning rate; optimization runs for a fixed number of iterations (26 main, 20 generalization, 12 baselines) with no early stopping. Interrogation times are clamped to $[t_{\min}, t_{\max}] = [0.3, 11]$. Seeds: 0/1 (training), 3 (oracle line search), 5/7/9 (evaluation), 11 (convergence Monte-Carlo).

Code. Every experiment is driven by a small, self-contained `numpy` script, all available at https://github.com/changhaohe-glitch/Quantum-Comm: `validate_math.py`, `verify_ai_agent.py`, `verify_ai_agent_matrix.py`, `exp_baselines.py`, `exp_ablation.py`, `exp_nonmarkov.py`, `exp2_robust.py`, and the figure generators `figures/gen_fig_*.py`. Each figure caption names the script that produces it.