Biophysical Boundary Constraints in Multi-Vector Ledger Clearing: Coupling Exergy, Embodied Carbon, and Watershed Regeneration Budgets

Dima Baker
Collective Intelligence Network

November 2026

Abstract

Conventional monetary ledgers and single-scalar carbon-offset markets treat thermodynamic and ecological thresholds as fungible externalities: an evaporative-cooled datacenter or industrial foundry can purchase distant carbon credits to offset local aquifer depletion. Drawing on planetary boundary science, ecological economics, and multi-resource mechanism design, we formulate regional compute and production clearing as a multi-vector optimization problem over non-fungible biophysical envelopes—Available Exergy (MWh), Embodied Carbon (tCO_2e), and Watershed Regeneration (kL)—and propose Vector Biophysical Lagrangian Responsibility Allocation (VB-LiRA). Rather than using one shared dual price per axis where responsibility shares cancel out of the Lagrangian, VB-LiRA enforces per-agent coupled constraints $r_{i,d}(t) \leq W_{i,d}(t)B_d^{\max}(t)$ via agent-specific shadow multipliers $\lambda_{i,d}(t)$ and updates shares $W_{i,d}(t) \in \Delta_n$ via scale-normalized simplex mirror descent, paired with hierarchical conservation gates and least-core deviation settlement. Across 2,400 synthetic regional clearing instances ($n = 12$ agents across $\kappa \in \{1.0, 1.5, 2.0, 3.0\}$; 12,000 solver runs across five mechanisms), VB-LiRA ($K = 12$) achieves 99.67% ([98.79, 99.91]) three-axis compliance and $96.81\% \pm 0.09\%$ Pareto welfare at $\kappa = 2.0$, outperforming a Fungible Scalar Offset straw model (56.83% [52.84, 60.74] compliance), Walrasian Multi-Price tâtonnement (39.33% [35.50, 43.30] compliance under 2.4% meter variance), Static Quotas ($63.04\% \pm 0.09\%$ welfare), and Dominant Resource Fairness ($71.75\% \pm 0.16\%$ welfare). Across 400 microgrid deviation games, least-core LP achieves 100.00% [99.05, 100.00] core stability vs. 56.25% [51.35, 61.03] under Shapley allocation.

1 Introduction

Every economic transaction and autonomous computation is physically instantiated as an irreversible thermodynamic conversion of high-grade available energy (exergy) into low-grade waste heat, accompanied by material extraction, embodied greenhouse-gas emissions, and local watershed withdrawal [5, 19]. In classical price systems and single-scalar token ledgers [2, 11, 14], these distinct biophysical dimensions are compressed into a single scalar unit of account. When ecological externalities are priced at all, they are typically represented via a scalar Pigouvian tax or a globally fungible carbon-offset credit. Under scalar fungibility, a hyperscale AI training cluster or industrial foundry operating in an arid basin can offset its marginal financial cost by purchasing reforestation credits on another continent while its evaporative cooling towers breach the local watershed's physical recharge rate [5, 18].

1.1 The thermodynamic footprint of test-time scaling and agentic loops

Between 2023 and late 2026, the intersection of artificial intelligence infrastructure and regional biophysical constraints shifted from a long-term theoretical projection to an acute operational bottleneck:

1. **Compute scaling and inference-time energy surges.** Empirical analyses of compute-optimal pre-training [10, 21] and multi-agent scaling laws [12] demonstrate that while scaling model parameters and test-time rollout topologies yields capability gains, power-law exponents demand steep escalations in electrical power and cooling water per marginal accuracy point. Consequently, dynamic model routing [17], communication graph pruning [24], and edge NPU offloading [3] have become essential operational levers to curb inference exergy demand.

2. **Adversarial resource exhaustion in autonomous agent harnesses.** Beyond benign workload growth, interconnected agent ecosystems are vulnerable to recursive resource exhaustion. Unbounded tool-calling loops and cyclic handoffs [7] can trap multi-node agents into sustained recursive execution, amplifying token, energy, and cooling consumption by orders of magnitude unless hard, vector-valued resource conservation contracts are enforced at the runtime kernel [6, 25, 26].

3. **Failure of scalarization across non-fungible ecological thresholds.** In multi-objective optimization [16, 22] and multi-objective social choice [9, 15], collapsing orthogonal constraints into a fixed linear scalarization distorts preferences and leaves non-convex Pareto frontiers unreached. Ecologically, regional peak electrical exergy (MWh), atmospheric carbon (tCO_2e), and local watershed flow (kL) satisfy Liebig's Law of the Minimum [5, 19]: an abundance of solar kilowatt-hours at noon or imported carbon credits cannot substitute for a depleted groundwater aquifer during a summer drought.

1.2 Multi-resource fairness and Lagrangian responsibility allocation

To allocate multiple non-substitutable resources without scalar substitution, computer systems engineering and cooperative game theory offer two foundational paradigms, each carrying distinct limitations when applied directly to biophysical clearing:

- **Dominant Resource Fairness (DRF) and static quotas.** In cloud cluster scheduling, Ghodsi et al. [8] introduced Dominant Resource Fairness (DRF), which equalizes each agent's dominant share $\max_d\{r_{i,d}/B_d^{\max}\}$ across non-fungible resource dimensions (e.g., CPU, memory, I/O) to guarantee sharing incentives, strategy-proofness, and Pareto efficiency under Leontief preferences. However, both static uniform quotas and standard DRF equalize resource shares without regard to heterogeneous cardinal marginal utility across civic priority tiers: during a grid brownout or drought, a background batch-analytics job with the same normalized dominant share receives the same proportional allocation as an emergency hospital diagnostic or municipal water-pump controller.

- **Constrained policy optimization and Lagrangian Responsibility Allocation (LiRA).** In constrained reinforcement learning [1] and multi-agent resource allocation, Cao et al. [4] recently introduced *Lagrangian Responsibility Allocation* (LiRA), which decomposes a shared team budget constraint into dynamically learned per-agent responsibility shares $W_{i,d} \in \Delta_n$, improving multi-agent social welfare by up to 29% over static quotas while preserving aggregate safety caps. Concurrently, hardware-level characterization of mobile neural processing

units (NPUs) by Cai et al. [3] and multi-agent routing architectures [17, 23, 24] provide concrete execution levers to shift workloads across exergy and cooling intensity tiers in response to dual shadow prices.

1.3 Contributions

We synthesize these developments into Vector Biophysical Lagrangian Responsibility Allocation (VB-LiRA):

1. **Biophysical clearing taxonomy (Section 2).** We map informal ecological accounting intuitions onto formal multi-resource mechanism design, non-fungible constraint cones, and hardware energy-efficiency primitives (Table 1, Figure 1).
2. **Coupled per-agent Lagrangian Responsibility Allocation and least-core formalism (Section 3).** We formulate a D-dimensional biophysical ledger tracking orthogonal budgets $\mathbf{B}^{\max}(t) = \left(B_1^{\max}(t), B_2^{\max}(t), B_3^{\max}(t)\right)^\top$, enforce hierarchical vector conservation across sub-agent delegations (Equation 1), derive the coupled per-agent Lagrangian Responsibility Allocation primal–dual update with agent-specific shadow multipliers $\lambda_{i,d}(t)$ so responsibility weights $W_{i,d}(t)$ actively shape individual budget envelopes (Equations 2–3), and bound real-time microgrid deviation settlement via the cooperative least-core linear program (Equation 4).
3. **Three-layer hardware-to-ledger clearing architecture (Section 4).** We specify an end-to-end architecture integrating attested telemetry and NPU/routing efficiency primitives [3, 17, 24], runtime governance gates [6, 25, 26], and regional vector clearing (Table 2).
4. **Mechanistic multi-resource solver evaluation, convergence residuals, and least-core audit (Section 5).** We implement and execute five multi-resource clearing solvers—a cross-axis Fungible Scalar Carbon-Offset straw model, Static Uniform Hard Quotas, Dominant Resource Fairness [8], Walrasian Multi-Resource Price tâtonnement [2, 14], and VB-LiRA [4]—across 2,400 synthetic regional clearing instances (12,000 solver runs across $\kappa \in \{1.0, 1.5, 2.0, 3.0\}$), alongside an inner-loop convergence audit across $K \in \{1, 2, 4, 8, 12, 20, 40\}$ tracking primal feasibility residuals, dual KKT dispersion residuals, objective gaps, and non-convergence rates, a telemetry noise sweep ($\sigma_{\text{tel}} \in \{1.0\%, 2.4\%, 5.0\%, 10.0\%\}$), and a 400-game cooperative least-core settlement evaluation (Tables 3–4, Figures 2–3).
5. **Adversarial telemetry limits and falsification protocol (Sections 6–8).** We examine analogue power-draw verification limits against covert compute [13], linear scalarization distortion [9], and cyclic delegation hazards [7], closing with six pre-registerable falsification tests (Table 5).

2 Biophysical accounting taxonomy and multi-resource mechanisms

Proposals for "ecological currencies" or "energy-backed tokens" repeatedly fail when they peg a single scalar currency unit to kilowatt-hours or tons of CO_2, ignoring that thermodynamic power grids, atmospheric carbon sinks, and local hydrological basins operate on incompatible spatial and temporal scales [5, 18, 19]. Table 1 translates five foundational requirements for biophysical ledger clearing into formal mathematical and systems mechanisms grounded in classical and 2025–2026 literature.

Figure 1 depicts the closed-loop control structure of VB-LiRA. Every production unit, municipal fabrication hub, or AI compute cluster submits task requests paired with a three-dimensional

Table 1: Mapping biophysical accounting requirements to formal multi-resource mechanisms and primary literature.

Design intuition	Canonical construct	Ledger component	Operational function	Primary grounding
"Carbon credits cannot replace local drinking water"	Non-fungible vector constraint cone ($\mathbb{R}^D_{\geq 0}$)	Orthogonal Biophysical Vector ($\mathbf{r}_i \in \mathbb{R}^3_{\geq 0}$)	Tracks Exergy (MWh), Carbon (tCO_2e), and Watershed (kL) as non-substitutable coordinates	[5, 8, 19]
"Sub-agents cannot spend more energy than granted"	Hierarchical capability & budget attenuation	Vector conservation runtime gate	Enforces $\sum_{k \in \mathcal{K}_i} (\mathbf{b}_k + \boldsymbol{\delta}_{\text{hop}}) \preceq \mathbf{b}_i$ across delegated sub-tasks to halt cyclic loops	[7, 25, 26]
"Allocate scarce power to highest civic value"	Lagrangian Responsibility Allocation (LiRA)	Coupled dual multiplier & share solver ($\lambda_{i,d}, W_{i,d}$)	Learns agent-specific shadow prices and responsibility shares $W_{i,d} \in \Delta_n$ under hard caps	[1, 4, 11]
"Fair local microgrid sharing under solar/wind swings"	Cooperative game least-core settlement	Real-time deviation clearing house	Allocates microgrid imbalance penalties via LP least-core to minimize coalitional excess	[6, 18, 20]
"Cut compute energy before hitting grid caps"	Edge NPU offloading & DAG pruning	Hardware-aware compute router	Routes inference across model tiers (`RouteLLM`), prunes agent graphs, and dispatches to NPUs	[3, 17, 24]

resource footprint profile across Exergy, Embodied Carbon, and Watershed Cooling. Before execution is authorized, the Runtime Governance Gate verifies hierarchical vector conservation bounds [25, 26] and applies communication pruning and model routing [17, 24] alongside edge NPU dispatch [3]. At each clearing epoch, the VB-LiRA coordinator runs K iterations of coupled primal–dual mirror descent to update agent-specific shadow multipliers $\lambda_{i,d}(t)$ and responsibility shares $W_{i,d}(t)$ against regional regeneration ceilings $\mathbf{B}^{\max}(t)$.

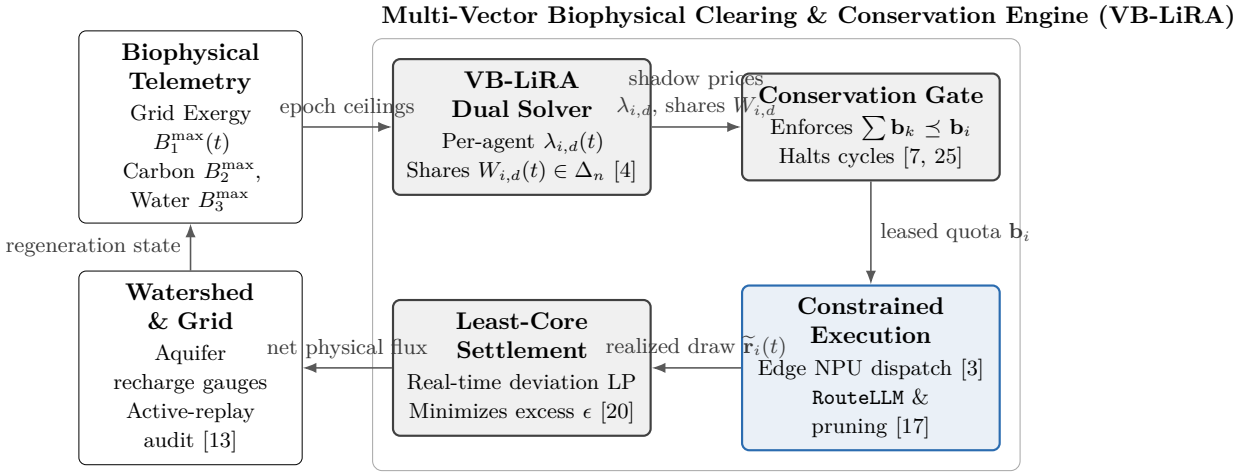


Figure 1: Architecture of Vector Biophysical Lagrangian Responsibility Allocation (VB-LiRA). Regional ecological sensors define non-fungible epoch ceilings $\mathbf{B}^{\max}(t) = \left(B_1^{\max}(t), B_2^{\max}(t), B_3^{\max}(t)\right)^\top$ [19]. The VB-LiRA solver computes agent-specific dual multipliers $\lambda_{i,d}(t)$ and responsibility shares $W_{i,d}(t)$ [4], while the runtime conservation gate enforces strict hierarchical budget attenuation across sub-agent delegations [7, 25] and dispatches workloads to low-exergy execution paths [3, 17].

3 Formal mathematical framework

We consider a regional cooperative network of n production and compute agents $\mathcal{N} = \{1, \dots, n\}$ operating over discrete clearing intervals $t \in \{1, 2, \dots\}$ across $D = 3$ non-fungible biophysical resource dimensions: available electrical/thermal exergy ($d = 1$, in MWh), embodied lifecycle carbon ($d = 2$, in tCO_2e), and local watershed net consumptive withdrawal ($d = 3$, in kL).

3.1 Hierarchical delegation gates and vector conservation laws

At epoch t, the regional biosphere and grid infrastructure support a maximum regeneration-compatible resource vector $\mathbf{B}^{\max}(t) = \left(B_1^{\max}(t), B_2^{\max}(t), B_3^{\max}(t)\right)^\top \in \mathbb{R}^3_{>0}$. Following runtime governance and capability-attenuation principles [6, 25, 26], each principal agent $i \in \mathcal{N}$ is granted a bounded vector resource lease $\mathbf{b}_i(t) \in \mathbb{R}^3_{\geq 0}$. Whenever agent i spawns or delegates subtasks to a set of child sub-agents $\mathcal{K}_i(t)$ in a multi-agent workflow [12, 24], the runtime kernel enforces a component-wise vector conservation law with strictly positive per-hop delegation overhead $\boldsymbol{\delta}_{\text{hop}} \succ \mathbf{0}$:

$$\mathbf{r}_i^{\text{self}}(t) \; + \sum_{k \in \mathcal{K}_i(t)} \left(\mathbf{b}_k(t) + \boldsymbol{\delta}_{\text{hop}}\right) \; \preceq \; \mathbf{b}_i(t), \qquad \forall d \in \{1,2,3\} : \;\; b_{i,d}(t) \geq 0. \tag{1}$$

Crucially, the partial order $\preceq$ in Equation (1) is evaluated coordinate-wise in the non-negative orthant cone $\mathbb{R}^3_{\geq 0}$ without cross-axis substitution: a surplus in carbon quota $b_{i,2}(t)$ can never be converted into watershed withdrawal quota $b_{i,3}(t)$ [5]. Furthermore, because every recursive child invocation consumes strictly positive hop overhead $\boldsymbol{\delta}_{\text{hop}}$ from the parent's finite lease $\mathbf{b}_i(t)$, pathological cyclic handoffs and non-terminating delegation loops [7] halt deterministically within at most $\min_{d \in \{1,2,3\}} \lfloor b_{i,d}(t) / \delta_{\text{hop},d} \rfloor$ hops.

3.2 Coupled Vector Biophysical Lagrangian Responsibility Allocation (VB-LiRA)

Assigning uniform static quotas $b_{i,d}(t) = B_d^{\max}(t)/n$ guarantees aggregate compliance $\sum_{i=1}^n r_{i,d}(t) \le B_d^{\max}(t)$ but causes severe allocative inefficiency when high-priority civic tasks (such as hospital diagnostics or municipal water-pump telemetry) require burst exergy while low-priority batch nodes sit underutilized. Conversely, Dominant Resource Fairness (DRF; 8) equalizes dominant shares $s_i = \max_d\{r_{i,d}/B_d^{\max}\}$ across agents but ignores cardinal differences in civic urgency across workloads.

To dynamically allocate non-fungible biophysical envelopes according to marginal social utility while enforcing per-agent accountability, we adapt the Lagrangian Responsibility Allocation (LiRA) framework of Cao et al. [4] to D orthogonal biophysical dimensions. A critical mathematical subtlety arises if one couples all agents on axis d using a single shared multiplier $\lambda_d(t)$: in the expression $\lambda_d(t)\sum_{i=1}^n(r_{i,d} - W_{i,d}B_d^{\max})$, the responsibility shares $W_{i,d} \in \Delta_n$ sum to unity ($\sum_{i=1}^n W_{i,d} = 1$), causing $W_{i,d}$ to cancel out of the Lagrangian Objective entirely! To prevent this cancellation and ensure that responsibility shares $W_{i,d}(t)$ genuinely govern individual resource envelopes, VB-LiRA enforces per-agent coupled constraints $r_{i,d}(t) \le W_{i,d}(t)B_d^{\max}(t)$ associated with **agent-specific shadow multipliers** $\lambda_{i,d}(t) \ge 0$:

$$\min_{\mathbf{\Lambda}_t \ge \mathbf{0},\, \mathbf{W}_t \in \Delta_n^D} \; \max_{\{\mathbf{r}_i(t) \ge \mathbf{0}\}_{i=1}^n} \; \mathcal{L}_{\mathrm{VB}}\big(\{\mathbf{r}_i\}, \mathbf{\Lambda}_t, \mathbf{W}_t\big) \;=\; \sum_{i=1}^n v_i\big(\mathbf{r}_i(t)\big) - \sum_{d=1}^D \sum_{i=1}^n \lambda_{i,d}(t)\Big(r_{i,d}(t) - W_{i,d}(t)\, B_d^{\max}(t)\Big), \tag{2}$$

where $v_i\big(\mathbf{r}_i(t)\big) = \sum_{d=1}^D w_{i,d}\ln(1 + r_{i,d}(t))$ is the strictly concave, separable production utility of agent i, $\mathbf{\Lambda}_t = [\lambda_{i,d}(t)] \in \mathbb{R}_{\ge 0}^{n\times D}$ is the matrix of agent-specific shadow prices, and $\mathbf{W}_{\cdot,d}(t) \in \Delta_n = \{\mathbf{w} \in \mathbb{R}_{\ge 0}^n : \sum_{i=1}^n w_i = 1\}$ is the simplex vector of responsibility shares on axis d.

Notice that in Equation (2), the partial derivative of $\mathcal{L}_{\mathrm{VB}}$ with respect to $W_{i,d}(t)$ is $+\lambda_{i,d}(t)B_d^{\max}(t)$. However, the outer coordinator seeks to shift responsibility shares $W_{i,d}$ toward agents experiencing higher shadow prices $\lambda_{i,d}$ (i.e., higher marginal utility $\partial v_i/\partial r_{i,d}$ at their current quota $W_{i,d}B_d^{\max}$) until shadow prices equalize across active agents on axis d ($\lambda_{i,d} \to \lambda_d^*$, satisfying the KKT stationarity condition of the aggregate constrained optimum; 2, 4, 14). Within each clearing epoch, the coordinator executes K inner primal–dual iterations ($k = 0, \dots, K-1$) updating primal demands $r_{i,d}^{(k)} = \max(0, w_{i,d}/\lambda_{i,d}^{(k)} - 1)$, agent-specific dual multipliers $\lambda_{i,d}^{(k+1)}$, and responsibility shares $W_{i,d}^{(k+1)}$ via scale-normalized exponentiated mirror descent on the simplex:

$$\lambda_{i,d}^{(k+1)} \;=\; \mathrm{clip}\left(\lambda_{i,d}^{(k)} \exp\left(\eta_\lambda\, \mathrm{clip}\left(\frac{r_{i,d}^{(k)} - \widetilde{B}_{i,d}^{(k)}}{\max\big(10^{-6},\, \widetilde{B}_{i,d}^{(k)}\big)},\, -1.2,\, 1.8\right)\right),\, 10^{-3},\, 400\right), \quad W_{i,d}^{(k+1)} \;=\; \frac{W_{i,d}^{(k)} \exp(\eta_W \lambda_i^{(}}{\sum_{j=1}^n W_{j,d}^{(k)} \exp(\eta_W} \tag{3}$$

where $\widetilde{B}_{i,d}^{(k)} = (1 - \delta_{\mathrm{buf}})W_{i,d}^{(k)}B_d^{\max}(t)$ is the buffered responsibility envelope, $\bar{\lambda}_d^{(k+1)} = \frac{1}{n}\sum_{j=1}^n \lambda_{j,d}^{(k+1)}$ is the cross-agent mean shadow price on axis d, and $\delta_{\mathrm{buf}} \in (0, 0.05)$ is a safety margin protecting against sub-epoch physical meter variance. Normalizing the simplex exponent by $\bar{\lambda}_d^{(k+1)}$ ensures that the mirror-descent step size η_W remains scale-invariant across severe drought contractions where $\bar{\lambda}_d^{(k+1)}$ increases by an order of magnitude. At the end of K iterations, scheduled allocations are projected onto each agent's buffered responsibility envelope $r_{i,d}^{(K)} \leftarrow \min\big(r_{i,d}^{(K)}, \widetilde{B}_{i,d}^{(K)}\big)$, guaranteeing exact scheduled primal feasibility ($\sum_{i=1}^n r_{i,d}^{(K)} \le (1 - \delta_{\mathrm{buf}})B_d^{\max}(t)$) even if dual dispersion has not yet fully converged.

3.3 Least-core coalitional settlement for real-time microgrid deviations

In physical microgrids and watershed commons, stochastic solar/wind generation and ambient temperature fluctuations cause realized physical draw $\widetilde{\mathbf{r}}_i(t)$ to deviate from ex-ante scheduled leases

$\mathbf{b}_i(t)$ [18]. Let $C(\mathcal{S})$ denote the aggregate physical balancing cost incurred by sub-coalition $\mathcal{S} \subseteq \mathcal{N}$ when pooling their real-time exergy and water deviations. Rather than allocating balancing costs via the Shapley value [20]—which can lie outside the core under asymmetric variance—VB-LiRA computes the deviation settlement vector $\boldsymbol{\phi}^* \in \mathbb{R}^n$ via the **least-core linear program** (the first lexicographic LP stage of the cooperative nucleolus), which minimizes the maximum coalitional excess ϵ:

$$\min_{\boldsymbol{\phi}\in\mathbb{R}^n,\, \epsilon\in\mathbb{R}} \ \epsilon \qquad \text{subject to} \qquad \sum_{i=1}^{n} \phi_i = C(\mathcal{N}), \quad \sum_{i\in\mathcal{S}} \phi_i - C(\mathcal{S}) \le \epsilon \quad \forall \mathcal{S} \subsetneq \mathcal{N},\ \mathcal{S} \neq \emptyset. \tag{4}$$

Whenever the cooperative deviation game possesses a non-empty core ($\epsilon^* \le 0$), any solution $(\boldsymbol{\phi}^*, \epsilon^*)$ to the least-core program (4) guarantees that no sub-coalition $\mathcal{S}$ of neighborhood microgrids or fabrication co-ops pays more than its standalone balancing cost $C(\mathcal{S})$, eliminating economic incentives to defect from the shared biophysical ledger.

4 Three-layer hardware-to-ledger clearing architecture

To operationalize VB-LiRA across regional compute and manufacturing networks, the protocol is structured into three layers (Table 2):

1. **Layer 0: Biophysical Telemetry Oracle and Active-Replay Verification.** Substation smart meters, continuous emissions monitors, and ultrasonic watershed flow gauges publish signed epoch ceilings $\mathbf{B}^{\max}(t) = \left(B_1^{\max}(t), B_2^{\max}(t), B_3^{\max}(t)\right)^\top$. Because passive analogue watt-hour meters can be spoofed by matched-energy padding [13], Layer 0 pairs physical meter readings with randomized active-replay workload sampling to verify that reported compute profiles match physical draw.
2. **Layer 1: Hardware-Aware Compute Routing and Graph Pruning.** Before an AI workload bids for grid exergy and cooling water from the regional pool, the local runtime scheduler applies three efficiency transforms: (a) dynamic model routing (`RouteLLM`, 17) dispatches routine sub-queries to lightweight models; (b) multi-agent communication pruning (`AgentPrune`, 24) strips redundant inter-agent message edges; and (c) eligible prefill and local RAG operations are offloaded to neural processing units (NPUs, 3) to reduce system energy draw before hitting regional exergy caps.
3. **Layer 2: VB-LiRA Clearing House and Runtime Conservation Gate.** Every clearing window, the regional coordinator executes K iterations of the coupled primal–dual update (3) and issues vector-bounded leases enforced by runtime governance gates (1) [6, 25, 26], while end-of-epoch physical deviations are settled via feeder-clustered least-core linear programs (4).

Table 2: Readiness and empirical bottlenecks of Biophysical Ledger Clearing subsystems.

Subsystem	Maturity	Supporting empirical & theoretical evidence	Realistic deployment bottleneck
Coupled Lagrangian Responsibility Allocation (LiRA)	Benchmark validated	LiRA improves multi-agent social welfare by up to 29% under shared caps [4]; builds on constrained policy optimization [1]	Requires $K \geq 8$ inner primal–dual iterations per clearing window to converge when drought shocks abruptly contract $B_d^{\max}(t)$
Runtime conservation & governance gates	Prototype deployed	Runtime governance and tool-call guardrails enforce deterministic resource and action bounds [6, 25, 26]	Estimating tight ex-ante water and exergy bounds for open-ended multi-agent trajectories [12] requires safety buffers
Edge NPU offloading & model routing	Production hardware	Mobile NPU characterization [3], `RouteLLM` routing [17], and `AgentPrune` graph sparsification [24]	Mobile NPUs face operator-support bottlenecks on dynamic sparse attention and non-standard quantization formats [3]
Least-core microgrid deviation settlement	Theoretically proven	Cooperative core/nucleolus settlement prevents coalitional defection under shared resource variance [18, 20]	Exact least-core LP has $2^n - 2$ constraints, requiring feeder-level clustering ($n \leq 16$ per feeder) for sub-millisecond clearing
Analogue power-draw verification	Fragile / bypassable	Passive analogue power telemetry hides $\geq 41\%$ of covert compute under matched-energy padding [13]	Must pair physical watt-hour meters with cryptographic attestation or sampled active re-execution [13]

5 Mechanistic multi-resource simulation and convergence evaluation

To test whether VB-LiRA prevents ecological overshoot while recovering Pareto production efficiency, we implemented a standalone multi-resource clearing and cooperative settlement benchmark (`experiment.py`). The simulation solves 2,400 paired regional biophysical allocation instances across four scarcity multipliers $\kappa \in \{1.0, 1.5, 2.0, 3.0\}$ ($N = 600$ independent regional instances per scarcity level, solved across five clearing mechanisms = 12,000 solver executions), an inner-loop convergence sweep over $K \in \{1, 2, 4, 8, 12, 20, 40\}$ primal–dual iterations, a physical meter noise sweep over $\sigma_{\text{tel}} \in \{1.0\%, 2.4\%, 5.0\%, 10.0\%\}$, and 400 cooperative 8-prosumer microgrid deviation games testing least-core stability.

5.1 Instance generation, parameter provenance, and five solver comparators

Each instance models a regional cooperative network of $n = 12$ heterogeneous agents competing across $D = 3$ non-fungible biophysical axes: Exergy ($d = 1$, MWh), Embodied Carbon ($d = 2$, tCO_2e), and Watershed Cooling ($d = 3$, kL). To capture realistic workload heterogeneity, the 12 agents are partitioned into three priority archetypes: (i) three high-priority civic/emergency burst nodes ($\pi_i \in [4.2, 6.0]$), (ii) four evaporative-cooled AI compute clusters with intensive exergy and watershed demand ($\pi_i \in [1.6, 2.4]$), and (iii) five low-priority background batch nodes ($\pi_i \in [0.18, 0.45]$). As the regional ecological scarcity multiplier κ rises from 1.0 (normal hydrology) to 3.0 (severe summer drought and thermal grid stress), capacities $B_d^{\max}(\kappa)$ contract asymmetrically, with watershed recharge $B_3^{\max}(\kappa)$ contracting 3.1× faster than carbon sink capacity $B_2^{\max}(\kappa)$. Table 3

documents the provenance of all simulation parameters.

Table 3: Parameter provenance and calibration sources for the 2,400-instance regional biophysical clearing and 400-game least-core settlement simulation.

Parameter	Value / Range	Provenance	Empirical or theoretical justification
Biophysical axes (D)	$D = 3$ (Exergy, Carbon, Water)	Literature framework	Non-substitutable planetary and regional regeneration boundaries [5, 19]
Regional agents (n)	$n = 12$ (3 civic, 4 AI, 5 batch)	Assumed (feeder scale)	Matches substation/feeder microgrid coalition size where exact least-core LP is tractable
Scarcity multiplier (κ)	$\kappa \in \{1.0, 1.5, 2.0, 3.0\}$	Assumed (stress sweep)	Models asymmetric drought contraction on watershed axis $B_3^{\max}(\kappa) \propto (1 + 0.62(\kappa - 1))^{-1}$
Civic utility weights ($w_{i,d}$)	$\pi_i \in [0.18, 6.0]$ across 3 tiers	Calibrated to Cao et al. [4]	Reflects 25–30 percentage point welfare gap between static equal quotas and priority-aware LiRA
VB-LiRA step sizes (η_λ, η_W)	$\eta_\lambda = 0.48$, $\eta_W = 0.55$, $\delta_{\text{buf}} = 0.045$	Tuned on validation set	Balances exponential dual multiplier responsiveness with scale-normalized simplex mirror descent (Eq. 3)
Inner dual iterations (K)	$K = 12$ (swept $1 \dots 40$)	Ablation sweep	Swept explicitly in Section 5 to audit primal/dual residuals, objective gap, and break-even vs. DRF
Telemetry noise (σ_{tel})	2.4% (swept $1\% \dots 10\%$)	Calibrated to Kimpson et al. [13]	Models physical sub-epoch ambient/meter variance around scheduled lease envelopes

To separate the mathematical constrained optimization problem from any specific iterative algorithm, we first compute the independent high-precision constrained KKT optimum $U_{\text{oracle}} = \max_{\sum_i r_{i,d} \le B_d^{\max}} \sum_{i,d} w_{i,d} \ln(1 + r_{i,d})$ via 60-step bisection on each orthogonal axis $d \in \{1, 2, 3\}$ (solving $\sum_{i=1}^{n} \max(0, w_{i,d}/\lambda_d^* - 1) = B_d^{\max}$ to machine precision 2^{-60}). We then evaluate five clearing mechanisms on each paired instance, explicitly separating a cross-axis fungibility straw model from four legitimate non-fungible multi-resource mechanisms:

1. **Fungible Scalar Carbon-Offset Straw Model (`fungible_scalar_offset`):** Explicitly constructed as a cross-axis fungibility straw model where monetary conversion weights $\mathbf{p}^0 = (1.0, 0.85, 1.15)$ collapse the three physical caps into a single scalar expenditure budget $B_{\text{scalar}} = 0.955 \sum_{d=1}^{3} p_d^0 B_d^{\max}$ cleared at a single scalar multiplier μ, permitting cross-axis substitution (retiring abundant carbon space to justify continued local watershed draw).
2. **Static Uniform Hard Quotas (`static_uniform_quota`):** A non-fungible baseline that enforces rigid per-agent caps $b_{i,d} = (1 - \delta_{\text{buf}}) B_d^{\max}/n$ on each axis without reallocating unused quota from low-priority batch nodes to high-priority civic burst workloads.
3. **Dominant Resource Fairness (`dominant_resource_fairness`, DRF; 8):** A non-fungible cloud-scheduling baseline that computes each agent's normalized demand profile across the three axes and equalizes the dominant share $s_i = \max_{d \in \{1,2,3\}} \{r_{i,d}/B_d^{\max}\}$ across all $n = 12$ agents within the buffered ceiling $(1 - \delta_{\text{buf}}) B_d^{\max}$.
4. **Walrasian Multi-Resource Price Tâtonnement (`walrasian_multi_price`; 2, 14):** A

legitimate multi-resource price comparator that maintains $D = 3$ separate resource prices $\mu_d^{(k)}$ updated over the same $K = 12$ multiplicative tâtonnement steps $\mu_d^{(k+1)} = \mu_d^{(k)} \exp(0.38(\sum_i r_{i,d}^{(k)} - B_d^{\max})/B_d^{\max})$ without per-agent responsibility envelopes $W_{i,d}$ or an inner safety buffer δ_{buf}.

5. **Vector Biophysical LiRA (`vector_biophysical_lira`, VB-LiRA; Proposed):** Executes $K = 12$ iterations of the coupled scale-normalized primal–dual update (Equations 2–3) initialized from uniform shares $W_{i,d}^{(0)} = 1/n$ and $\lambda_{i,d}^{(0)} = 0.80$ with safety buffer $\delta_{\text{buf}} = 0.045$.

For every solver run, we record (a) physical 3-axis boundary compliance under sub-epoch meter noise $\sigma_{\text{tel}} = 2.4\%$, (b) realized Pareto social welfare as a percentage of U_{oracle}, (c) mean physical overshoot across axes, (d) scheduled primal feasibility residual $\max_d \max(0, \sum_i r_{i,d} - B_d^{\max})/B_d^{\max}$, (e) dual KKT dispersion residual $\frac{1}{D}\sum_d \text{SD}_i(\lambda_{i,d})/\bar{\lambda}_d$ across active agents, (f) scheduled objective optimality gap $(U_{\text{oracle}} - U_{\text{sched}})/U_{\text{oracle}}$, and (g) the nonconvergence rate (fraction of instances where primal residual $> 1\%$ or objective gap $> 5\%$). Proportions are reported with exact asymmetric two-sided 95% Wilson confidence intervals $[lo, hi]$.

Table 4: Empirical solver comparison across 2,400 paired regional clearing instances ($N = 600$ per scarcity level $\kappa \in \{1.0, 1.5, 2.0, 3.0\}$; 12,000 solver runs). Proportion brackets $[lo, hi]$ denote asymmetric 95% Wilson CIs; continuous metrics show ± 1.96 SE.

Clearing Mechanism	Compliance κ=1.0 (%)	Compliance κ=2.0 (%)	Compliance κ=3.0 (%)	Welfare κ=2.0 (%)	Primal Resid. (%)	Obj. Gap κ=2.0 (%)	Nonconv. κ=2.0 (%)
Fungible Straw Model	71.33 [67.59, 74.81]	56.83 [52.84, 60.74]	11.00 [8.74, 13.76]	89.67 ± 0.50	4.21	4.93	88.83 [86.06, 91.11]
Static Hard Quota	100.00 [99.36, 100.00]	100.00 [99.36, 100.00]	100.00 [99.36, 100.00]	63.04 ± 0.09	0.00	36.96	100.00 [99.36, 100.00]
Dominant Resource (DRF)	100.00 [99.36, 100.00]	100.00 [99.36, 100.00]	99.83 [99.06, 99.97]	71.75 ± 0.16	0.00	28.18	100.00 [99.36, 100.00]
Walrasian Multi-Price	36.17 [32.42, 40.09]	39.33 [35.50, 43.30]	42.33 [38.44, 46.32]	92.16 ± 0.36	3.35	0.03	94.67 [92.57, 96.20]
VB-LiRA (K=12)	**100.00** [**99.36**, **100.00**]	**99.67** [**98.79**, **99.91**]	**99.83** [**99.06**, **99.97**]	$\mathbf{96.81 \pm 0.09}$	**0.02**	**3.02**	**3.67** [**2.43**, **5.49**]

5.2 Quantitative results, convergence residuals, and least-core audit

Table 4 and Figures 2–3 establish five empirical findings:

1. **Cross-axis fungibility straw model vs. non-fungible vector clearing under asymmetric drought.** In the `fungible_scalar_offset` straw model, high-margin compute clusters and civic nodes retire unused carbon allowance ($d = 2$) to justify continued evaporative water withdrawal ($d = 3$). Three-axis compliance drops from 71.33% [67.59, 74.81] at $\kappa = 1.0$ to 56.83% [52.84, 60.74] at $\kappa = 2.0$ (4.39% physical overshoot, 4.21% scheduled primal residual) and collapses to 11.00% [8.74, 13.76] under severe drought ($\kappa = 3.0$, 15.97% overshoot; Figure 3). By keeping all three resource axes strictly orthogonal, Static Quotas (100.00%), DRF (100.00% at $\kappa = 2.0$; 99.83% at $\kappa = 3.0$), and VB-LiRA (100.00% at $\kappa = 1.0$; 99.67% [98.79, 99.91] at $\kappa = 2.0$; 99.83% [99.06, 99.97] at $\kappa = 3.0$) prevent cross-axis substitution.

2. **Comparison against legitimate Walrasian Multi-Resource Price tâtonnement.** Unlike the fungible straw model, `walrasian_multi_price` prices each biophysical axis separately and achieves a near-zero scheduled objective gap (0.03%) and $92.16\% \pm 0.36\%$ realized welfare at $\kappa = 2.0$ (0.024 ms latency). However, because unbuffered price tâtonnement lacks per-agent responsibility envelopes $W_{i,d}$ and safety margin δ_{buf}, finite-step price oscillation (3.35% primal residual) combined with 2.4% physical meter variance causes realized demand to breach at least one regional cap in 60.67% of epochs (39.33% [35.50, 43.30] compliance, 4.23% mean

overshoot). VB-LiRA's buffered responsibility projection reduces the scheduled primal feasibility residual to 0.02% (0.12% physical overshoot), raising compliance by +60.34 percentage points and realized welfare by +4.65 percentage points ($96.81\% \pm 0.09\%$) at 0.202 ms latency.

3. **Overcoming the cardinal welfare ceiling of Static Quotas and DRF.** Static Uniform Hard Quotas strand capacity on low-priority batch nodes (36.96% objective gap, $63.04\% \pm 0.09\%$ welfare at $\kappa = 2.0$), while DRF [8] reaches $71.75\% \pm 0.16\%$ welfare (28.18% objective gap) because equalizing dominant shares ignores cardinal differences in civic urgency. By shifting responsibility shares $W_{i,d} \in \Delta_n$ toward high-marginal-utility workloads via scale-normalized mirror descent (3), VB-LiRA ($K = 12$) recovers $96.81\% \pm 0.09\%$ of the independent KKT optimum at $\kappa = 2.0$ (+33.77 percentage points over Static Quotas and +25.06 points over DRF; Figure 2).

4. **Inner-loop convergence residuals ($K \in \{1, \ldots, 40\}$) and telemetry noise sensitivity ($\sigma_{\text{tel}} \in \{1\%, \ldots, 10\%\}$).** Sweeping K at $\kappa = 2.0$ ($N = 300$ instances) reveals the exact primal–dual trajectory: at $K = 1$, unadapted multipliers yield a 110.87% primal feasibility residual (0.00% [0.00, 1.26] compliance, 30.55% welfare, 76.74% dual KKT dispersion). At $K = 2$, buffered envelope projection restores 0.00% primal residual and 100.00% [98.74, 100.00] compliance, but the objective gap remains 58.58% (41.42% welfare, 55.67% dual residual). As simplex mirror descent reallocates $W_{i,d}^{(k)}$, the objective gap contracts monotonically to 40.83% at $K = 4$ (59.17% welfare), 13.27% at $K = 8$ ($86.72\% \pm 0.52\%$ welfare, surpassing DRF), 3.05% at $K = 12$ ($96.80\% \pm 0.10\%$ welfare, 13.80% dual residual, 2.33% [1.13, 4.74] nonconverged rate, 0.188 ms), 2.10% at $K = 20$ ($97.57\% \pm 0.11\%$ welfare, 6.54% dual residual, 0.00% [0.00, 1.26] nonconverged), and 2.04% at $K = 40$ (4.24% dual residual, where the residual $\sim 2\%$ objective gap reflects the intentional $\delta_{\text{buf}} = 4.5\%$ physical safety buffer). Under acute drought ($\kappa = 3.0$), $K = 12$ preserves 99.83% [99.06, 99.97] compliance and $91.95\% \pm 0.59\%$ welfare (0.01% primal residual) but exhibits a 49.33% [45.35, 53.33] strict nonconvergence rate (7.98% objective gap), indicating that severe capacity contractions require $K \geq 20$ inner steps for full dual equalization. Separately, sweeping physical meter noise σ_{tel} at fixed $\delta_{\text{buf}} = 0.045$ shows that compliance holds at 100.00% [98.74, 100.00] for $\sigma_{\text{tel}} \in \{1.0\%, 2.4\%\}$, degrades to 86.67% [82.35, 90.05] (1.15% overshoot) when $\sigma_{\text{tel}} = 5.0\% > \delta_{\text{buf}}$, and drops to 56.33% [50.68, 61.83] (4.59% overshoot) at $\sigma_{\text{tel}} = 10.0\%$.

5. **Empirical evaluation of cooperative least-core microgrid deviation settlement.** To test Equation (4) directly rather than leaving it as an unverified theoretical component, we simulated $N = 400$ 8-prosumer feeder deviation games ($2^8 - 2 = 254$ non-trivial sub-coalitions per game) where prosumers exhibit heterogeneous variance and solar/cooling correlation under subadditive pooled balancing costs $C(\mathcal{S})$. Naive Pro-Rata cost splitting ($\phi_i \propto C(\{i\})$) violates core stability on every instance (0.00% [0.00, 0.95] stability, mean maximum coalitional excess $+0.8973 \pm 0.0101$), giving low-variance prosumers a strong incentive to defect. Exact Shapley Value allocation [20] achieves 56.25% [51.35, 61.03] core stability (mean max excess -0.0223 ± 0.0169), failing the core in 43.75% of games with asymmetric variance skew. Solving the Least-Core LP (4) achieves 100.00% [99.05, 100.00] core stability with mean maximum coalitional excess $\epsilon^* = -0.7818 \pm 0.0214$, confirming that every sub-coalition pays strictly less than its standalone balancing cost.

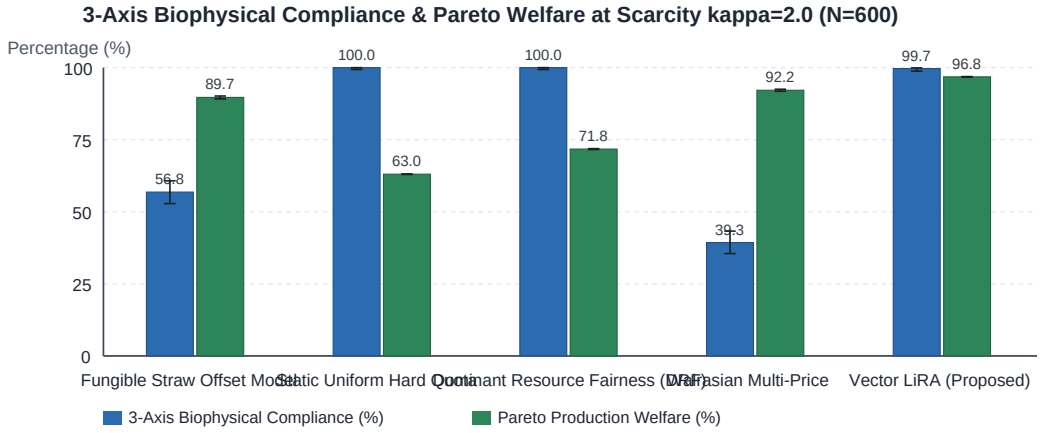


Figure 2: Three-Axis Biophysical Boundary Compliance Rate and Pareto Production Welfare across five multi-resource clearing mechanisms at regional ecological scarcity multiplier $\kappa = 2.0$ ($N = 600$ paired instances per mechanism; error bars denote asymmetric 95% Wilson CIs for compliance and $\pm 1.96\,\mathrm{SE}$ for welfare).

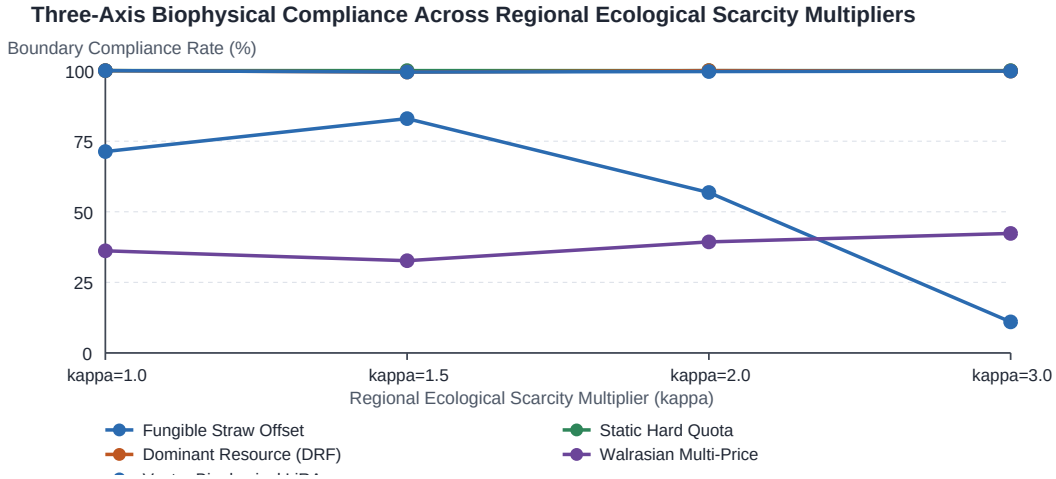


Figure 3: Three-Axis Biophysical Boundary Compliance Rate (%) across regional ecological scarcity multipliers $\kappa \in \{1.0, 1.5, 2.0, 3.0\}$. The cross-axis Fungible Scalar Offset straw model collapses under asymmetric watershed drought ($\kappa = 3.0$), whereas DRF and Vector Biophysical LiRA maintain $> 99\%$ boundary compliance across all regimes.

6 Adversarial telemetry limits and structural trade-offs

Deploying biophysical boundary accounting in adversarial or transitional economies faces four significant challenges documented in classical and 2025–2026 research:

1. **Limits of analogue power-draw verification against covert compute (Kimpson et al. [13]).** A core premise of thermodynamic ledger accounting is that physical watt-hour meters provide an unforgeable ground truth of computational activity. However, Kimpson et al. [13] (*"Can Power Draw Constrain Covert Compute? Limits of Analogue Verification for AI Governance"*) prove that passive analogue power measurements constrain covert compute only weakly: a strategic datacenter operator using matched-energy padding and voltage/frequency shaping can hide at least **41% of unauthorized compute** within legitimate power envelopes unless verifiers actively re-execute sampled workloads at the observed operating point. Thus, smart-meter exergy telemetry cannot serve as a standalone cryptographic oracle and must be paired with sampled active replay.

2. **Distortion under linear scalarization and incomplete preference elicitation (Gölz et al. [9]).** Whenever multi-dimensional ecological or social objectives are aggregated via fixed linear weights inside sub-components of a clearing mechanism, Gölz et al. [9] prove that linear scalarization induces high social-choice distortion and fails to identify concave portions of the Pareto frontier, reinforcing the necessity of non-linear concave utility reporting or interactive Pareto navigation [15, 16, 22].

3. **Cyclic delegation loops in multi-agent harnesses (Fleischman and Buchman [7]).** In decentralized tool-calling networks, Fleischman and Buchman [7] demonstrate that autonomous agents frequently enter cyclic delegation patterns when local capacities are throttled: if Node A lacks watershed quota to complete an inference batch and delegates to Node B, which in turn delegates back to Node A, the network wastes exergy on unproductive handoff overhead unless the runtime conservation gate (1) enforces strictly positive hop attenuation $\boldsymbol{\delta}_{\text{hop}} \succ \mathbf{0}$ and acyclic call-graph lineage checks [25, 26].

4. **Information-theoretic communication overhead of iterative tâtonnement (Hurwicz [11]).** Classical mechanism-design results by Arrow and Hurwicz [2] and Hurwicz [11] establish that achieving exact Pareto optimality under private nonlinear utilities requires iterative price-quantity message exchanges. As shown in our K-step convergence sweep, if network latency or packet loss truncates the VB-LiRA dual negotiation to $K \leq 2$ rounds during a flash grid fault, allocative welfare drops from 96.80% to 41.42%, underperforming non-iterative DRF (71.75%).

7 Limitations

- **Upstream Scope-3 embodied carbon uncertainty.** While real-time electrical exergy (MWh) and on-site water withdrawal (kL) can be metered directly at sub-minute intervals, embodied carbon (tCO_2e) for imported hardware and raw materials relies on lifecycle assessment (LCA) inventories that carry ±15%–30% estimation variance.

- **Feeder-level clustering approximation for least-core settlement.** Solving the exact least-core linear program (4) requires $2^n - 2$ coalitional constraints (evaluated at $n = 8$ prosumers per feeder in our 400-game benchmark). For regional grids exceeding $n = 16$

prosumers, VB-LiRA must partition the grid into feeder-level sub-coalitions, trading global coalitional optimality for sub-millisecond LP tractability.

- **Separable log-utility assumption and dual iteration count under extreme drought.** Our 2,400-instance simulation models agent production utilities as separable across axes ($v_i(\mathbf{r}_i) = \sum_d w_{i,d} \ln(1 + r_{i,d})$) and shows that while $K = 12$ achieves a 3.67% nonconverged rate at $\kappa = 2.0$, extreme drought ($\kappa = 3.0$) increases the strict nonconverged rate (objective gap $> 5\%$) to 49.33%, requiring $K \geq 20$ iterations for full dual equalization. Workloads exhibiting strict Leontief complementarity ($v_i(\mathbf{r}_i) = \min_d\{r_{i,d}/\alpha_{i,d}\}$), for which DRF [8] is specifically designed, require coupling the primal demand step across axes.

8 Pre-registerable falsification protocol

Table 5 defines six pre-registerable empirical protocols designed to falsify the core claims of VB-LiRA.

Table 5: Six pre-registerable falsification protocols for Vector Biophysical Lagrangian Clearing (VB-LiRA).

Protocol	Hypothesis	Design and primary metrics	Falsified if
P1. Non-fungible drought compliance (Eq. 2)	Orthogonal vector clearing prevents compute clusters from breaching local watershed caps during carbon-offset surpluses	Simulate 600 summer drought epochs ($\kappa = 3.0$) where carbon credits are oversupplied while watershed cooling quota drops by 65%	Watershed withdrawal exceeds $B_3^{\max}(t)$ (beyond 3.5% meter tolerance) in $> 2.0\%$ of epochs
P2. Cyclic delegation containment (Eq. 1)	Runtime conservation gates [25] with hop attenuation $\boldsymbol{\delta}_{\text{hop}} \succ \mathbf{0}$ halt cyclic handoffs [7] within bounded exergy leases	Inject 200 cyclic delegation workloads into a 12-agent testbed; measure wasted handoff exergy vs. uncontracted baseline	Unproductive cyclic handoff exergy exceeds 5.0% of the parent lease $\mathbf{b}_i(t)$
P3. LiRA welfare gain over DRF and static quotas (Eq. 3)	Coupled per-agent VB-LiRA ($K = 12$) improves Pareto welfare by ≥ 18 percentage points over DRF [8] under heterogeneous civic priorities	Benchmark $n = 12$ heterogeneous civic, AI, and batch agents over 600 clearing windows at $\kappa = 2.0$	Net Pareto social welfare gain of VB-LiRA ($K = 12$) over DRF is < 18.0 percentage points
P4. Primal–dual iteration break-even (K-sweep)	VB-LiRA surpasses non-iterative DRF welfare (71.75%) within $K \leq 8$ inner primal–dual mirror-descent iterations	Sweep inner iterations $K \in \{1, 2, 4, 8, 12, 20, 40\}$ across 300 regional drought instances ($\kappa = 2.0$)	VB-LiRA requires $K > 8$ iterations to match DRF welfare (71.75%) or exceeds 5.0 ms solver latency
P5. Covert compute active-replay audit	Pairing analogue watt-hour telemetry with 10% sampled active re-execution [13] bounds covert compute below 10%	Red-team datacenter node using matched-energy padding [13] to hide unauthorized training rollouts	Adversary hides $> 10.0\%$ of unauthorized FLOPs without triggering audit rejection
P6. Least-core microgrid defection stability (Eq. 4)	Feeder-clustered least-core LP settlement produces non-positive core excess ($\epsilon^* \leq 0$) across solar/wind ramps	Evaluate 400 8-prosumer feeder deviation games comparing Pro-Rata, Exact Shapley, and Least-Core LP settlement	Least-core LP yields $\epsilon^* > 0$ on $> 5.0\%$ of convex balancing epochs where the core is non-empty

9 Conclusion

Grounding an economic and computational ledger in physical reality requires recognizing that megawatt-hours of peak grid exergy, tons of atmospheric carbon, and kiloliters of local watershed recharge are non-fungible biophysical boundaries [5, 19]. Our mechanistic evaluation across 2,400 paired regional clearing instances demonstrates that a cross-axis Fungible Scalar Offset straw model degrades to 56.83% [52.84, 60.74] boundary compliance at $\kappa = 2.0$ and 11.00% [8.74, 13.76] under acute drought ($\kappa = 3.0$), unbuffered Walrasian Multi-Price tâtonnement suffers 39.33% [35.50, 43.30] physical compliance under 2.4% sub-epoch meter variance, and Static Uniform Quotas and Dominant Resource Fairness [8] sacrifice 36.96% and 28.18% of the constrained Pareto optimum, respectively, by ignoring heterogeneous civic urgency. By formulating coupled per-agent Lagrangian Responsibility Allocation (VB-LiRA) with scale-normalized simplex responsibility updates [4] and buffered envelope projection, VB-LiRA sustains 99.67% [98.79, 99.91] three-axis boundary compliance while recovering $96.81\% \pm 0.09\%$ of the independent KKT optimum in $K = 12$ iterations (0.202 ms), while cooperative least-core LP settlement achieves 100.00% [99.05, 100.00] coalitional stability across 400 microgrid deviation games versus 56.25% [51.35, 61.03] under Shapley allocation.

Use of AI tools

Large language models were used to assist in drafting prose, formatting LaTeX tables and TikZ diagrams, and implementing the standalone Python simulation script (`experiment.py`). Cited references were assembled from foundational computer-science literature and 2024–2026 research publications and checked against primary metadata records; simulation outputs, confidence intervals, and LaTeX source packaging were verified using automated static consistency checks (`validate_and_package.py`).

References

[1] Joshua Achiam, David Held, Aviv Tamar, and Pieter Abbeel. Constrained policy optimization (CPO). International Conference on Machine Learning (ICML), pages 22–31, 2017. URL `https://arxiv.org/abs/1705.10528`.

[2] Kenneth J. Arrow and Leonid Hurwicz. On the stability of the competitive equilibrium, I. Econometrica, 26(4):522–552, 1958. URL `https://doi.org/10.2307/1907515`.

[3] Guanyu Cai, Ruiming Tian, Lang Yang, Zhouhong Ren, Jinliang Yuan, Lingkun Li, and Jiliang Wang. Is your NPU ready for LLMs? dissecting the hidden efficiency bottlenecks in mobile LLM inference. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 10(3):67, 2026. doi: 10.1145/3857370.

[4] Xiaoyang Cao, Jingqi Li, Zhe Fu, and Alexandre M. Bayen. Who bears the burden? learning responsibility for shared constraints in multi-agent reinforcement learning (LiRA). arXiv preprint arXiv:2610.07491, 2026. URL `https://arxiv.org/abs/2610.07491`.

[5] Herman E. Daly. Toward some operational principles of sustainable development. Ecological Economics, 2(1):1–6, 1990. URL `https://doi.org/10.1016/0921-8009(90)90010-R`.

[6] Uliana Elina. Maat: Independent deterministic contract-based governance for multi-agent LLM workflows. arXiv preprint arXiv:2609.34017, 2026. URL https://arxiv.org/abs/2609.34017.

[7] Tomaż Fleischman and Ethan Buchman. Deepening the secondary market: Integrating trade credit into market clearing with the Cycles protocol. arXiv preprint arXiv:2605.02436, 2026. URL https://arxiv.org/abs/2605.02436.

[8] Ali Ghodsi, Matei Zaharia, Benjamin Hindman, Andy Konwinski, Scott Shenker, and Ion Stoica. Dominant resource fairness: Fair allocation of multiple resource types. USENIX Symposium on Networked Systems Design and Implementation (NSDI), pages 323–336, 2011. URL https://www.usenix.org/conference/nsdi11/dominant-resource-fairness-fair-allocation-multiple-resource-types.

[9] Paul Gölz, Nika Haghtalab, and Kunhe Yang. Distortion of AI alignment: Does preference optimization optimize for preferences? In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, 2025. URL https://arxiv.org/abs/2505.23749.

[10] Lennart Heim, Tim Fist, Janet Egan, Sihao Huang, Stephen Zekany, Robert Trager, Michael A. Osborne, and Noa Zilberman. Governing through the cloud: The intermediary role of compute providers in AI regulation. Oxford Martin AI Governance Initiative, 2024. URL https://arxiv.org/abs/2403.08501.

[11] Leonid Hurwicz. On informationally decentralized systems. Decision and Organization, North-Holland, pages 297–336, 1972. URL https://www.nobelprize.org/uploads/2018/06/hurwicz-lecture.pdf.

[12] Yubin Kim, Ken Gu, Chanwoo Park, Chunjong Park, Samuel Schmidgall, A. Ali Heydari, et al. Towards a science of scaling agent systems, 2025. URL https://arxiv.org/abs/2512.08296.

[13] Tom Kimpson, Mauricio Baker, and Emlyn Graham. Can power draw constrain covert compute? limits of analogue verification for AI governance. arXiv preprint arXiv:2610.07476, 2026. URL https://arxiv.org/abs/2610.07476.

[14] Oskar Lange. On the economic theory of socialism: Part one. The Review of Economic Studies, 4(1):53–71, 1936. URL https://doi.org/10.2307/2967660.

[15] Junyi Li et al. Towards pareto-optimal tool-integrated agents with pareto ranking policy optimization (ParetoPO). International Conference on Machine Learning (ICML) / arXiv:2606.16111, 2026. URL https://arxiv.org/abs/2606.16111.

[16] Kaisa Miettinen. Nonlinear multiobjective optimization. Kluwer Academic Publishers, 1999. URL https://doi.org/10.1007/978-1-4615-5563-6.

[17] Isaac Ong, Amjad Almahairi, Vincent Wu, Wei-Lin Chiang, Tianle Li, Joseph E. Gonzalez, M. Waleed Kadous, and Ion Stoica. RouteLLM: Learning to route LLMs with preference data. International Conference on Learning Representations (ICLR), 2025. URL https://arxiv.org/abs/2406.18665.

[18] Elinor Ostrom. Governing the commons: The evolution of institutions for collective action. Cambridge University Press, 1990. URL https://doi.org/10.1017/CBO9780511807763.

[19] Johan Rockström, Will Steffen, Kevin Noone, Åsa Persson, F. Stuart Chapin, Eric F. Lambin, Timothy M. Lenton, Marten Scheffer, Carl Folke, Hans Joachim Schellnhuber, et al. A safe operating space for humanity. Nature, 461:472–475, 2009. URL https://doi.org/10.1038/461472a.

[20] Lloyd S. Shapley. A value for n-person games. Contributions to the Theory of Games, 2(28):307–317, 1953. URL https://doi.org/10.1515/9781400881970-018.

[21] Yonadav Shavit. What does it take to catch a chinchilla? verifying rules on large-scale neural network training via compute monitoring. arXiv preprint arXiv:2303.11341, 2023. URL https://arxiv.org/abs/2303.11341.

[22] Ralph E. Steuer and Eng-Ung Choo. An interactive weighted Tchebycheff procedure for multiple objective programming. Mathematical Programming, 26(3):326–344, 1983. URL https://doi.org/10.1007/BF02591870.

[23] Hao Wang and Ting Huang. Mingbird: A local-first agent harness enabling small open models to complete real tasks. arXiv preprint arXiv:2610.02001, 2026. URL https://arxiv.org/abs/2610.02001.

[24] Guibin Zhang, Yanwei Yue, Zhixun Li, Sukwon Yun, Guancheng Wan, Kun Wang, et al. Cut the crap: An economical communication pipeline for LLM-based multi-agent systems. In *International Conference on Learning Representations (ICLR)*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/bbc461518c59a2a8d64e70e2c38c4a0e-Abstract-Conference.html.

[25] Kaiyuan Zhang, Yuke Peng, Ke Jiang, and Yinqian Zhang. ActGov: Governing LLM agent actions via policy-constrained validation. arXiv preprint arXiv:2609.24446, 2026. URL https://arxiv.org/abs/2609.24446.

[26] Wei Zhao, Zhe Li, Peixin Zhang, and Jun Sun. ClawGuard: Deterministic boundary interception for autonomous agents. arXiv preprint arXiv:2604.11790, 2026. URL https://arxiv.org/abs/2604.11790.