CONTINUAL LEARNING AS ASYMMETRIC NASH BARGAINING WITH ABSENT PLAYERS

Anonymous authors
Paper under double-blind review

ABSTRACT

Gradient-based multi-task learning methods assume that all task gradients are available when an update is selected. Continual learning violates this assumption: past tasks are absent and can only be represented through imperfect proxies, such as replay buffers, regularization statistics, or frozen models. Unlike ordinary minibatch noise, these proxies may introduce systematic, non-vanishing errors in the gradient and curvature information that defines each task utility. We formulate task-incremental continual learning as an asymmetric n-player Nash bargaining problem in which absent players are represented by uncertain proxies, utilities are second order, and disagreement points are tolerable-forgetting budgets. Taking the worst case over bounded proxy errors yields an exact robust bargaining objective whose feasible set is nonempty and bounded without a trust region. We first characterize how proxy errors perturb the bargaining solution, deriving a sensitivity bound governed by the proxy error bounds, bargaining slack, and curvature conditioning. We then connect this perturbation to catastrophic forgetting through a bound that separates the ideal bargaining contribution, proxy-induced update error, and a third-order curvature remainder. Finally, we derive a damage certificate that is computable at the returned update and involves no oracle quantity. Together, these results identify bargaining slack as a central source of fragility in continual learning and provide a principled basis for bargaining with imperfect representations of past tasks. Our source code is available at `https://anonymous.4open.science/r/CL-Nash-bargaining-8584/`

1 INTRODUCTION

Continual learning (CL) requires making decisions that affect objectives no longer directly observable. We model this dynamic as an asymmetric cooperative game between the fully observable current task and *absent* past tasks. These absent players must be represented through imperfect proxies, such as replay buffers or frozen prototypes. While framing multi-task optimization as a Nash bargaining game provides a principled analytical lens for balancing task objectives (Navon et al., 2022), naively translating this framework to the CL setting encounters fundamental structural obstructions due to the nature of absent players:

- **Loss of Directional Identifiability at Historical Optima:** Because the model has recently converged on past tasks, their true gradient norms $\|g_i\|$ naturally approach zero. Under a standard first-order utility model, $u_i(\Delta) = -\langle \hat{g}_i, \Delta \rangle$, the direction an absent player protects is dictated by the non-vanishing proxy estimation error e_i rather than the true task signal. As the ratio $\|e_i\|/\|g_i\|$ becomes large, this loss of directional identifiability causes the bargain to blindly optimize noise, while the true curvature-driven vulnerability near the optimum remains invisible.
- **The Disagreement Point Dilemma:** Standard multi-task bargaining assumes a disagreement point of $d_i = 0$, demanding strict loss improvement for all tasks. In CL, incorporating curvature reveals that past tasks near their optima have inherently non-positive utility ($u_i(\Delta) \approx -\frac{1}{2}\Delta^\top \hat{H}_i \Delta \leq 0$), collapsing the individually rational feasible set to the origin. CL seeks bounded degradation, not perpetual improvement. Conversely, heuristic conflict-dependent disagreement points (e.g., $d_i \propto \langle g_i, g_{\text{new}} \rangle$) are normatively backwards: they re-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

duce an old task's bargaining weight precisely when it faces severe conflict, inadvertently withdrawing protection where catastrophic forgetting is most imminent.

To resolve these structural flaws, we reformulate CL as an asymmetric Nash bargaining problem utilizing second-order utilities. By incorporating positive semidefinite curvature operators ($\hat{H}_i \succeq 0$), absent players bargain not by pulling with unreliable gradients, but by vetoing high-curvature directions. Furthermore, we replace the standard disagreement point with a principled tolerable-forgetting budget, $d_i = -\tau_i$. This establishes a formal stability-plasticity trade-off, ensuring bounded degradation for past tasks while maintaining a rigorous framework for continuous adaptation.

2 Related Work

Bargaining and multi-task optimization. Multi-task learning (MTL) jointly trains a model on several tasks, often by minimizing a weighted sum of task losses. To address gradient conflicts, MGDA (Sener & Koltun, 2018) seeks a common descent direction, PCGrad (Yu et al., 2020) projects away conflicting components, and CAGrad (Liu et al., 2021) balances average descent with worst-task local improvement. Nash's bargaining solution (Nash, 1950) provides an axiomatic rule for selecting mutually beneficial outcomes relative to a disagreement point. Building on this, Nash-MTL (Navon et al., 2022) formulates task balancing as a bargaining problem, while AuxiNash (Shamsian et al., 2023) introduces asymmetric bargaining power. Applying these formulations to continual learning requires accounting for past tasks whose data are no longer fully available and whose objectives must therefore be represented through proxies.

Continual learning methods. Continual learning methods retain information about past tasks through stored examples, parameter statistics, or model predictions. EWC (Kirkpatrick et al., 2017) penalizes parameter changes according to their estimated importance for previous tasks. GEM (Lopez-Paz & Ranzato, 2017) uses episodic-memory gradients to constrain updates, while A-GEM (Chaudhry et al., 2019) uses an averaged reference gradient. Dark Experience Replay (Buzzega et al., 2020) preserves past predictions through logit matching, with DER++ additionally replaying labels. We study the uncertainty in such surrogate representations when they are used to construct bargaining utilities and disagreement points, analyzing how resulting perturbations affect catastrophic forgetting.

Subspace methods. Existing subspace-based approaches, such as Orthogonal Gradient Descent (Farajtabar et al., 2020) and Gradient Projection Memory (Saha et al., 2021), mitigate forgetting by artificially constraining the update to a subspace orthogonal to the critical directions of past tasks. In contrast, our proposed framework also operates within a low-dimensional subspace spanned by the players' gradients and curvature ranges, but for a fundamentally different reason: the optimal update naturally lies in this subspace rather than being confined to it by design (Proposition 4).

3 Continual Learning with Absent Players

3.1 A Two-Task Motivating Example

We investigate the structural failure of multi-task bargaining when a task leaves the training stream. Consider a two-player scenario with an absent past task A and a new live task B. Because task A has just finished training, the model sits near its optimum, meaning task A's true gradient is negligible ($\|g_A\| \approx 0$).

In standard first-order Nash bargaining, player A evaluates any candidate update Δ through the linear utility $u_A(\Delta) = -\langle \hat{g}_A, \Delta \rangle$, where $\hat{g}_A = g_A + e_A$ is computed from a buffer. Since the Nash solution is invariant to positive rescaling of u_A, the small norm of g_A does not reduce the influence of A. Instead,

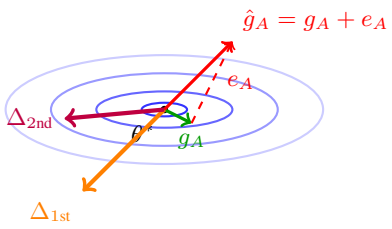

Figure 1: **Proxy Dominance:** Near the optimum θ^*, the vanishing true gradient g_A leaves the proxy $\hat{g}_A$ dominated by noise e_A. A first-order update (Δ_{1st}) follows this noise across steep contours, whereas a curvature-aware update (Δ_{2nd}) safely navigates the flat valley.

when $\|e_A\|$ is comparable to $\|g_A\|$, the half-space $\{\Delta : \langle \hat{g}_A, \Delta \rangle < 0\}$ that the bargain protects is determined largely by the proxy error. Consequently, the mechanism may select a step that heavily damages A while satisfying its illusory first-order requirement.

This damage is invisible to the first-order proxy but becomes strictly apparent once curvature is introduced: near a local minimum, the damage inflicted by an update is dominated by the second-order geometry of the loss surface.

3.2 Players, Utilities, and Disagreement Points

Let $\theta \in \mathbb{R}^p$ be the parameters and $\Delta \in \mathbb{R}^p$ a candidate update. Player 0 is the live task, with full data available. Players $i = 1, \ldots, n-1$ are absent tasks represented by a buffer of m_i examples together with a curvature operator.

Definition 1 (Curvature utility)**.** For each player, the utility is defined as the second-order predicted decrease of L_i: $u_i(\Delta) = -\langle \hat{g}_i, \Delta \rangle - \frac{1}{2}\Delta^\top \hat{H}_i \Delta$, where $\hat{g}_i$ is an estimate of $\nabla L_i(\theta)$ and $\hat{H}_i \succeq 0$ is a curvature operator.

Nash-MTL and AuxiNash use zero disagreement utility (Navon et al., 2022; Shamsian et al., 2023). We instead encode retention requirements through task-specific budgets, allowing bounded modeled loss increases on absent tasks.

Definition 2 (Tolerable-forgetting disagreement point)**.** Set $\tau_0 = 0$ for the live task and prescribe $\tau_i > 0$ for each absent task. Define $d_i = -\tau_i$ and the nominal proxy surplus $\hat{s}_i(\Delta) = \hat{u}_i(\Delta) - d_i = \tau_i - \widehat{\mathrm{dam}}_i(\Delta)$, where $\widehat{\mathrm{dam}}_i(\Delta) = -\hat{u}_i(\Delta)$ is the predicted loss increase. Thus, $\hat{s}_i(\Delta) > 0$ precisely when the predicted loss increase is below τ_i.

The tolerances (τ_i) express the stability–plasticity trade-off through modeled retention constraints. Smaller tolerances impose tighter limits on modeled deterioration for absent tasks, whereas larger tolerances allow a broader set of updates that may support adaptation to the current task.

3.3 Gradient and Curvature Uncertainty

Since the data of past tasks are unavailable, their gradients and curvature can only be approximated from stored examples. Consequently, the quadratic utility used in the bargain may differ from the oracle model.

Assumption 1 (Derivative approximation errors)**.** At the current parameters $\theta \in \mathbb{R}^p$, for each task $i \in \{0, \ldots, n-1\}$, let $g_i := \nabla L_i(\theta)$ and $H_i := \nabla^2 L_i(\theta)$ denote the gradient and Hessian of the twice continuously differentiable task loss $L_i : \mathbb{R}^p \to \mathbb{R}$. Let $\hat{g}_i$ and $\hat{H}_i$ denote their respective estimates. There exist specified finite constants $r_i, \zeta_i \geq 0$ such that:

$$\|\hat{g}_i - g_i\|_2 \leq r_i, \qquad \|\hat{H}_i - H_i\|_2 \leq \zeta_i. \tag{1}$$

We impose the following conditions:

1. *(Bias)* Define $b_i := \mathbb{E}[\hat{g}_i] - g_i$. The bias b_i is not required to vanish.
2. *(Non-vanishing)* $r_i \geq \|b_i\| + \sigma_i/\sqrt{m_i}$ with σ_i the per-example gradient dispersion, so $r_i \not\to 0$ as $m_i \to \infty$.
3. *(Endogeneity)* $r_i^{(t)} \approx r_i^{(0)} + \kappa_i \|\theta_t - \theta_{t_i}\|$, and the drift is produced by the updates the bargain itself selects.

3.4 The Bargaining Program

Under Assumption 1, taking the worst case over the bounded uncertainty set yields the robust slack $\hat{s}_i^R(\Delta)$:

$$\hat{s}_i^R(\Delta) = \tau_i - \langle \hat{g}_i, \Delta \rangle - \tfrac{1}{2}\Delta^\top \hat{H}_i \Delta - r_i\|\Delta\| - \tfrac{\zeta_i}{2}\|\Delta\|^2. \tag{2}$$

The bargaining program is then formulated to maximize the logarithmic objective $\hat{\varphi}(\Delta)$ over the open convex domain $\hat{\mathcal{S}} = \{\Delta \in \mathbb{R}^p : \hat{s}_i^R(\Delta) > 0 \text{ for all } i\}$:

$$\hat{\Delta} \in \arg\max_{\Delta \in \hat{\mathcal{S}}} \hat{\varphi}(\Delta), \qquad \text{where} \quad \hat{\varphi}(\Delta) = \sum_{i=0}^{n-1} \alpha_i \log \hat{s}_i^R(\Delta). \tag{3}$$

4 Sensitivity of the Bargaining Update

4.1 Feasibility and Boundedness

To analyze the proxy bargain, we define the comprehensive utility set $\mathcal{U} = \{y \in \mathbb{R}^n : \exists \Delta \in \mathbb{R}^p,\ y \le u^R(\Delta)\}$ and its individually rational subset $\mathcal{U}_d = \{y \in \mathcal{U} : y \ge d\}$. Let I denote the identity on $\mathbb{R}^p$, $\bar{H} = \sum_i \hat{H}_i$, $Z = \sum_i \zeta_i$, and aggregate curvature $\gamma = \lambda_{\min}(\bar{H}) + Z$.

Theorem 1 (Convexity, attainment, and uniqueness)**.** *The surplus $\hat{s}_i^R$ is continuous and concave, making $\hat{\mathcal{S}}$ open and convex. If $\hat{\mathcal{S}}$ is nonempty and bounded, then $\mathcal{U}_d$ is compact and convex, and the objective equation 3 attains its maximum in $\hat{\mathcal{S}}$ at a unique utility vector $y^\star > d$. If $\gamma > 0$, the maximizing update $\hat{\Delta}$ is strictly unique.*

Boundedness without Trust Regions. A critical structural advantage of our second-order formulation is that the feasible set inherently constrains the step size. As detailed in Appendix B, strict feasibility $\hat{\mathcal{S}} \neq \emptyset$ holds if and only if the current task gradient $\hat{g}_0 \neq 0$. Furthermore, whenever $\gamma > 0$, every feasible update is strictly bounded by $D_{\max} := (\|\bar{g}\| + \sqrt{\|\bar{g}\|^2 + 2\gamma\bar{\tau}})/\gamma$, yielding a uniform positive upper bound $U_{\max}$ for all proxy surpluses. This guarantees optimization stability without imposing arbitrary external trust regions.

4.2 Sensitivity to Derivative Perturbations

To cleanly isolate the effect of derivative estimation, we construct a reference bargaining problem retaining identical budgets and uncertainty penalties, but evaluated on the unobservable oracle variables g_i and H_i. We define the reference surplus $s_i^R(\Delta) = \tau_i - \langle g_i, \Delta\rangle - \frac{1}{2}\Delta^\top H_i \Delta - r_i\|\Delta\| - \frac{\zeta_i}{2}\|\Delta\|^2$ on the domain $\mathcal{S} = \{\Delta : s_i^R(\Delta) > 0\ \forall i\}$, with corresponding objective $\varphi(\Delta)$. Let $\Delta^\star$ denote the oracle maximizer.

Theorem 2 (Strong concavity and sensitivity)**.** *If $\hat{g}_0 \neq 0$ and $\gamma > 0$, then $\hat{\varphi}$ is $\hat{m}$-strongly concave on $\hat{\mathcal{S}}$ with $\hat{m} = \alpha_{\min}\gamma/U_{\max} > 0$, satisfying $\nabla^2\hat{\varphi}(\Delta) \preceq -\frac{\alpha_{\min}}{U_{\max}}(\bar{H} + ZI)$. For any oracle maximizer $\Delta^\star \in \mathcal{S} \cap \hat{\mathcal{S}}$, the update deviation is bounded by:*

$$\|\hat{\Delta} - \Delta^\star\| \le \frac{1}{\hat{m}}\|\nabla\hat{\varphi}(\Delta^\star) - \nabla\varphi(\Delta^\star)\|. \tag{4}$$

Corollary 1 (Explicit sensitivity certificate)**.** *Under the hypotheses of Theorem 2 and Assumption 1, let $D = \|\Delta^\star\|$, $v_i = \|g_i + H_i\Delta^\star\|$, $\mu = \min_i s_i^R(\Delta^\star) > 0$, and $\hat{\mu} = \min_i \hat{s}_i^R(\Delta^\star) > 0$. Then the deviation is bounded explicitly: $\|\hat{\Delta} - \Delta^\star\| \le \frac{U_{\max}}{\alpha_{\min}\gamma}\sum_i \alpha_i \left[\frac{r_i + \zeta_i D}{\hat{\mu}} + \frac{(v_i + r_i + \zeta_i D)\left(r_i D + \frac{1}{2}\zeta_i D^2\right)}{\hat{\mu}\,\mu}\right]$.*

Remark 1 (Cross-feasibility)**.** These analytical bounds fundamentally rely on the oracle maximizer remaining strictly feasible for the proxy problem ($\Delta^\star \in \hat{\mathcal{S}}$). Because $\hat{\mathcal{S}}$ is an open set, cross-feasibility naturally holds whenever proxy errors are sufficiently small relative to optimal surplus margins (empirically confirmed in Appendix F.3).

5 From Bargaining Sensitivity to Forgetting

The following theorem connects bargaining sensitivity to the actual one-step loss change on a past task. It formally separates the ideal capability of the bargain from the damage induced by proxy distortion and local geometric variations.

Theorem 3 (One-step loss bound)**.** *Suppose the local regularity assumptions hold and the oracle objective attains its maximum at $\Delta^\star \in \mathcal{S}$. Let $D = \|\Delta^\star\|$, $\mu = \min_j s_j^R(\Delta^\star) > 0$, $V_i = \sup_{\Delta \in [\Delta^\star, \hat{\Delta}]} \|g_i + H_i \Delta\|$, and $F_i = L_i(\theta + \hat{\Delta}) - L_i(\theta)$. Then the actual loss increase is bounded by $F_i \le \tau_i - s_i^R(\Delta^\star) - r_i D - \frac{\zeta_i}{2} D^2 + V_i \|\hat{\Delta} - \Delta^\star\| + \frac{\rho_i}{6} \|\hat{\Delta}\|^3$. In particular, this simplifies to $F_i \le \tau_i - \mu + V_i \|\hat{\Delta} - \Delta^\star\| + \frac{\rho_i}{6} \|\hat{\Delta}\|^3$.*

Theorem 3 provides a crucial mechanistic understanding of catastrophic forgetting. However, because it relies on the unobservable oracle update $\Delta^\star$, it serves as an analytical tool. To bridge this gap during training, we derive a practical damage certificate evaluated solely at the realized proxy update $\hat{\Delta}$.

Theorem 4 (Computable damage certificate)**.** *Under the local regularity assumptions, let $\hat{\Delta} \in \hat{\mathcal{S}}$ and suppose $\|\hat{g}_i - g_i\| \le r_i$ and $\|\hat{H}_i - H_i\|_2 \le \zeta_i$. Then, independently of the existence of an oracle maximizer, the one-step loss is certified by: $F_i \le \mathcal{C}_i(\hat{\Delta}) := \tau_i - \hat{s}_i^R(\hat{\Delta}) + \frac{\rho_i}{6} \|\hat{\Delta}\|^3$.*

Certified Damage under Shrinkage. A direct consequence of Theorem 4 is a practical admission control mechanism. If a strictly feasible update yields a certificate exceeding the budget ($\mathcal{C}_i(\hat{\Delta}) > \tau_i$), we can scale it by a factor $\beta^\star = \min\{1, \min_i(\tau_i / \mathcal{C}_i(\hat{\Delta}))\}$. As proven in Appendix D, the admitted update $\beta^\star \hat{\Delta}$ strictly satisfies $F_i \le \mathcal{C}_i(\beta^\star \hat{\Delta}) \le \tau_i$ for every absent task i.

6 Experiments

Our empirical evaluation aims to validate the mechanisms predicted by the bargaining model rather than claiming a new state-of-the-art performance metric. Specifically, we sequentially investigate (1) the structural collapse of first-order utilities when past tasks converge, (2) the validity of the computable damage certificate on a real training stream, (3) the physical theory-practice gap in deep network curvature proxies, and (4) the downstream implications on continual learning benchmarks. Table 1 outlines the configuration for all deep learning experiments.

Stream	CIFAR-100, 20 tasks of 5 classes, class order permuted per seed
Backbone	reduced ResNet-18, GroupNorm, width 16 (7.1×10^5 parameters)
Training	5 epochs per task, minibatch 32, SGD at 0.1, three seeds
Memory	reservoir buffer of 2000 examples (4% of the data), identical for all methods
Bargain	one player per task, all present at every step, $\tau_i = 0.05$, $\zeta_i = 0.01$, $r_i = 0$
Curvature	weighted Gauss-Newton on each player's own buffered examples and classes
Metrics	A_{20}, mean accuracy over all tasks after the last one; forgetting, its mean drop

Table 1: Setting of the real-stream experiments.

6.1 Mechanism Validation: Proxy Dominance at Convergence

To empirically verify the structural vulnerability of first-order proxies discussed in Section 3.1, we compare first-order and second-order (curvature) bargains on a controlled quadratic sandbox (details in Appendix F.3). We sweep the norm $\kappa = \|g_{\text{old}}\|$ of the absent tasks' true gradients over four decades to simulate tasks progressively converging to their optima.

As illustrated in Figure 2a, **the conditioning constant $\lambda_{\min}(K)$ of the first-order utility collapses rapidly** from 0.55 at $\kappa = 1$ to 5.5×10^{-9} at $\kappa = 10^{-4}$. In contrast, the curvature conditioning $\lambda_{\min}(\bar{H})$ remains stable at 0.38 regardless of the gradient norm. Consequently, the first-order sensitivity certificate degrades exponentially (Figure 2b). At a matched step norm, **the first-order bargain inflicts a worst-case damage of up to** $2.0\tau_i$, breaching the tolerable-forgetting budget on approximately 80% of instances (Figure 2c). By incorporating curvature, the proposed second-order bargain successfully limits the maximum damage to $0.15\tau_i$, proving that second-order utilities are structurally necessary to maintain directional identifiability and protect absent tasks.

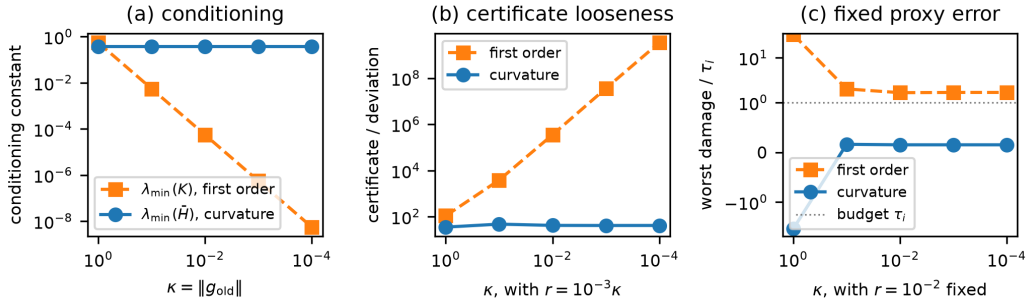

Figure 2: **First-order against curvature bargain as absent tasks converge.** (a) The conditioning constants collapse quadratically for first-order utilities. (b) Ratio of the sensitivity certificate to the measured update deviation. (c) Worst realized damage to an absent task in units of its budget τ_i; note how the first-order approach severely breaches the acceptable threshold.

6.2 Certificate Coverage on a Real Stream

Having established the necessity of second-order utilities, we evaluate the computable damage certificate $\mathcal{C}_i(\hat{\Delta})$ (Theorem 4) on a realistic training trajectory. We record the realized increase of each absent task's loss on a *probe slice*—a disjoint subset comprising one-tenth of its buffered examples that never enters the replay loss. The uncertainty constants (r_i, ζ_i, ρ_i) are calibrated dynamically via a three-variable linear program over a rolling 30-step window.

Metric	**Mean $\pm$ s.d. over three seeds**
Calibrated coverage	0.9608 ± 0.0005
Fixed-constant coverage	0.4067 ± 0.0124
Replay-slice calibrated coverage	0.9573 ± 0.0004
Full-task-loss coverage	0.9863 ± 0.0023
Median certificate margin in units of τ_i	8.4622 ± 0.8559
Fraction with $\beta^\star < 1$ (Admission control triggered)	0.9999 ± 0.0000

Table 2: Results for the exact update on three fixed seeds of the real-stream protocol. Coverage evaluates the fraction of step/task pairs where realized loss increase does not exceed the theoretical certificate.

The results in Table 2 demonstrate that **the calibrated certificate successfully bounds the true damage**, covering $96.08\% \pm 0.05$ of probe-slice observations and $98.63\% \pm 0.23$ of the full task loss across three seeds. However, this theoretical validity reveals a practical selectivity issue: the median certificate margin is 8.46 times the budget τ_i, causing the admission control mechanism (Proposition 5) to aggressively shrink the proposed update on virtually every step (99.99%). This indicates that while the theoretical framework is mathematically sound, the empirical estimates fed into the certificate are overly conservative in deep networks.

6.3 The Theory-Practice Gap: Curvature Proxy vs. True Hessian

To understand why the admission control becomes excessively tight, we investigate the fundamental gap between the theoretical assumptions and the physical geometry of deep neural networks. Theorem 4 requires the operator norm error $\|\hat{H}_i - H_i\|_2 \leq \zeta_i$, where H_i is the true Hessian of the task loss. Because the Gauss-Newton proxy $\hat{H}_i$ is strictly positive semidefinite ($\hat{H}_i \succeq 0$), this condition is mathematically viable only if $\zeta_i \geq -\lambda_{\min}(H_i)$.

As shown in Figure 3, **the true Hessian is highly indefinite at every task boundary**, with $\lambda_{\min}(H_i)$ ranging between -9.6 and -17.6. While the buffer operator captures the correct overall scale ($\|\hat{H}_i\|_2$ between 25 and 60), its operator norm error spans 16 to 42—which is roughly $1,600\times$ to $4,200\times$ larger than the floor $\zeta_i = 0.01$ used during training. The inherent negative curvature of

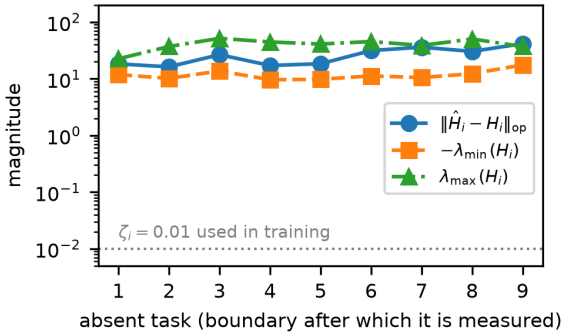

Figure 3: **Spectrum of the true Hessian and error of the buffer curvature operator** at the nine task boundaries, against the uncertainty floor $\zeta_i = 0.01$ used in training.

deep loss landscapes accounts for 29% to 65% of this error. Consequently, evaluating the certificate with a mathematically valid ζ_i would add at least $4.8\|\Delta\|^2$ to every player's robust slack penalty, fully explaining why theoretical guarantees in continual learning often demand highly restrictive update shrinkage. This identifies proxy unreliability as a profound structural limitation for deep CL algorithms that rely on strict mathematical bounding.

6.4 Continual-Learning Benchmarks

Finally, we assess how these mechanisms impact overall classification performance. Crucially, **the parameter τ_i explicitly models the stability-plasticity trade-off**, operating as a principled bound on degradation rather than a mere technical fix to prevent empty feasible sets. By assigning $d_i = -\tau_i$, the system tolerates a strictly controlled amount of damage to past tasks to accommodate the assimilation of new data.

Method	$\mathbf{A_{20}}$ **(%)**	**Forgetting**	**Worst-task (%)**	**Wall (s)**
Replay	68.3 ± 0.8	-0.007	53.4 ± 3.7	91
Averaged-gradient replay (A-GEM)	37.6 ± 3.1	$+0.172$	25.7 ± 4.8	195
Regularisation (EWC)	32.0 ± 1.6	$+0.182$	21.5 ± 0.7	688
First-order bargain, $d_i = -\tau_i$	62.4 ± 0.9	-0.033	49.8 ± 2.4	1144
First-order bargain, $d_i = 0$	61.3 ± 0.3	-0.028	49.8 ± 1.9	1063
This work (Curvature bargain)	$\mathbf{67.4 \pm 0.8}$	$\mathbf{-0.029}$	$\mathbf{56.9 \pm 1.8}$	11438

Table 3: CIFAR-100 (20 tasks of 5 classes). All methods share an identical reservoir buffer of 2000 examples. A_{20} denotes the mean accuracy over all tasks after the sequence concludes.

Table 3 positions the curvature bargain against established baselines at matched memory capacity. The empirical outcomes validate our theoretical framework:

- **Second-order vs. First-order:** The curvature bargain reaches 67.4%, consistently outperforming both first-order variants (62.4% and 61.3%) by a margin of 5 to 6 points across all seeds. This confirms that *even in deep neural networks, second-order utilities are indispensable for maintaining task protection* when gradients naturally diminish. Furthermore, incorporating the tolerable-forgetting budget ($d_i = -\tau_i$) yields a tangible improvement over the rigid zero-disagreement approach ($d_i = 0$), demonstrating the practical benefit of modeling controlled degradation.
- **Analytical Limit vs. Replay:** The curvature bargain does not outperform standard replay (68.3%). This result is scientifically consistent with the findings in Section 6.3: because the deep Hessian proxy contains massive, unavoidable inaccuracies, the bargaining framework cannot flawlessly execute its theoretically optimal updates. Thus, attempting to negotiate updates strictly through heavily distorted proxies offers no overall advantage over directly

minimizing the replayed loss, despite requiring a substantially higher computational overhead.

Ultimately, these experiments successfully validate the mechanics of asymmetric Nash bargaining in continual learning while honestly highlighting the exact boundary where theory breaks down in practice: the geometric fidelity of the deep learning proxy.

7 Discussion and Conclusion

We formulate continual learning with absent tasks as a robust bargain with quadratic utilities and fixed forgetting budgets. The analysis gives conditions for feasibility, boundedness, and uniqueness, and bounds sensitivity to derivative errors through surplus margins and curvature conditioning. This comparison requires an attained reference solution that is also proxy-feasible. Separately, the damage certificate bounds one-step loss increase without requiring a reference optimizer, provided the derivative-error and Hessian-Lipschitz bounds hold. The cumulative result requires these conditions at every step; it does not establish unconditional long-term retention.

The controlled experiments are consistent with the theoretical bounds. On the evaluated task-incremental CIFAR-100 stream, the curvature bargain achieves higher mean accuracy than the two first-order bargains, but remains below replay (67.4% versus 68.3%) at substantially greater computational cost. Its higher average worst-task accuracy is not consistent across seeds. With one dataset, three seeds, and one operating point per method, these results suggest a benefit of quadratic utilities within the tested bargaining formulations, but do not establish a general advantage over replay.

Limitations and Future Work. Bridging the gap between our theoretical framework and practical deep neural network regimes remains a critical challenge. First, the theoretical guarantees (particularly the damage certificate and sensitivity bounds) rely on strict assumptions regarding bounded curvature errors and reference quantities unavailable during training. As demonstrated in our empirical analysis, the negative spectrum of the true Hessian in large-scale networks frequently violates the curvature floors used in training. Future investigations must establish whether alternative network architectures or specialized regularization strategies can better satisfy these geometric prerequisites without sacrificing model plasticity.

Second, the reliable estimation of proxy errors, specifically the gradient uncertainty radius, is inherently difficult. Current heuristic estimators struggle to maintain strong correlations with the true proxy error across diverse optimization trajectories. Consequently, it remains unresolved whether the breakdown of theoretical coverage in real streams stems from the out-of-sample invalidity of telemetry-fitted constants or a fundamental inadequacy of the local quadratic approximation in highly non-convex landscapes. Developing robust estimators for these local geometric properties is a vital direction for future research.

Finally, our empirical scope is currently limited to a single dataset and three seeds, largely due to the computational overhead of incorporating second-order utilities. The cost of computing the curvature term scales quadratically with the number of bargained absent tasks. While exact forward-only evaluations can partially alleviate this burden, they do not change the underlying asymptotic complexity. For the bargaining framework to become a practical alternative to standard replay methods, future work must explore more efficient curvature approximations and selective task participation mechanisms that drastically reduce the computational footprint.

AI use statement

Generative AI tools assisted with language refinement, LaTeX editing, and the presentation of theoretical contributions, experimental findings, and limitations. The authors remain responsible for the correctness of the mathematical arguments, the experimental results, and the final manuscript.

Ethics statement

This work studies continual learning through mathematical analysis, synthetic experiments, and evaluation on the public CIFAR-100 benchmark. The reported experiments do not involve recruiting human participants or collecting new personal data. Although bargaining weights and forgetting

budgets express task-level priorities, they do not by themselves ensure fairness across individuals or demographic groups. Similarly, the loss certificates depend on the stated mathematical assumptions and do not by themselves guarantee reliable predictions under changing data distributions. Applications involving sensitive data would require separate consideration of privacy when retaining examples in replay buffers, together with evaluation of potential bias and predictive reliability.

Reproducibility statement

Sections 3–5 describe the bargaining formulation and theoretical results, with supporting derivations and proofs in Appendices B–E. Section 6 and Appendix F.1 specify the CIFAR-100 task construction, preprocessing, architecture, training settings, replay memory, baselines, random seeds, and evaluation metrics. Appendix F.1 also distinguishes the separate trajectory used for the true-Hessian diagnostics from the main benchmark configuration. Section 6.1 and Appendices F.3 and F.4 describe the controlled experiments, while Appendix F.2 details the probe-slice construction and calibration using preceding training steps. Additional uncertainty-estimator comparisons and measurements of the reduced dimension and computational cost appear in Appendices F.5 and F.6. Our source code is available at `https://anonymous.4open.science/r/CL-Nash-bargaining-8584/`

References

Pietro Buzzega, Matteo Boschini, Angelo Porrello, Davide Abati, and Simone Calderara. Dark experience for general continual learning: A strong, simple baseline. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Arslan Chaudhry, Marc'Aurelio Ranzato, Marcus Rohrbach, and Mohamed Elhoseiny. Efficient lifelong learning with A-GEM. In *International Conference on Learning Representations (ICLR)*, 2019.

Mehrdad Farajtabar, Navid Azizan, Alex Mott, and Ang Li. Orthogonal gradient descent for continual learning. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017.

Bo Liu, Xingchao Liu, Xiaojie Jin, Peter Stone, and Qiang Liu. Conflict-averse gradient descent for multi-task learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

David Lopez-Paz and Marc'Aurelio Ranzato. Gradient episodic memory for continual learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

John F. Nash. The bargaining problem. *Econometrica*, 18(2):155–162, 1950.

Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, and Ethan Fetaya. Multi-task learning as a bargaining game. In *International Conference on Machine Learning (ICML)*, 2022.

R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

Gobinda Saha, Isha Garg, and Kaushik Roy. Gradient projection memory for continual learning. In *International Conference on Learning Representations (ICLR)*, 2021.

Ozan Sener and Vladlen Koltun. Multi-task learning as multi-objective optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

Aviv Shamsian, Aviv Navon, Neta Glazer, Kenji Kawaguchi, Gal Chechik, and Ethan Fetaya. Auxiliary learning as an asymmetric bargaining game. In *International Conference on Machine Learning*, pp. 30689–30705. PMLR, 2023.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

A Notation and Discussion of Assumptions

B Feasibility and Bargaining Geometry

B.1 Proof of Theorem 1

Proof. We prove the three statements in order.

1. Since $\hat{H}_i$ is positive semidefinite and r_i, ζ_i are nonnegative, each surplus function $\hat{s}_i^R$ is continuous and concave. By continuity, each set $\{\Delta : \hat{s}_i^R(\Delta) > 0\}$ is open. Since there are finitely many players, their intersection $\hat{\mathcal{S}}$ is open.

   For any $\Delta^1, \Delta^2 \in \hat{\mathcal{S}}$ and $t \in [0,1]$, concavity gives

   $$\hat{s}_i^R\big(t\Delta^1 + (1-t)\Delta^2\big) \geq t\,\hat{s}_i^R(\Delta^1) + (1-t)\hat{s}_i^R(\Delta^2) > 0 \qquad \text{for every } i.$$

   Thus, $\hat{\mathcal{S}}$ is convex.

   Additionally, take any $y, y' \in \hat{\mathcal{U}}$. By definition, there exist updates $\Delta^1, \Delta^2 \in \mathbb{R}^p$ such that $y \leq \hat{u}^R(\Delta^1)$ and $y' \leq \hat{u}^R(\Delta^2)$. For every $t \in [0,1]$, componentwise concavity of $\hat{u}^R$ yields

   $$ty + (1-t)y' \leq t\hat{u}^R(\Delta^1) + (1-t)\hat{u}^R(\Delta^2) \leq \hat{u}^R(t\Delta^1 + (1-t)\Delta^2).$$

   Hence $ty + (1-t)y' \in \hat{\mathcal{U}}$, proving that $\hat{\mathcal{U}}$ is convex. Downward comprehensiveness follows directly from the definition.

2. Assume that $\hat{\mathcal{S}}$ is nonempty and bounded. To prove that the objective attains its maximum on $\hat{\mathcal{S}}$, we first maximize the continuous Nash product on the compact set $\overline{\hat{\mathcal{S}}}$ and then show that every maximizer lies in $\hat{\mathcal{S}}$.

   Choose $\Delta^+ \in \hat{\mathcal{S}}$ and define

   $$K = \{\Delta \in \mathbb{R}^p : \hat{s}_i^R(\Delta) \geq 0 \text{ for every } i\}.$$

   Continuity of the surplus function gives $\overline{\hat{\mathcal{S}}} \subseteq K$. Conversely, take any $\Delta \in K$. For every $t \in (0,1]$, concavity gives

   $$\hat{s}_i^R\big((1-t)\Delta + t\Delta^+\big) \geq (1-t)\hat{s}_i^R(\Delta) + t\,\hat{s}_i^R(\Delta^+) > 0 \qquad \text{for every } i.$$

   Thus $(1-t)\Delta + t\Delta^+ \in \hat{\mathcal{S}}$. Letting $t \downarrow 0$ shows that $\Delta \in \overline{\hat{\mathcal{S}}}$. Therefore, $K = \overline{\hat{\mathcal{S}}}$. Since $\hat{\mathcal{S}}$ is bounded, K is closed and bounded, and hence compact by the Heine–Borel theorem.

   To establish compactness of $\hat{\mathcal{U}}_d$, consider

   $$\mathcal{A} = \{(\Delta, y) \in K \times \mathbb{R}^n : d \leq y \leq \hat{u}^R(\Delta)\}.$$

   This set is closed by continuity of $\hat{u}^R$. It is bounded because K is compact, $\hat{u}^R$ is continuous on K, and $d \leq y \leq \hat{u}^R(\Delta)$. Hence $\mathcal{A}$ is compact. Its image under the continuous coordinate projection $(\Delta, y) \mapsto y$ is precisely $\hat{\mathcal{U}}_d$, so $\hat{\mathcal{U}}_d$ is compact. Moreover, $\hat{\mathcal{U}}_d$ is convex because it is the intersection of the convex set $\hat{\mathcal{U}}$ with $\{y : y \geq d\}$.

   Define $P(\Delta) = \prod_{i=0}^{n-1} \hat{s}_i^R(\Delta)^{\alpha_i}$ on K. Since the surpluses are continuous and nonnegative on K and $\alpha_i > 0$, P is continuous and attains its maximum on the nonempty compact set K by the extreme value theorem. Moreover, $P(\Delta^+) > 0$, while $P = 0$ on $K \setminus \hat{\mathcal{S}}$. Hence every maximizer lies in $\hat{\mathcal{S}}$.

Consequently, every maximizer belongs to $\hat{\mathcal{S}}$. On $\hat{\mathcal{S}}$, all logarithm arguments are positive and

$$\log P(\Delta) = \sum_{i=0}^{n-1} \alpha_i \log \hat{s}_i^R(\Delta) = \hat{\varphi}(\Delta).$$

Since the logarithm is strictly increasing, the logarithmic bargaining objective also attains its maximum.

We next establish uniqueness of the maximizing utility vector. Every $y \in \hat{\mathcal{U}}_d$ satisfies $d \le y \le \hat{u}^R(\Delta)$ for some $\Delta \in K$. Monotonicity of the Nash product in its nonnegative surpluses therefore gives

$$\prod_{i=0}^{n-1} (y_i - d_i)^{\alpha_i} \le P(\Delta).$$

Conversely, $\hat{u}^R(\Delta) \in \hat{\mathcal{U}}_d$ for every $\Delta \in K$. Hence the maximum Nash product over $\hat{\mathcal{U}}_d$ equals the maximum of P over K.

On the convex set $\hat{\mathcal{U}} \cap \{y : y > d\}$, define

$$f(y) = \sum_{i=0}^{n-1} \alpha_i \log(y_i - d_i).$$

Its Hessian is

$$\nabla^2 f(y) = -\operatorname{diag}\left(\frac{\alpha_i}{(y_i - d_i)^2} \right)_{i=0}^{n-1} \prec 0.$$

Thus f is strictly concave, which establishes uniqueness of its maximizing utility vector $y^\star$.

Since $y^\star > d$ and all weights are positive, any update satisfying $\hat{u}^R(\Delta) \ge y^\star$ must satisfy equality. Otherwise its utility vector would have a strictly larger Nash product. By the equivalence of the two maximization problems and uniqueness of the optimal utility vector, every maximizing update realizes $y^\star$.

3. Assume additionally that $\gamma > 0$. We prove uniqueness of the update by establishing strict concavity of $\hat{\varphi}$ on $\hat{\mathcal{S}}$.

   Take distinct $\Delta^1, \Delta^2 \in \hat{\mathcal{S}}$ and $t \in (0,1)$ and let $\Delta_t = t\Delta^1 + (1-t)\Delta^2$, where $t \in (0,1)$, and define

   $$c_i := (\Delta^1 - \Delta^2)^\top (\hat{H}_i + \zeta_i I)(\Delta^1 - \Delta^2) \ge 0.$$

   The nonnegativity follows from $\hat{H}_i + \zeta_i I \succeq 0$. Expanding the quadratic form gives

   $$\begin{aligned} \Delta_t^\top (\hat{H}_i + \zeta_i I)\Delta_t &= t^2 (\Delta^1)^\top (\hat{H}_i + \zeta_i I)\Delta^1 \\ &\quad + 2t(1-t)(\Delta^1)^\top (\hat{H}_i + \zeta_i I)\Delta^2 \\ &\quad + (1-t)^2 (\Delta^2)^\top (\hat{H}_i + \zeta_i I)\Delta^2 \\ &= t(\Delta^1)^\top (\hat{H}_i + \zeta_i I)\Delta^1 \\ &\quad + (1-t)(\Delta^2)^\top (\hat{H}_i + \zeta_i I)\Delta^2 - t(1-t)c_i. \end{aligned} \tag{5}$$

   Here, symmetry of $\hat{H}_i + \zeta_i I$ is used to combine the cross terms. Moreover, the triangle inequality and positive homogeneity of the Euclidean norm yield

   $$\|\Delta_t\|_2 \le t\|\Delta^1\|_2 + (1-t)\|\Delta^2\|_2.$$

Since $r_i \geq 0$, multiplying this inequality by $-r_i$ reverses its direction. Substituting it and equation 5 into the surplus gives

$$\begin{aligned}\hat{s}_i^R(\Delta_t) &= \tau_i - \langle \hat{g}_i, \Delta_t \rangle - \tfrac{1}{2}\Delta_t^\top(\hat{H}_i + \zeta_i I)\Delta_t - r_i\|\Delta_t\|_2 \\ &\geq t\Big[\tau_i - \langle \hat{g}_i, \Delta^1 \rangle - \tfrac{1}{2}(\Delta^1)^\top(\hat{H}_i + \zeta_i I)\Delta^1 - r_i\|\Delta^1\|_2\Big] \\ &\quad + (1-t)\Big[\tau_i - \langle \hat{g}_i, \Delta^2 \rangle - \tfrac{1}{2}(\Delta^2)^\top(\hat{H}_i + \zeta_i I)\Delta^2 - r_i\|\Delta^2\|_2\Big] \\ &\quad + \frac{t(1-t)}{2}c_i \\ &= t\,\hat{s}_i^R(\Delta^1) + (1-t)\hat{s}_i^R(\Delta^2) + \frac{t(1-t)}{2}c_i. \end{aligned} \tag{6}$$

Each quadratic form is nonnegative because $\hat{H}_i + \zeta_i I \succeq 0$. Moreover,

$$\begin{aligned}&\sum_{i=0}^{n-1}(\Delta^1 - \Delta^2)^\top(\hat{H}_i + \zeta_i I)(\Delta^1 - \Delta^2) \\ &\qquad = (\Delta^1 - \Delta^2)^\top(\bar{H} + ZI)(\Delta^1 - \Delta^2) \\ &\qquad \geq \gamma\|\Delta^1 - \Delta^2\|_2^2 > 0. \end{aligned} \tag{7}$$

The final inequality uses $\gamma > 0$ and $\Delta^1 \neq \Delta^2$. Therefore, at least one player has a strictly positive quadratic contribution.

Monotonicity and concavity of the logarithm consequently yield

$$\begin{aligned}\log \hat{s}_i^R(\Delta_t) &\geq \log\big(t\,\hat{s}_i^R(\Delta^1) + (1-t)\hat{s}_i^R(\Delta^2)\big) \\ &\geq t\log \hat{s}_i^R(\Delta^1) + (1-t)\log \hat{s}_i^R(\Delta^2). \end{aligned} \tag{8}$$

Multiplying by $\alpha_i > 0$ and summing gives

$$\hat{\varphi}(\Delta_t) > t\,\hat{\varphi}(\Delta^1) + (1-t)\hat{\varphi}(\Delta^2).$$

Thus $\hat{\varphi}$ is strictly concave on the convex domain $\hat{\mathcal{S}}$ and has at most one maximizer.

□

B.2 Strict Feasibility and Boundedness

Because we deferred the formal statements of boundedness from Section 4.1 to streamline the main text, we present them here alongside their proofs.

Proposition 1 (Strict feasibility)**.** *Under the standing assumptions, $\hat{\mathcal{S}} \neq \emptyset$ if and only if $\hat{g}_0 \neq 0$.*

Proposition 2 (Bounds on feasible updates)**.** *Under the standing assumptions, set $\bar{g} = \sum_i \hat{g}_i$ and $\bar{\tau} = \sum_i \tau_i$.*

1. *If $\gamma > 0$, then every $\Delta \in \hat{\mathcal{S}}$ satisfies*

$$\|\Delta\| < D_{\max} := \frac{\|\bar{g}\| + \sqrt{\|\bar{g}\|^2 + 2\gamma\bar{\tau}}}{\gamma}. \tag{9}$$

2. *For any player i with $r_i > \|\hat{g}_i\|$, independently of γ, every $\Delta \in \hat{\mathcal{S}}$ also satisfies $\|\Delta\| < \tau_i/(r_i - \|\hat{g}_i\|_2)$.*

Proof of Proposition 2. Let $\Delta \in \hat{\mathcal{S}}$ and $q = \|\Delta\|$. Summing the slacks and applying Cauchy–Schwarz and the minimum-eigenvalue bound gives

$$\begin{aligned}0 < \sum_i \hat{s}_i^R(\Delta) &= \bar{\tau} - \langle \bar{g}, \Delta \rangle - \tfrac{1}{2}\Delta^\top(\bar{H} + ZI)\Delta - \sum_i r_i q \\ &\leq \bar{\tau} + \|\bar{g}\|q - \tfrac{\gamma}{2}q^2. \end{aligned}$$

For $\gamma > 0$, the quadratic on the right has nonnegative root $D_{\max}$ and nonpositive root $(\|\bar{g}\| - \sqrt{\|\bar{g}\|^2 + 2\gamma\bar{\tau}})/\gamma$. Its strict positivity implies $q < D_{\max}$ and hence equation 9. The discriminant is nonnegative since $\bar{\tau} \geq 0$, and $\gamma > 0$ justifies division. If $\bar{\tau} = \|\bar{g}\| = 0$, strict positivity is impossible and the assertion is vacuous.

For the second bound, the individual feasibility inequality yields

$$\begin{aligned}\tau_i &> \langle \hat{g}_i, \Delta \rangle + \tfrac{1}{2}\Delta^\top(\hat{H}_i + \zeta_i I)\Delta + r_i q \\ &\geq (r_i - \|\hat{g}_i\|)q,\end{aligned}$$

using positive semidefiniteness and Cauchy–Schwarz. Division by $r_i - \|\hat{g}_i\| > 0$ proves the claim. Such a player is absent, since $r_0 = 0$, and therefore has $\tau_i > 0$. □

Proposition 3 (Exact boundedness)**.** *Suppose $\hat{\mathcal{S}} \neq \emptyset$ and define*

$$\mathcal{N} = \big\{ v \in \mathbb{R}^p : \ (\hat{H}_i + \zeta_i I)v = 0 \text{ and } \langle \hat{g}_i, v \rangle + r_i \|v\| \leq 0 \ \text{ for every } i \big\}.$$

Then $\mathcal{N}$ is the recession cone of $\overline{\hat{\mathcal{S}}}$, and $\hat{\mathcal{S}}$ is bounded if and only if $\mathcal{N} = \{0\}$. In particular, if $\zeta_i > 0$ for some i, or more generally if $\gamma > 0$, then $\hat{\mathcal{S}}$ is bounded.

Proof of Proposition 3. Write $A_i = \hat{H}_i + \zeta_i I \succeq 0$. As established in the proof of Theorem 1,

$$K := \{\Delta : \hat{s}_i^R(\Delta) \geq 0 \text{ for every } i\} = \overline{\hat{\mathcal{S}}}.$$

This identity uses only continuity and concavity of the surpluses and $\hat{\mathcal{S}} \neq \emptyset$. Thus K is nonempty, closed, and convex.

For $\Delta \in K$, $v \in \mathbb{R}^p$, and $t \geq 0$,

$$\begin{aligned}\hat{s}_i^R(\Delta + tv) = \hat{s}_i^R(\Delta) - t\big(\langle \hat{g}_i, v \rangle + \langle A_i\Delta, v \rangle\big) - \frac{t^2}{2} v^\top A_i v \\ - r_i\big(\|\Delta + tv\|_2 - \|\Delta\|_2\big).\end{aligned} \tag{10}$$

If $v \in \operatorname{rec}(K)$, this expression is nonnegative for every $\Delta \in K$ and $t \geq 0$. Since

$$\big|\|\Delta + tv\|_2 - \|\Delta\|_2\big| \leq t\|v\|_2,$$

division by t^2 and passage to the limit as $t \to \infty$ give $v^\top A_i v \leq 0$. Positive semidefiniteness therefore implies $A_i v = 0$. Dividing equation 10 by t and using $\|\Delta + tv\|_2/t \to \|v\|_2$ then yields

$$\langle \hat{g}_i, v \rangle + r_i\|v\|_2 \leq 0.$$

Thus $\operatorname{rec}(K) \subseteq \mathcal{N}$.

Conversely, if $v \in \mathcal{N}$, then $A_i v = 0$ for every i. The triangle inequality and $r_i \geq 0$ give, for every $\Delta \in K$ and $t \geq 0$,

$$\hat{s}_i^R(\Delta + tv) \geq \hat{s}_i^R(\Delta) - t\big(\langle \hat{g}_i, v \rangle + r_i\|v\|_2\big) \geq 0.$$

Hence $\Delta + tv \in K$, proving $\mathcal{N} \subseteq \operatorname{rec}(K)$. Consequently, $\operatorname{rec}(K) = \mathcal{N}$.

A nonempty closed convex set in $\mathbb{R}^p$ is bounded if and only if its recession cone is $\{0\}$ (Rockafellar, 1970). Moreover, $\hat{\mathcal{S}}$ is bounded if and only if its closure K is bounded: any closed ball containing $\hat{\mathcal{S}}$ also contains K, and the converse follows from $\hat{\mathcal{S}} \subseteq K$. Therefore, $\hat{\mathcal{S}}$ is bounded if and only if $\mathcal{N} = \{0\}$.

Finally, if $\gamma > 0$ and $v \in \mathcal{N}$, then

$$0 = \sum_i v^\top A_i v = v^\top(\bar{H} + ZI)v \geq \gamma\|v\|_2^2,$$

so $v = 0$. If some $\zeta_i > 0$, then $Z > 0$ and $\gamma = \lambda_{\min}(\bar{H}) + Z > 0$ under the standing assumptions. □

Corollary 2 (Uniform slack bound)**.** *Under the standing assumptions, suppose $\gamma > 0$ and let $D_{\max}$ be defined by equation 9. Then, for every $\Delta \in \hat{\mathcal{S}}$ and every $i \in \{0, \dots, n-1\}$,*

$$0 < \hat{s}_i^R(\Delta) \le U_{\max} := \max_{0 \le j \le n-1} \big(\tau_j + \|\hat{g}_j\| D_{\max}\big).$$

In particular, $U_{\max} > 0$ whenever $\hat{\mathcal{S}} \neq \varnothing$.

Proof of Corollary 2. The nonaffine terms in equation 2 are nonpositive. By Cauchy–Schwarz and Proposition 2,

$$\hat{s}_i^R(\Delta) \le \tau_i - \langle \hat{g}_i, \Delta \rangle \le \tau_i + \|\hat{g}_i\| \|\Delta\| \le \tau_i + \|\hat{g}_i\| D_{\max} \le U_{\max}.$$

Strict positivity follows from feasibility. Evaluation at any feasible point then gives $U_{\max} > 0$. □

B.3 Representer property

To ensure a positive aggregate curvature bound even when the proxy curvature matrices are singular, we impose the following sufficient condition.

Assumption 2 (Curvature-uncertainty floor)**.** There is a constant $\zeta_{\min} > 0$ with $\zeta_i \ge \zeta_{\min}$ for every player.

This assumption guarantees $\gamma > 0$ and, when $\hat{g}_0 \neq 0$, existence and uniqueness of the proxy bargaining update. The following proposition shows that this update lies in the subspace generated by the proxy gradients and the ranges of the proxy curvature matrices. Consequently, the bargaining problem can be restricted to this subspace without changing its solution.

Proposition 4. *Let Assumption 2 hold, let $\hat{g}_0 \neq 0$, and let $\Delta^\star$ be the maximizer of equation 3. Define*

$$\mathcal{V} = \operatorname{span}\{\hat{g}_0, \dots, \hat{g}_{n-1}\} + \sum_{i=0}^{n-1} \operatorname{range}(\hat{H}_i).$$

Then $\Delta^\star \in \mathcal{V}$, and $\Delta^\star$ is the maximizer of the restriction of $\hat{\varphi}$ to $\mathcal{V}$. Consequently $\dim \mathcal{V} \le n + \sum_i \operatorname{rank}(\hat{H}_i)$, and with $\hat{H}_i = J_i^\top J_i$ for an output Jacobian $J_i \in \mathbb{R}^{k_i \times p}$ on stored inputs, $\dim \mathcal{V} \le n + \sum_i k_i$, independently of p.

Proof. By Propositions 1 and 2 and Theorem 1, the maximizer exists, is unique, and lies in the open set $\hat{\mathcal{S}}$, which excludes the origin; so $\hat{\varphi}$ is differentiable at $\Delta^\star$ and $\nabla\hat{\varphi}(\Delta^\star) = 0$. Writing $\lambda_i = \alpha_i / \hat{s}_i^R(\Delta^\star) > 0$ and substituting $\nabla \hat{s}_i^R$ gives the stationarity condition,

$$\Big(\sum_i \lambda_i \hat{H}_i + \rho I\Big)\Delta^\star = -\sum_i \lambda_i \hat{g}_i, \qquad \rho = \frac{\sum_i \lambda_i r_i}{\|\Delta^\star\|} + \sum_i \lambda_i \zeta_i.$$

Assumption 2 gives $\rho \ge \sum_i \lambda_i \zeta_i > 0$, so

$$\Delta^\star = -\frac{1}{\rho}\Big(\sum_i \lambda_i \hat{g}_i + \sum_i \lambda_i \hat{H}_i \Delta^\star\Big) \in \operatorname{span}\{\hat{g}_i\} + \sum_i \operatorname{range}(\hat{H}_i) = \mathcal{V}.$$

For the second claim, note that $\mathcal{V}$ is invariant under the gradient map: for any $\Delta \in \mathcal{V} \cap \hat{\mathcal{S}}$,

$$\nabla\hat{\varphi}(\Delta) = -\sum_i \lambda_i(\Delta)\Big(\hat{g}_i + \hat{H}_i\Delta + r_i \tfrac{\Delta}{\|\Delta\|} + \zeta_i \Delta\Big) \in \mathcal{V},$$

since each summand lies in $\mathcal{V}$. Let $\Delta_{\mathcal{V}}$ maximize $\hat{\varphi}$ over $\mathcal{V} \cap \hat{\mathcal{S}}$, which is nonempty because $\Delta^\star \in \mathcal{V}$. First-order optimality within the subspace gives $P_{\mathcal{V}} \nabla\hat{\varphi}(\Delta_{\mathcal{V}}) = 0$ with $P_{\mathcal{V}}$ the orthogonal projector onto $\mathcal{V}$; since $\nabla\hat{\varphi}(\Delta_{\mathcal{V}}) \in \mathcal{V}$ we get $\nabla\hat{\varphi}(\Delta_{\mathcal{V}}) = 0$, so $\Delta_{\mathcal{V}}$ is a stationary point of the unrestricted concave program and hence equals $\Delta^\star$. The dimension counts are immediate. □

C Proofs of the Sensitivity Results

C.1 Proof of Theorem 2

Proof. We first establish strong concavity of the proxy objective and then use it to bound the difference between the proxy and reference maximizers.

Since $\hat{g}_0 \neq 0$ and $\gamma > 0$, Propositions 1 and 2 imply that $\hat{\mathcal{S}}$ is nonempty and bounded. Theorem 1 therefore gives a unique proxy maximizer $\hat{\Delta} \in \hat{\mathcal{S}}$. Corollary 2 also gives

$$0 < \hat{s}_i^R(\Delta) \leq U_{\max} < \infty \qquad \text{for every } \Delta \in \hat{\mathcal{S}} \text{ and every } i,$$

with $U_{\max} > 0$. Because $\tau_0 = 0$, both current-task surpluses vanish at the origin. Consequently, neither $\mathcal{S}$ nor $\hat{\mathcal{S}}$ contains zero. The norm terms are therefore twice continuously differentiable throughout these domains, and all logarithm arguments are positive. For every $\Delta \in \hat{\mathcal{S}}$, differentiation gives

$$\nabla \hat{s}_i^R(\Delta) = -\hat{g}_i - \hat{H}_i \Delta - r_i \frac{\Delta}{\|\Delta\|} - \zeta_i \Delta,$$

$$\nabla^2 \hat{s}_i^R(\Delta) = -\hat{H}_i - \zeta_i I - \frac{r_i}{\|\Delta\|} \left(I - \frac{\Delta\Delta^\top}{\|\Delta\|^2} \right).$$

All norm denominators are positive because $\Delta \neq 0$. The matrix in parentheses is the orthogonal projection onto the subspace perpendicular to Δ, and hence is positive semidefinite. Since $r_i \geq 0$,

$$\nabla^2 \hat{s}_i^R(\Delta) \preceq -(\hat{H}_i + \zeta_i I).$$

Applying the chain rule to the logarithmic objective yields

$$\begin{aligned} \nabla^2 \hat{\varphi}(\Delta) &= \sum_{i=0}^{n-1} \alpha_i \left[\frac{\nabla^2 \hat{s}_i^R(\Delta)}{\hat{s}_i^R(\Delta)} - \frac{\nabla \hat{s}_i^R(\Delta) \nabla \hat{s}_i^R(\Delta)^\top}{\left(\hat{s}_i^R(\Delta)\right)^2} \right] \\ &\preceq -\sum_{i=0}^{n-1} \frac{\alpha_i}{\hat{s}_i^R(\Delta)} (\hat{H}_i + \zeta_i I) \\ &\preceq -\frac{\alpha_{\min}}{U_{\max}} (\bar{H} + ZI). \end{aligned} \tag{11}$$

Finally, $\bar{H} + ZI \succeq \gamma I$ gives

$$\nabla^2 \hat{\varphi}(\Delta) \preceq -\frac{\alpha_{\min}\gamma}{U_{\max}} I = -\hat{m} I, \qquad \hat{m} = \frac{\alpha_{\min}\gamma}{U_{\max}} > 0.$$

By the Hessian criterion for strong concavity on an open convex domain, $\hat{\varphi}$ is $\hat{m}$-strongly concave on $\hat{\mathcal{S}}$.

Now let $\Delta^\star$ be any maximizer of φ on $\mathcal{S}$ satisfying the additional condition $\Delta^\star \in \hat{\mathcal{S}}$. Both objectives are differentiable on their open domains. Fermat's necessary condition gives

$$\nabla \varphi(\Delta^\star) = 0, \qquad \nabla \hat{\varphi}(\hat{\Delta}) = 0.$$

Strong concavity of $\hat{\varphi}$ implies strong monotonicity of $-\nabla\hat{\varphi}$. Because both $\Delta^\star$ and $\hat{\Delta}$ belong to $\hat{\mathcal{S}}$, this property and the Cauchy–Schwarz inequality give

$$\begin{aligned} \hat{m}\|\Delta^\star - \hat{\Delta}\|^2 &\leq -\langle \nabla\hat{\varphi}(\Delta^\star) - \nabla\hat{\varphi}(\hat{\Delta}), \Delta^\star - \hat{\Delta} \rangle \\ &= -\langle \nabla\hat{\varphi}(\Delta^\star) - \nabla\varphi(\Delta^\star), \Delta^\star - \hat{\Delta} \rangle \\ &\leq \|\nabla\hat{\varphi}(\Delta^\star) - \nabla\varphi(\Delta^\star)\| \, \|\Delta^\star - \hat{\Delta}\|. \end{aligned} \tag{12}$$

If $\Delta^\star = \hat{\Delta}$, the claimed bound holds immediately because its left-hand side is zero. Otherwise, $\hat{m}\|\Delta^\star - \hat{\Delta}\| > 0$, so division yields

$$\|\hat{\Delta} - \Delta^\star\| \leq \frac{1}{\hat{m}} \|\nabla\hat{\varphi}(\Delta^\star) - \nabla\varphi(\Delta^\star)\|.$$

Only stationarity of the reference maximizer was used. Neither its uniqueness nor strong concavity of φ is required. □

C.2 Proof of Corollary 1

Proof. All estimates are evaluated at $\Delta^\star$. Subtracting the oracle penalties from equation 2 cancels the identical regularizers. Cauchy–Schwarz and the operator-norm estimate yield

$$|\hat{s}_i^R - s_i^R| \le \|\hat{g}_i - g_i\| D + \tfrac{1}{2}\|\hat{H}_i - H_i\|_2 D^2 \le r_i D + \tfrac{1}{2}\zeta_i D^2,$$
$$\|\nabla \hat{s}_i^R - \nabla s_i^R\| \le \|\hat{g}_i - g_i\| + \|\hat{H}_i - H_i\|_2 D \le r_i + \zeta_i D.$$

The oracle slack gradient is

$$\nabla s_i^R(\Delta^\star) = -g_i - H_i\Delta^\star - r_i\frac{\Delta^\star}{D} - \zeta_i\Delta^\star,$$

and hence its norm is at most $v_i + r_i + \zeta_i D$ by the triangle inequality. Here $D > 0$ since the origin is excluded from $\mathcal{S}$.

The exact decomposition

$$\nabla\hat{\varphi} - \nabla\varphi = \sum_i \alpha_i \left[\frac{\nabla \hat{s}_i^R - \nabla s_i^R}{\hat{s}_i^R} + \nabla s_i^R \frac{s_i^R - \hat{s}_i^R}{\hat{s}_i^R s_i^R}\right]$$

and the positive bounds $\hat{s}_i^R \ge \hat{\mu}$ and $s_i^R \ge \mu$ give the sum in the explicit certificate as an upper bound on the gradient discrepancy. Substitution into equation 4, using $1/\hat{m} = U_{\max}/(\alpha_{\min}\gamma)$, proves the estimate. Retaining the exact oracle slack-gradient norm proves the final assertion. □

D Proof of the Forgetting Bound

D.1 Proof of Theorem 3

Proof. We estimate the Taylor remainder and compare q_i at the two updates. Taylor's theorem with integral remainder gives

$$F_i - q_i(\hat{\Delta}) = \int_0^1 (1-t)\hat{\Delta}^\top\big[\nabla^2 L_i(\theta + t\hat{\Delta}) - H_i\big]\hat{\Delta}\, dt,$$
$$|F_i - q_i(\hat{\Delta})| \le \rho_i\|\hat{\Delta}\|^3 \int_0^1 t(1-t)\, dt = \tfrac{\rho_i}{6}\|\hat{\Delta}\|^3. \tag{13}$$

The inequality follows from the Hessian Lipschitz bound and $|x^\top A x| \le \|A\|_2\|x\|^2$. It includes the case $\hat{\Delta} = 0$.

The function $\nabla q_i(\Delta) = g_i + H_i\Delta$ is continuous, so V_i is finite on the compact segment $[\Delta^\star, \hat{\Delta}]$. The fundamental theorem of calculus and Cauchy–Schwarz imply

$$q_i(\hat{\Delta}) - q_i(\Delta^\star) = \int_0^1 \langle g_i + H_i\big(\Delta^\star + t(\hat{\Delta} - \Delta^\star)\big), \hat{\Delta} - \Delta^\star\rangle\, dt$$
$$\le V_i\|\hat{\Delta} - \Delta^\star\|.$$

If the updates coincide, both sides of the difference estimate are zero. The oracle slack identity is

$$q_i(\Delta^\star) = \tau_i - s_i^R(\Delta^\star) - r_i D - \tfrac{\zeta_i}{2}D^2. \tag{14}$$

Combining these estimates proves the sharp forgetting bound. Finally, $s_i^R(\Delta^\star) \ge \mu > 0$ and $r_i, \zeta_i \ge 0$ give the simplified estimate. □

D.2 Proof of Theorem 4

Proof. Estimate equation 13 depends only on local loss regularity. By Cauchy–Schwarz and the operator-norm inequality,

$$q_i(\hat{\Delta}) \le \langle \hat{g}_i, \hat{\Delta}\rangle + \tfrac{1}{2}\hat{\Delta}^\top\hat{H}_i\hat{\Delta} + r_i\|\hat{\Delta}\| + \tfrac{\zeta_i}{2}\|\hat{\Delta}\|^2$$
$$= \tau_i - \hat{s}_i^R(\hat{\Delta}).$$

Combining this inequality with equation 13 proves the computable certificate. Feasibility gives $\hat{s}_i^R(\hat{\Delta}) > 0$; the argument involves no division by a slack or curvature constant. □

D.3 Certified damage under shrinkage

We explicitly state and prove the admission control proposition here, mapping to the summarized control mechanism discussed in Section 5.

Proposition 5 (Certified damage under shrinkage)**.** *Let $\hat{\Delta} \in \hat{\mathcal{S}}$ and $\beta \in [0,1]$. Then $\beta\hat{\Delta} \in \hat{\mathcal{S}}$ and*

$$\mathcal{C}_i(\beta\hat{\Delta}) \;\le\; \beta\,\mathcal{C}_i(\hat{\Delta}) \qquad \text{for every } i.$$

Consequently, if $\hat{\Delta}$ is strictly feasible and

$$\beta^\star = \min\Big\{1,\; \min_{i:\mathcal{C}_i(\hat{\Delta})>\tau_i} \frac{\tau_i}{\mathcal{C}_i(\hat{\Delta})}\Big\},$$

then the admitted update $\beta^\star\hat{\Delta}$ satisfies $F_i \le \mathcal{C}_i(\beta^\star\hat{\Delta}) \le \tau_i$ for every i.

Proof of Proposition 5. Concavity of $\hat{s}_i^R$ with $\hat{s}_i^R(0) = \tau_i$ gives $\hat{s}_i^R(\beta\hat{\Delta}) \ge (1-\beta)\tau_i + \beta\hat{s}_i^R(\hat{\Delta}) > 0$, which proves feasibility and, rearranged, $\tau_i - \hat{s}_i^R(\beta\hat{\Delta}) \le \beta\big(\tau_i - \hat{s}_i^R(\hat{\Delta})\big)$. The remainder term scales as $\beta^3 \le \beta$ with a nonnegative coefficient. Adding the two estimates gives $\mathcal{C}_i(\beta\hat{\Delta}) \le \beta\mathcal{C}_i(\hat{\Delta})$. For the consequence, fix i. If $\mathcal{C}_i(\hat{\Delta}) \le \tau_i$ then $\mathcal{C}_i(\beta^\star\hat{\Delta}) \le \beta^\star\mathcal{C}_i(\hat{\Delta}) \le \max\{\mathcal{C}_i(\hat{\Delta}), 0\} \le \tau_i$, using $\tau_i > 0$ and $\beta^\star \in (0,1]$. Otherwise $\mathcal{C}_i(\hat{\Delta}) > \tau_i > 0$ and $\beta^\star \le \tau_i/\mathcal{C}_i(\hat{\Delta})$, so $\mathcal{C}_i(\beta^\star\hat{\Delta}) \le \beta^\star\mathcal{C}_i(\hat{\Delta}) \le \tau_i$. Theorem 4 applies at $\beta^\star\hat{\Delta}$ because it is strictly feasible. □

D.4 Cumulative loss and budget allocation

Corollary 3. *Fix an absent task i, an integer $T_i \ge 1$, and updates $\theta_{t+1} = \theta_t + \hat{\Delta}_t$ for $t = 0, \dots, T_i - 1$. Suppose Theorem 4 applies at every step to the same function L_i, with positive budgets $\tau_i^{(t)}$, proxy slacks $\hat{s}_{i,t}^R$, and Hessian Lipschitz constants $\rho_{i,t} \ge 0$. Define*

$$F_{i,t} = L_i(\theta_{t+1}) - L_i(\theta_t), \qquad R_{i,t} = \tfrac{\rho_{i,t}}{6}\|\hat{\Delta}_t\|^3.$$

Then, for every $k = 1, \dots, T_i$,

$$L_i(\theta_k) - L_i(\theta_0) = \sum_{t=0}^{k-1} F_{i,t} \le \sum_{t=0}^{k-1} \Big[\tau_i^{(t)} - \hat{s}_{i,t}^R(\hat{\Delta}_t) + R_{i,t}\Big]. \tag{15}$$

For $B_i > 0$, the sufficient conditions

$$\sum_{t=0}^{T_i-1} \tau_i^{(t)} \le B_i, \qquad R_{i,t} \le \hat{s}_{i,t}^R(\hat{\Delta}_t) \quad (t = 0, \dots, T_i - 1) \tag{16}$$

imply $L_i(\theta_k) - L_i(\theta_0) \le B_i$ for every $k = 1, \dots, T_i$. The allocation $\tau_i^{(t)} = B_i/T_i$ satisfies the first condition.

Proof. Applying Theorem 4 at each step gives $F_{i,t} \le \tau_i^{(t)} - \hat{s}_{i,t}^R(\hat{\Delta}_t) + R_{i,t}$. Summation and telescoping yield equation 15. Under equation 16, each summand is at most $\tau_i^{(t)}$, so positivity of the budgets implies

$$L_i(\theta_k) - L_i(\theta_0) \le \sum_{t=0}^{k-1} \tau_i^{(t)} \le \sum_{t=0}^{T_i-1} \tau_i^{(t)} \le B_i.$$

Since $T_i \ge 1$ and $B_i > 0$, the equal allocation is well defined, positive, and sums to B_i. □

E Gram Matrix and First-Order Utility

This appendix examines whether a bargain with first-order utilities alone reaches an outcome that protects absent tasks. It states the first-order program, derives three of its properties from the Gram matrix of the players' gradients, and tests each property on the quadratic sandbox of Appendix F.3.

The first-order program. Setting $\hat{H}_i = 0$ and $\zeta_i = 0$ gives the utilities $u_i(\Delta) = -\langle \hat{g}_i, \Delta \rangle$. Without curvature the positive-surplus domain is in general unbounded, and first-order bargaining methods impose a trust region instead:

$$\max_{\|\Delta\| \le \epsilon} \sum_{i=0}^{n-1} \alpha_i \log\big(\tau_i - \langle \hat{g}_i, \Delta \rangle\big), \tag{17}$$

with $\tau_0 = 0$ and either $\tau_i > 0$ or $\tau_i = 0$ for the absent players, the latter being the disagreement point of Nash-MTL. Let $G \in \mathbb{R}^{n \times p}$ have rows $\hat{g}_i$ and $K = GG^\top$. Stationarity of equation 17 gives $\Delta = -c \sum_i \lambda_i \hat{g}_i$ with $\lambda_i = \alpha_i / (\tau_i - \langle \hat{g}_i, \Delta \rangle)$ and $c \ge 0$, so $\Delta = G^\top \beta$, and since $\langle \hat{g}_i, G^\top \beta \rangle = (K\beta)_i$ and $\|G^\top \beta\|^2 = \beta^\top K \beta$, the program depends on the data only through K.

Lemma 1 (Gram conditioning)**.** $\lambda_{\min}(K) \le \min_i \|\hat{g}_i\|^2$.

Proof. $e_i^\top K e_i = \|\hat{g}_i\|^2$ for the i-th coordinate vector, and $\lambda_{\min}(K) \le e_i^\top K e_i$. □

Lemma 2 (Direction of a proxy gradient)**.** *If $\|\hat{g}_i - g_i\| \le r_i < \|g_i\|$, then $\sin \angle(\hat{g}_i, g_i) \le r_i / \|g_i\|$. If $r_i > \|g_i\|$, every direction in $\mathbb{R}^p$ is the direction of some $\hat{g}_i$ with $\|\hat{g}_i - g_i\| \le r_i$.*

Proof. The ball of radius r_i about g_i is seen from the origin within the cone of half-angle $\arcsin(r_i / \|g_i\|)$ when $r_i < \|g_i\|$. When $r_i > \|g_i\|$ the ball contains a neighbourhood of the origin, hence a point in every direction. □

Lemma 3 (Individual rationality at $d_i = 0$)**.** *With $\tau_i = 0$ for every player, the set $\{\Delta : \langle \hat{g}_i, \Delta \rangle < 0$ for all $i\}$ is nonempty if and only if $0 \notin \operatorname{conv}\{\hat{g}_0, \dots, \hat{g}_{n-1}\}$.*

Proof. This is Gordan's theorem of the alternative: either $G\Delta < 0$ has a solution, or there is $y \ge 0$, $y \ne 0$, with $G^\top y = 0$, and normalising y to sum to one exhibits 0 as a convex combination of the rows of G. □

The three lemmas describe how the first-order bargain behaves as an absent task converges. Its conditioning constant $\lambda_{\min}(K)$ vanishes at least as fast as $\|\hat{g}_i\|^2$ (Lemma 1), whereas $\gamma = \lambda_{\min}(\bar{H}) + Z$ does not involve the gradients. When the proxy error does not vanish with the gradient, as Assumption 1 states for a buffer, the half-space the bargain protects is eventually determined by the error rather than by the task (Lemma 2). With $d_i = 0$ the bargain has no individually rational update whenever the proxy gradients surround the origin (Lemma 3).

F Experimental Protocols and Additional Results

F.1 Real-Stream Protocol

Stream. CIFAR-100 is split into 20 tasks of 5 classes. The class order is a permutation drawn from the seed, so every method at a given seed sees the same partition. Each task has 2500 training and 500 test images and occupies five logits of a single 100-way head. Training is task incremental: the cross-entropy of an example is taken over the five logits of its own task alone, at training and at evaluation time, which is what the reported accuracies measure. Inputs are normalised by the dataset statistics and, during training only, randomly cropped from a 4-pixel padding and flipped horizontally.

Loop. Tasks arrive in order. For each task the learner makes five passes over its examples in minibatches of 32, a fixed 395 steps per task and 7900 over the stream. After the last step on a task, its examples are offered to a reservoir buffer holding 2000 in total, so the buffer stays a uniform sample of everything seen. The accuracy of the just-finished task is recorded at that moment, and is the reference the forgetting measure subtracts from. Nothing is reset between tasks.

Methods. Replay concatenates a buffer minibatch of 32 to the current one and takes one step on the joint loss. The averaged-gradient method uses the buffer only to form a reference gradient and projects the current gradient onto the halfspace of non-negative inner product with it when they conflict. The regularisation baseline adds a diagonal empirical Fisher penalty, averaged over retained tasks, with each task's Fisher estimated once at its boundary on 300 of its examples. The bargains give the live task one player and each buffered task one more, take each player's gradient on 32 of its own buffered examples restricted to its own classes, and solve their program at every step; a step at which the solver returns no feasible update applies no update and is counted as skipped.

Measurement. After the last task, each task's test accuracy is measured on its own 500 images over its own five logits. A_{20} is the mean of those twenty numbers, forgetting is the mean over the first nineteen of the accuracy each had at its own boundary minus the accuracy it has at the end, and the worst-task accuracy is the minimum of the twenty. Each configuration is run at three seeds, which fix the class order, the initialisation and the batch order together.

Item	Setting
Stream	CIFAR-100 in 20 tasks of 5 classes, class order permuted per seed
Backbone	reduced ResNet-18, GroupNorm, width 16 (7.1×10^5 parameters), one 100-way head
Training	task incremental, 5 epochs per task, minibatch 32, 395 steps per task
Augmentation	random crop from 4-pixel padding, horizontal flip
Memory	reservoir buffer of 2000 examples for every method
Replay, A-GEM	SGD, learning rate 0.1, replay minibatch 32
EWC	SGD, learning rate 0.1, $\lambda = 3$ averaged over retained tasks, Fisher on 300 examples per task
Bargains	one player per task, gradients on 32 buffered examples, uniform α_i, $\tau_i = 0.05$, $r_i = 0$
Curvature	weighted Gauss-Newton on the player's own examples and classes, $\zeta_i = 0.01$
Certificate	probe slice one tenth of each buffer, window 30 steps, fixed-constant ablation at 0.05
Seeds	0 to 2

Table 4: Protocol of the real-stream experiments.

The robust gradient term is not used in training ($r_i = 0$); gradient-proxy error enters only through the calibrated certificate below. The true-Hessian measurement runs on a separate trajectory, trained by replay on the ten-task imbalanced split of CIFAR-100 at width 20 with a 500-example buffer and five epochs per task, with the true Hessian of each task taken on 512 of its retained training examples.

F.2 Probe Slice and Calibration of the Certificate

Theorem 4 is conditional on (r_i, ζ_i, ρ_i) being valid at the current step. Estimating them on the replay data would underestimate them, since the optimiser has been pushed to make the loss small on exactly those examples, so a disjoint slice of each buffer is held out from the replay loss.

Definition 3 (Probe slice)**.** For each absent task i, the stored examples are partitioned into a replay slice, which enters the replay loss, and a probe slice $\mathcal{P}_i$, which never does. Write $L_i^{\mathcal{P}}$ for the loss of task i on $\mathcal{P}_i$.

At step t the telemetry residual

$$\epsilon_{i,t} = \left[L_i^{\mathcal{P}}(\theta_{t+1}) - L_i^{\mathcal{P}}(\theta_t)\right] - \left[\langle \hat{g}_i, \hat{\Delta}_t \rangle + \tfrac{1}{2}\hat{\Delta}_t^\top \hat{H}_i \hat{\Delta}_t\right] \tag{18}$$

aggregates the gradient-proxy error, the curvature-proxy error and the failure of the quadratic model over the realised step.

Proposition 6 (Certificate calibration)**.** *Fix a window W of past steps and let $q_{i,t} = \|\hat{\Delta}_t\|$. Any $(r_i, \zeta_i, \rho_i) \geq 0$ satisfying*

$$r_i q_{i,t} + \tfrac{\zeta_i}{2} q_{i,t}^2 + \tfrac{\rho_i}{6} q_{i,t}^3 \geq \epsilon_{i,t} \qquad (t \in W) \tag{19}$$

makes the certificate of Theorem 4 hold on W with F_i replaced by the probe-loss increase. Among feasible triples, the one minimising $\sum_{t \in W} \left(r_i q_{i,t} + \frac{\zeta_i}{2} q_{i,t}^2 + \frac{\rho_i}{6} q_{i,t}^3\right)$ solves a three-variable linear program and gives the tightest such certificate on W.

Proof. The certified damage at step t is $\langle \hat{g}_i, \hat{\Delta}_t \rangle + \frac{1}{2}\hat{\Delta}_t^\top \hat{H}_i \hat{\Delta}_t + r_i q_{i,t} + \frac{\zeta_i}{2} q_{i,t}^2 + \frac{\rho_i}{6} q_{i,t}^3$, which by equation 18 is at least the probe-loss increase exactly when equation 19 holds. Objective and constraints are linear in (r_i, ζ_i, ρ_i), and the feasible set is nonempty because a large r_i satisfies every constraint whenever $q_{i,t} > 0$. □

The fit is in sample: validity on W does not imply validity at step $t+1$, which is why the certificate scores each step with the fit from the preceding window only. In the experiments the probe slice is a tenth of each task's buffered examples, the window is 30 steps, the fixed-constant ablation uses $r_i = \zeta_i = \rho_i = 0.05$, and admission control uses the calibrated certificate of the absent players only, since $\tau_0 = 0$.

F.3 Quadratic Sandbox

Task i has curvature $H_i = A^\top A/\mathrm{rank}$ with A Gaussian and N examples drawn from $\mathcal{N}(c_i, S_i)$, so that $g_i(\theta) = H_i(\theta - \bar{z}_i)$ is exact and a buffer's proxy error is $H_i(\bar{z}_i^B - \bar{z}_i)$ in closed form. Every quantity of the theory, g_i, H_i, r_i, ζ_i and the realised forgetting, is therefore available, and the certificate can be compared against the quantity it certifies.

Result	Check	Outcome
Theorem 1, concavity	largest Hessian eigenvalue of $\hat{\varphi}$ at 1000 points	-1.4×10^{-2}
Proposition 1	strictly feasible start found, 5000 instances, $n \in \{2, \ldots, 10\}$	**100%** at every n
Proposition 2(2)	$\|\Delta^\star\| < \tau_i/(r_i - \|\hat{g}_i\|)$ where $r_i > \|\hat{g}_i\|$, 33 instances	**no violation**

Table 5: Structural results on random instances with $p \in \{2, \ldots, 11\}$. The bound $\|\Delta^\star\| < \tau_i/r_i$ without the gradient term, which the proposition does not state, is violated on 4 of 950 instances, so the $\|\hat{g}_i\|$ in its denominator is strictly necessary.

r	Median Deviation	Median Bound	Median Ratio	Violations
10^{-4}	1.39×10^{-4}	1.98×10^{-3}	13.4	**0**/150
10^{-3}	1.48×10^{-3}	1.86×10^{-2}	12.4	**0**/150
10^{-2}	1.44×10^{-2}	1.98×10^{-1}	13.1	**0**/150
3×10^{-2}	4.86×10^{-2}	5.86×10^{-1}	12.1	**0**/150

Table 6: The explicit certificate against the measured deviation. Its ratio to the deviation is constant in r, and the fitted log-log slope of the deviation in r is 1.02.

Sweep	Values	Median Cert. / Dev.	Solved
$\mathrm{rank}(H_i)$ at $\zeta = 0$	$2, 4$ ($\lambda_{\min}(\bar{H}) = 0$)	none	0/240
	$8, 16, 24$ ($\lambda_{\min}(\bar{H}) = 0.35, 1.18, 1.72$)	13.6, 5.8, **4.3**	**360**/360
ζ at $\lambda_{\min}(\bar{H}) = 0$	$10^{-3}, 10^{-2}, 5 \times 10^{-2}, 10^{-1}, 5 \times 10^{-1}$	2.80, 2.80, 2.76, 2.77, **2.18**	**600**/600
n	$2, 5, 10, 15, 20$	**3.1**, 11.5, 11.2, 11.8, 11.2	**500**/500

Table 7: Conditioning of the sensitivity certificate. Without a floor and with $\lambda_{\min}(\bar{H}) = 0$ no certificate exists, as Assumption 2 anticipates; with $\lambda_{\min}(\bar{H}) = 0$ held at zero, Z alone keeps the ratio flat over a $500\times$ range of ζ.

F.4 Forgetting Decomposition with Known Constants

Binary logistic regression over $d = 20$ features with four players, three of them absent and represented by a 40-example buffer, gives a loss that is not its own second-order model while every constant is exact. $H_i = \frac{1}{N}\sum_j \sigma_j(1-\sigma_j)x_j x_j^\top$ is positive semidefinite and is the true Hessian, so the only curvature error is the one caused by estimating it from the buffer; $\rho_i \le B^3/(6\sqrt{3})$ with $B = \max_j \|x_j\|$, since $\sup_z |\sigma''(z)| = 1/(6\sqrt{3})$; and V_i is exact, because the norm of an affine function attains its supremum over a segment at an endpoint. The budget sets the step size, so sweeping it exercises the third-order remainder; 60 instances per budget.

τ	$\|\hat{\Delta}\|$	Budget Term	Discrepancy Term	Remainder Share	Bound Holds	Coverage, True r_i	Coverage, Split-Half r_i
0.02	0.09	1.6×10^{-2}	2.1×10^{-3}	0.07%	**100%**	**100%**	98%
0.1	0.38	8.4×10^{-2}	7.2×10^{-3}	0.94%	**100%**	**100%**	**100%**
0.5	1.21	0.45	1.6×10^{-2}	5.8%	**100%**	**100%**	**100%**
2	2.56	1.90	1.8×10^{-2}	12.3%	**100%**	**100%**	**100%**
8	3.97	7.87	9.9×10^{-3}	11.3%	**100%**	**100%**	**100%**

Table 8: One-step loss bounds with every constant known. The median margin of the certificate is $0.18\tau_i$ to $0.25\tau_i$. The split-half radius estimate underestimates the true radius on 83% of instances and still leaves coverage at 98% or above.

F.5 Curvature and Radius Estimators

On the replay trajectory of Section 6.3, three buffer curvature estimators were compared with the Gauss-Newton matrix on each task's full retained training data at the first seven boundaries. All three track the target across boundaries (Spearman correlation 1.00 for the Gauss-Newton and function-space Jacobian estimators, 0.93 for a stale diagonal Fisher). The weighted Gauss-Newton estimator has a median relative error of 0.05 to 0.16.

Estimator of r_i	Sandbox, Unbiased Buffer	Sandbox, Biased Buffer	Real Trajectory
Split-half buffer disagreement	0.227	0.402	**0.667**
Parameter drift	0.045		0.533
Telemetry fit, Proposition 6			0.550

Table 9: Spearman correlation of radius estimators with the true proxy error; on the real trajectory the true error is the distance to the gradient on the task's full retained data, at nine boundaries.

F.6 Reduced Dimension and Cost

Examples / Player	Bound $n + \sum_i k_i$	$\dim \mathcal{V}$	Smallest / Largest S.V.	s / Newton iter.	Peak Memory (MB)
1	505	505	1.5×10^{-2}	**1.3**	**43**
2	1005	1005	7.3×10^{-3}	1.8	170
4	2005	2005	4.3×10^{-3}	3.4	676
8	4005	4005	2.4×10^{-3}	3.7	2688

Table 10: Dimension of the exact subspace on the reduced ResNet-18 at width 20 with five absent tasks and a 100-class head. The rank is taken at the smallest relative threshold at which a duplicated copy of one player's rows adds no rank.

On a two-layer network with $p \approx 212$, where the exact subspace can be formed explicitly, the gradient-span path fits a wall-clock exponent of 1.44 in n and the exact path an exponent of 1.09, at $42\times$ to $88\times$ the absolute cost, because forming the reduced matrices column by column costs $n \cdot \dim \mathcal{V}$ Jacobian-vector products.