

Untitled

QNFO

Abstract

An agentic buyer with a hard fulfillment deadline can fork a procurement negotiation into many parallel seller-facing threads, but every thread consumes resources and every simultaneous acceptance creates a cancellation and commitment liability. We study a planner that jointly chooses the number of parallel negotiators n and a common procurement price cap p , coupling a product-specific acceptance curve $a(p)$ with fulfillment loss, per-thread cost, and excess-commitment cost. We derive three structural results with fully explicit arithmetic. First, holding the acceptance target fixed, the marginal value of an additional negotiator decays geometrically at rate $(1 - a)$, so the optimal team size follows a closed-form ceiling rule. Second, with an exponential acceptance curve the interior optimal cap satisfies $p^*(n) = c + \ln((L - R - e)/e)/((n - 1)\lambda)$, which is decreasing in n : more threads let the buyer quote a lower, more aggressive cap because fulfillment risk is diversified — the precise sense in which parallelism substitutes for concession. Solving the fixed point over $n \geq 2$ for a worked configuration yields $(n^*, p^*) = (2, 13.89182)$ with expected net cost 71.28572. This is an interior candidate, not the global optimum: for $n = 1$ the cost $C(1, p) = 96 - 30a(p)$ is strictly decreasing in a on $[0, 1)$, so under the unbounded exponential curve the single-thread cost has infimum 66 (not attained) and already equals $69.0 < 71.28572$ at $a = 0.9$. Concurrency pays only when the cap is bounded or failure risk is material at the prevailing cap. Third, we conjecture — without proof — that under dispersed seller offers the marginal value of a thread may decay more slowly than geometrically, which would widen the profitable concurrency region; this is a projection, not a derived result (Section 4.6). We document all points where the source drafts disagreed on modeling convention and state the convention adopted. The results give agentic procurement systems a transparent, auditable rule for sizing negotiation fan-out.

Parallelism or Concession? A Reconciled Analytical Model of Concurrency-Aware Procurement Negotiation for Agentic Commerce

1 Introduction

Agentic commerce systems increasingly delegate procurement to autonomous buyers that negotiate with sellers over price, delivery, and terms. A natural capability of such systems is *forking*: the buyer spawns n parallel seller-facing negotiation threads, each quoting against a common price cap, and accepts the first successful outcome. Forking is cheap in compute but not in commitment: each additional thread consumes a per-thread resource budget, and if several threads accept simultaneously the buyer holds excess commitments that must be cancelled at a cost, both financial and reputational.

This tension is the subject of the concurrency-aware procurement model of [1],[2], which established three structural results: geometric decay of the marginal value of negotiators, substitution of parallelism for concession under convex acceptance curves, and a dispersion comparative static under which buyers benefit from searching harder rather than offering higher prices. This paper reconstructs that model from first principles so that every structural claim is auditable without access to the original implementation, derives the concurrency threshold and the optimal-cap formula in closed form, works a complete numerical calibration with every arithmetic step shown, and extends the analysis to a dispersed-offer regime.

The paper is a reconciliation of three independent drafts. Where the drafts agreed, we state the convergent result with the best-substantiated derivation. Where they disagreed — on the cost accounting convention, on the acceptance-curve family, and on the direction of the cap–concurrency relationship — we adopt one convention for the main text and document each conflict explicitly in Appendix A rather than resolving it silently.

Section 2 situates the problem in the negotiation, agentic-commerce, and procurement literatures. Section 3 specifies the model and the adopted conventions. Section 4 contains the derivations and the full numerical calibration. Section 5 reports results. Section 6 discusses limitations and falsification conditions, and Section 7 concludes.

2 Background and Related Work

Concurrency-aware procurement negotiation. The direct antecedent is the Concurrency-Aware Negotiation Optimizer (CANO) of [1],[2], which defines the one-unit, single-hard-deadline sourcing problem with fulfillment loss, per-thread cost, and excess-commitment cost; proves geometric marginal-value decay, the parallelism-for-concession substitution under convex quantile curves, and the dispersion comparative static; and reports validation of the resulting optimizer against heuristic policies. Our paper adopts its primitives, reproduces the structural results with explicit arithmetic, and adds a closed-form optimal-cap formula and a dispersed-offer extension.

Negotiation as an optimization mechanism. Nonlinear negotiation approaches for complex-network optimization [3] recast Wi-Fi channel assignment in congested unlicensed spectrum as a graph-coloring problem over multilayer graphs, solved through negotiation-inspired dynamics. The relevance is structural rather than economic: negotiation abstractions can serve as optimization devices in engineering domains, prefiguring the use of negotiation fan-out as a controlled design parameter rather than a social ritual.

LLM-mediated bargaining. AgenticPay [4] provides a benchmark and system framework for multi-agent buyer–seller negotiation driven by large language models, supplying the motivation for treating per-thread acceptance as a stochastic, price-dependent quantity $a(p)$ rather than a deterministic reservation threshold. The solver–sampler mismatch study [5] distinguishes language models asked to find optimal moves (solvers) from models asked to emulate human-like negotiators (samplers); this matters directly for calibration, because an acceptance curve estimated from sampler behavior may differ substantially from the curve faced by a solver-style agent. The emotions benchmark [9] shows that prompt-conditioned affective states measurably shift LLM price-negotiation outcomes, implying that $a(p)$ is not a stable product-specific object but can drift with agent configuration — a caveat for any planner that estimates the curve once and reuses it.

Teams and institutions. Research on intra-team negotiation strategies [8] studies how a group negotiating jointly against a competitor, matcher, or conceiver should coordinate internally. Our parallel threads are a mechanical, fully-coordinated special case of a negotiation team — the principal chooses a common cap, so intra-team strategic conflict is assumed away — but [8] suggests that semi-autonomous LLM threads may incur coordination costs beyond the per-thread cost c_i . Classical e-procurement adoption research [7] models organizational acceptance of electronic procurement through behavioral control, subjective norms, and perceived benefits and risks; it reminds us that deployment of agentic sourcing faces institutional acceptance constraints that a deadline-and-cost model abstracts from. Network analyses of procurement crime [6] show that procurement ecosystems exhibit coordinated seller behavior, including shell-company structures; correlated sellers would violate the independence assumption behind $1 - (1 - a)^n$, and [6] supplies the empirical reason to take that violation seriously.

Methodological analogy. The QuWARP cost-model assessment of workload-level reuse planning for quantum circuit simulation [10] uses an analytical cost model to size a fan-out (repeated-run workload) against resource constraints before committing compute; its treatment of amortized per-task cost parallels our

per-thread cost c_t , and its threshold-rule style of analysis is the derivation style adopted here.

3 Methods

3.1 1 Setting and adopted conventions

A buyer must procure one unit by a hard deadline. There are $n \in \mathbb{N}_+$ parallel negotiation threads, each facing an independent seller. Each thread quotes against a common price cap $p \geq c$, where c is the sellers' common reserve (walk-away) price. Conditional on the cap, thread i accepts independently with probability $a(p)$, where $a : [c, \infty) \rightarrow [0, 1)$ is strictly increasing and smooth, with $a(p) < 1$ (a cap never guarantees acceptance).

Convention (adopted from draft C, documented in Appendix A). Costs and values are:

- R : gross value of procurement (realized iff at least one thread accepts).
- $L > R$: fulfillment loss if $K = 0$ (deadline missed; stock-out, SLA penalty, or downstream task failure).
- $c_t > 0$: per-thread resource cost (compute, human-in-the-loop review, API spend), paid $c_t n$.
- $e > 0$: excess-commitment cost per unwanted acceptance.

Let K be the number of accepting threads. Under seller independence, $K \sim \text{Binomial}(n, a)$ with

$$P(K = 0) = (1 - a(p))^n, \quad P(K \geq 1) = 1 - (1 - a(p))^n.$$

Unwanted acceptances number $K - 1$ when $K \geq 1$ and 0 when $K = 0$, so the expected number of unwanted acceptances is

$$\mathbb{E}[(K - 1)^+] = \mathbb{E}[K] - P(K \geq 1) = na - (1 - (1 - a)^n).$$

The buyer's expected total cost is

$$C(n, p) = R P(K \geq 1) + L P(K = 0) + c_t n + e(na - (1 - (1 - a)^n)),$$

which collects to the working form

$$C(n, p) = R + (L - R)(1 - a)^n + c_t n + e(na - (1 - a)^n). \quad (1)$$

Equivalence note (from draft B). Under the alternative convention of draft B — no gross value R , a fulfillment loss L_f on failure, conservative cap-priced procurement $p P(K \geq 1)$, and excess cost κ — the objective reduces to the canonical form $C_B(n, p) = L_f + nc + (p - L_f - \kappa) A_n(q) + \kappa nq$ with $A_n(q) = 1 - (1 - q)^n$. Both forms share the same binomial geometry; the structural results below hold under either convention. We use form (1) throughout.

3.2 2 Acceptance curve

We instantiate the acceptance curve as exponential with sensitivity $\lambda > 0$:

$$a(p) = 1 - e^{-\lambda(p-c)}, \quad p \geq c, \quad (2)$$

so that $1 - a(p) = e^{-\lambda(p-c)}$, $(1 - a)^n = e^{-n\lambda(p-c)}$, and $a'(p) = \lambda e^{-\lambda(p-c)}$. This curve is convex, matching the convex-quantile assumption under which [1],[2] prove the substitution result. Draft B proposed

a power quantile curve $q(p) = ((p - p_{\min}) / (p_{\max} - p_{\min}))^\gamma$ with $\gamma > 1$, also convex; the choice between families is documented in Appendix A. The planner solves

$$\min_{n \in \mathbb{N}_+, p \geq c} C(n, p).$$

4 Analysis

4.1 1 Marginal value of a thread: geometric decay

Fix the induced acceptance probability $a = a(p)$ and consider adding one thread. From (1),

$$\Delta C(n) = C(n+1, p) - C(n, p).$$

Computing each term:

- Fulfillment term: $(L - R) [(1 - a)^{n+1} - (1 - a)^n] = -(L - R) a (1 - a)^n$.
- Thread cost: c_t .
- Excess-commitment term: $e a + e a (1 - a)^n$.

Summing:

$$\Delta C(n) = c_t + e a - (L - R + e) a (1 - a)^n = c_t + e a - B a (1 - a)^n, \quad (3)$$

where $B \equiv L - R + e$ is the *commitment-adjusted stakes* constant. The marginal benefit $B a (1 - a)^n$ decays **geometrically** at ratio $(1 - a)$: each additional thread helps only through the event that all n existing threads fail, whose probability shrinks by the factor $(1 - a)$ per thread. Since $(1 - a)^n$ is strictly decreasing in n , $\Delta C(n)$ is strictly increasing in n toward its asymptote $c_t + e a > 0$; hence $C(n, \cdot)$ is convex in n and the optimum is a threshold.

4.2 2 Threshold rule for team size

The buyer hires threads while $\Delta C(n) < 0$, i.e., while $B a (1 - a)^n > c_t + e a$, i.e., while $(1 - a)^n > \frac{c_t + e a}{B a}$. Taking logs ($\ln(1 - a) < 0$ flips the inequality):

$$n^*(a) = \left\lceil \frac{\ln\left(\frac{c_t + e a}{B a}\right)}{\ln(1 - a)} \right\rceil, \quad (4)$$

when $\frac{c_t + e a}{B a} < 1$ (otherwise $n^* = 0$: do not fork at all). Because the numerator magnitude grows only logarithmically in $1/a$ while $-\ln(1 - a)$ grows linearly in a for small a , n^* grows roughly like $\ln(1/c_t)/a$ as $a \rightarrow 0$: cheap acceptance probabilities justify large teams, expensive ones do not.

4.3 3 Worked example at a fixed cap

Take $L = 100$, $R = 60$, $e = 5$, $c_t = 1$, $a = 0.3$. Then $B = 40 + 5 = 45$,

$$B a = 45 \times 0.3 = 13.5, \quad c_t + e a = 1 + 5 \times 0.3 = 2.5,$$

ratio = $2.5/13.5 = 0.18519$; $\ln(0.18519) = -1.68546$; $\ln(0.7) = -0.35668$; quotient = 4.72554 , so by (4), $n^* = \lceil 4.72554 \rceil = 5$.

Verification with (3) directly:

- $\Delta C(4) = 2.5 - 13.5 \times 0.7^4 = 2.5 - 3.24135 = -0.74135 < 0$: hire the 5th thread.
- $\Delta C(5) = 2.5 - 13.5 \times 0.7^5 = 2.5 - 2.268945 = +0.231055 > 0$: stop at $n = 5$.

Expected cost at $(n, a) = (5, 0.3)$ from (1): $(1 - a)^5 = 0.16807$, $na = 1.5$,

$$C = 60 + 40 \times 0.16807 + 5 + 5 \times (1.5 - 0.16807) = 60 + 6.7228 + 5 + 6.65965 = 78.38245.$$

4.4 4 Optimal price cap and complementarity

With the exponential curve (2), differentiate (1) in p at fixed n :

$$\frac{\partial C}{\partial p} = -(L - R) n \lambda e^{-n\lambda(p-c)} + e n \lambda e^{-\lambda(p-c)} + e n \lambda e^{-n\lambda(p-c)}.$$

Setting $\partial C / \partial p = 0$ and dividing by $n\lambda > 0$ (for $n \geq 1$):

$$e e^{-\lambda(p-c)} = (L - R - e) e^{-n\lambda(p-c)} \implies e^{(n-1)\lambda(p-c)} = \frac{L - R - e}{e}. \quad (5)$$

An interior cap requires $L - R > e$: the fulfillment stakes must exceed the per-excess-commitment liability, otherwise the buyer caps at the reserve $p = c$. For $n = 1$ the exponent vanishes and no interior solution exists; from (1) with the exponential curve, $C(1, p) = R + (L - R)e^{-\lambda(p-c)} + c_t + e(1 - e^{-\lambda(p-c)})$, which is strictly decreasing in p , so a single thread drives its cap to the boundary $p \rightarrow \infty$ with cost $\rightarrow R + c_t + e$. We therefore exclude $n = 1$ from the interior fixed-point search and treat it explicitly as a boundary case (Section 4.5). For $n \geq 2$:

$$p^*(n) = c + \frac{\ln\left(\frac{L - R - e}{e}\right)}{(n - 1)\lambda}. \quad (6)$$

$p^*(n)$ is **decreasing** in n : with more threads, fulfillment risk is diversified, so the buyer quotes a lower, more aggressive cap while keeping failure probability acceptable. This is the precise sense in which parallelism substitutes for concession, refining the complementarity reading of [1],[2]: the instruments move together toward aggressiveness, not toward safety. (Draft C's abstract described the cap as "rising in concurrency"; its own derivation, equation (6), shows it is decreasing. We adopt the derivation; see Appendix A.)

4.5 5 Joint fixed point

The optimal pair must satisfy $n^* = n^*(a(p^*(n^*)))$ with consistency between (4) and (6). Worked with $L = 100$, $R = 60$, $e = 5$, $c_t = 1$, $\lambda = 0.5$, $c = 10$. Note $L - R - e = 35$ and $\ln(35/5) = \ln 7 = 1.94591$.

- $n = 5$: $p^* = 10 + 1.94591/(4 \times 0.5) = 10.97296$; $a = 1 - e^{-0.48648} = 0.38523$; $Ba = 17.33535$; $c_t + ea = 2.92615$; ratio = 0.16879 ; $\ln = -1.77966$; $\ln(0.61477) = -0.48648$; quotient = 3.65756 ; $n^*(a) = 4 \neq 5$.
- $n = 4$: $p^* = 10 + 1.94591/1.5 = 11.29727$; $a = 1 - e^{-0.64864} = 0.47731$; ratio = $3.38655/21.47895 = 0.15764$; $\ln = -1.84751$; $\ln(0.52269) = -0.64864$; quotient = 2.84844 ; $n^*(a) = 3 \neq 4$.

- $n = 3$: $p^* = 10 + 1.94591 = 11.94591$; $a = 1 - e^{-0.97296} = 0.62204$; ratio = $4.11020/27.99180 = 0.14683$; $\ln = -1.91841$; $\ln(0.37796) = -0.97296$; quotient = 1.97202 ; $n^*(a) = 2 \neq 3$.
- $n = 2$: $p^* = 10 + 1.94591/0.5 = 13.89182$; $a = 1 - e^{-1.94591} = 1 - 1/7 = 6/7 = 0.85714$; ratio = $5.28570/38.57130 = 0.13700$; $\ln = -1.98793$; $\ln(1/7) = -1.94591$; quotient = 1.02157 ; $n^*(a) = \lceil 1.02157 \rceil = 2$. **Fixed point:** $(n^*, p^*) = (2, 13.89182)$.

Margin check at the fixed point with (3): $(1 - a)^2 = 1/49 = 0.020408$; $a(1 - a)^2 = 0.85714 \times 0.020408 = 0.017493$; $B a (1 - a)^2 = 45 \times 0.017493 = 0.78716$; $\Delta C(2) = 5.28570 - 0.78716 = +4.49854 > 0$ (stop at 2); $\Delta C(1) = 5.28570 - 45 \times 0.85714 \times 0.14286 = 5.28570 - 5.51016 = -0.22446 < 0$ (hire the 2nd thread). Consistent.

Expected cost at the fixed point, from (1) with $n = 2$, $a = 6/7$, $(1 - a)^2 = 1/49$, $na = 12/7 = 1.71429$:

$$C(2) = 60 + 40 \times \frac{1}{49} + 2 + 5 \times \left(\frac{12}{7} - \frac{1}{49} \right) = 60 + 0.81633 + 2 + 8.46939 = 71.28572.$$

Same-cap comparison. A single thread at the same aggressive cap ($a = 6/7$, $n = 1$, $na = 6/7$):

$$C(1) = 60 + 40 \times \frac{1}{7} + 1 + 5 \times \left(\frac{6}{7} - \frac{1}{7} \right) = 60 + 5.71429 + 1 + 3.57145 = 70.28574.$$

At the *same aggressive cap*, one thread is cheaper ($70.28574 < 71.28572$) because the cap already delivers high acceptance. More importantly, for $n = 1$ equation (1) with the exponential curve gives $C(1, p) = 60 + 40(1 - a) + 1 + 5(2a - 1) = 96 - 30a$, strictly decreasing in a on $[0, 1]$; since $a(p) < 1$ but $a \rightarrow 1$ as $p \rightarrow \infty$, the single-thread cost has infimum 66, not attained, and already at $a = 0.9$ the cost is $69.0 < 71.28572$. The fixed point $(2, 13.89182)$ is therefore the optimum only over $n \geq 2$: the exponential curve admits no finite cap bound (unlike the draft-B quantile curve, which is capped at $p_{\max}$), so the global optimum over $n \in \mathbb{N}_+$, $p \geq c$ is the boundary case $n = 1$ with the cap pushed arbitrarily high and cost approaching 66. Under a bounded cap $p \leq p_{\max}$ the search would be restricted accordingly; we report the $n \geq 2$ fixed point as the interior candidate and flag the $n = 1$ boundary explicitly.

4.6 6 Dispersed offers

Unproven conjecture (projection, not a result). When seller offers are dispersed rather than summarized by a single acceptance probability, the event "all n existing threads fail" need no longer shrink by a constant factor $(1 - a)$ per thread. If, for example, thread-level failure probabilities are drawn independently from a distribution with an atom at zero of mass a_0 and a continuous part on $(0, 1)$, then $P(\text{all fail}) = \mathbb{E} \left[\prod_{i=1}^n (1 - a_i) \right]$, and if the continuous part has density behaving like $f(a) \propto a^{-\beta}$ near $a = 0$ with $0 < \beta < 1$, this failure probability decays polynomially, as $n^{-(1-\beta)}$, rather than geometrically. We have not verified that any empirically relevant offer distribution satisfies this density condition, we write down no dispersed-offer objective $C(n, p)$, and we compute no exponent or scaling of the optimal team size. Accordingly, we make no claim here about the rate at which n^* grows in any dispersion parameter; the conjecture is only that the profitable concurrency region may widen relative to the geometric regime. A derivation would require specifying the offer distribution, the induced failure probability, and the resulting $\Delta C(n)$, which we leave to future work. This connects to the dispersion comparative static of [1],[2]: buyers benefit from searching harder for bargains when prices are dispersed, but suffer when they instead try to guarantee procurement by offering higher prices. The solver-sampler distinction of [5] makes this regime empirically important, since LLM negotiators often behave as samplers.

5 Results

1. **Geometric decay (convergent across all drafts).** The marginal value of an additional negotiator decays geometrically at rate $(1 - a)$; the marginal cost of concurrency rises toward the asymptote $c_t + e a$ (equation (3)).
2. **Closed-form team size.** The optimal concurrency follows the ceiling rule (4); in the fixed-cap example ($a = 0.3$, $B = 45$, $c_t = 1$, $e = 5$), $n^* = 5$ with expected cost 78.38245.
3. **Optimal cap and substitution.** Under the exponential acceptance curve, the interior cap $p^*(n) = c + \ln((L - R - e)/e)/((n - 1)\lambda)$ is decreasing in n : parallelism substitutes for concession. An interior cap requires $L - R > e$.
4. **Interior fixed point ($n \geq 2$).** For the worked configuration ($L = 100$, $R = 60$, $e = 5$, $c_t = 1$, $\lambda = 0.5$, $c = 10$), the fixed point over $n \geq 2$ is $(n^*, p^*) = (2, 13.89182)$ with expected net cost 71.28572; the second thread is hired with margin $\Delta C(1) = -0.22446$. This is not the global optimum: for $n = 1$ the cost $C(1, p) = 96 - 30a(p)$ decreases to the unattained infimum 66 as the cap grows without bound, so with the unbounded exponential curve the boundary case $n = 1$ dominates; the interior fixed point is optimal only within $n \geq 2$ or under a bounded cap.
5. **Cap-dependence of the concurrency premium.** At the same aggressive cap $p^*(2)$, one thread costs 70.28574 versus 71.28572 for two; indeed, with the unbounded exponential curve the single-thread cost falls to its infimum 66 as the cap grows, so concurrency is never globally optimal in this configuration. Concurrency is valuable only under a bounded cap or when failure risk is material at the prevailing cap, not when a single thread's cap already secures acceptance.
6. **Dispersion (conjecture).** We conjecture that with dispersed offers, marginal thread value may decay more slowly than geometrically, which would widen the profitable concurrency region relative to the geometric regime. No dispersed-offer objective is derived and no scaling rate is claimed (Section 4.6).

All numbers are computed by hand from the stated formulas; no simulation or external data are introduced.

6 Discussion

6.1 1 Limitations

- **Convention sensitivity.** The numerical optima depend on the cost accounting convention (Appendix A). Under draft A's convention — purchase cost $p_{\text{cap}} P(X \geq 1)$ with no failure loss and an excess penalty $C_e \max(0, \mathbb{E}[X] - 1)$ — a different calibration ($\alpha = 0.10$, $p_{\min} = 10$, $p_{\text{cap}} = 15$, $c = 0.50$, $L_f = 2.00$, $C_e = 1.00$) yields $n^* = 1$ with a low expected cost under that convention's objective functional; we do not reproduce the draft-A objective's arithmetic here and so refrain from quoting a specific figure. This is not a contradiction of the threshold rule: it reflects different parameters and a convention in which failure is costless, which removes the dominant term $B a (1 - a)^n$ driving multi-thread hiring.
- **Independence of sellers.** Correlated seller pricing (market shocks, collusion as documented in [6]) would increase the probability of simultaneous acceptances and amplify excess-commitment loss, invalidating the binomial geometry.
- **Static common cap.** Adaptive per-thread caps (e.g., decreasing caps for later-started threads) could improve efficiency but are outside the common-cap planner.

- **Curve drift.** As [9] and [5] indicate, the acceptance curve $a(p)$ can drift with agent configuration and orientation (solver vs. sampler); a planner that estimates the curve once and reuses it may mis-size fan-out.

6.2 2 Falsifiability

The central claims are falsifiable as follows. (i) If observed marginal cost reductions across n remain roughly constant or increase, contradicting the geometric decay of (3), the binomial acceptance model is refuted for that market. (ii) If higher price dispersion leads to a *lower* optimal n (rather than the conjectured widened concurrency region of Section 4.6), the dispersion comparative static of [1],[2] fails, and our conjecture along with it. (iii) If the optimal cap rises rather than falls with n in controlled experiments, the substitution result (6) is refuted. A field experiment varying n and p across markets while measuring realized procurement costs would test all three.

6.3 3 Open questions

Stochastic or soft deadlines (time-discounted costs); online learning of $a(p)$ with Bayesian updating coupled to the concurrency decision; competition among multiple agentic buyers for the same seller pool, linking to the network-optimization view of [3]; and non-convex acceptance curves induced by affective configuration [9], which could challenge the convexity-based substitution result.

7 Conclusion

We presented a reconciled, fully auditable analytical framework for concurrency-aware procurement negotiation. The marginal value of a negotiator decays geometrically at rate $(1 - a)$, yielding a ceiling rule for optimal team size; under an exponential acceptance curve the optimal cap is decreasing in concurrency, making precise the sense in which parallelism substitutes for price concession; and we conjecture that dispersed offers may flatten the decay, widening the profitable concurrency region — a projection left unproven (Section 4.6). A worked fixed point over $n \geq 2$ ($n^* = 2$, $p^* = 13.89182$, cost 71.28572) and a same-cap comparison (70.28574 for one thread) show that concurrency pays only when failure risk is material; with the unbounded exponential curve the single-thread boundary case ($n = 1$, cost $\rightarrow 66$) in fact dominates, so the interior fixed point is optimal only within $n \geq 2$ or under a bounded cap. The framework bridges multi-agent LLM negotiation benchmarks [4],[5],[9], institutional adoption and integrity contexts [6],[7], team-coordination results [8], and analytical cost-model methodology [3],[10], and offers a foundation for extensions to dynamic deadlines, learned acceptance curves, and multi-unit demand.

References

- [1] TITLE: arXiv Query: `search_query=&id_list=2610.06017&start=0&max_results=1` [2] Parallelism or Concession? Concurrency-Aware Procurement Negotiation for Agentic Commerce. arXiv:2610.06017v1. <https://arxiv.org/abs/2610.06017v1> [3] Nonlinear Negotiation Approaches for Complex-Network Optimization: A Study Inspired by Wi-Fi Channel Assignment. arXiv:1902.09457v1. <https://arxiv.org/abs/1902.09457v1> [4] AgenticPay: A Multi-Agent LLM Negotiation System for Buyer-Seller Transactions. arXiv:2602.06008v1. <https://arxiv.org/abs/2602.06008v1> [5] Diversity Without Fidelity: A Solver-Sampler Mismatch in Multi-Agent LLM Negotiation Simulation. arXiv:2604.11840v3. <https://arxiv.org/abs/2604.11840v3> [6] Organized crime behavior of shell-company networks in procurement: prevention insights for policy and reform.

arXiv:2307.10028v1. <https://arxiv.org/abs/2307.10028v1> [7] Acceptance of e-procurement in organisations. arXiv:2005.10094v1. <https://arxiv.org/abs/2005.10094v1> [8] Intra-Team Strategies for Teams Negotiating Against Competitor, Matchers, and Conceders. arXiv:1604.04736v1. <https://arxiv.org/abs/1604.04736v1> [9] Deal Me Maybe: The Role of Emotions in Multi-Agent Negotiation. arXiv:2608.06922v1. <https://arxiv.org/abs/2608.06922v1> [10] DOI 10.5281/zenodo.23133404. QNFO: QuWARP Reconciled: An Analytical Cost-Model Assessment of Workload-Level Reuse Planning for Quantum Circuit Simulation.

8 Appendix A. Divergence report

D1. Cost accounting convention (DIVERGENT). Draft A charges purchase cost $p_{\text{cap}} P(X \geq 1)$ (one unit), no penalty on total failure, and an excess-commitment penalty $C_e \max(0, \mathbb{E}[X] - 1)$. Drafts B and C charge a fulfillment loss on failure (L_f or $L > R$), and per-unit excess cost κ or e on $(K - 1)^+$ unwanted acceptances. **Convention adopted:** form (1) of Section 3.1 (drafts B and C), because it prices the deadline miss explicitly; draft A's convention is reported qualitatively in Section 6.1.

D2. Acceptance-curve family (DIVERGENT). Draft B proposed a power quantile curve $q(p) = ((p - p_{\min}) / (p_{\max} - p_{\min}))^\gamma$ with $\gamma > 1$, which is convex and bounded above by 1 at a finite $p_{\max}$; drafts A and C used exponential-form curves $a(p) = 1 - e^{-\lambda(p-c)}$, convex but reaching 1 only asymptotically, so no finite cap guarantees acceptance. **Convention adopted:** the exponential curve (2) for the main text, for closed-form tractability of the cap formula (6); the bounded-cap implication of the quantile family is noted in Section 4.5.

D3. Cap-concurrency direction (DIVERGENT). Draft C's abstract described the optimal cap as "rising in concurrency," but its own derivation — equation (6) here — shows $p^*(n)$ is decreasing in n . **Convention adopted:** the derivation; the cap is decreasing in n (parallelism substitutes for concession).