# Unconditional Characteristic-2 Two-Layer Recovery, Even-$q$ Double-Projective-Simplex Batch Codes, and Deterministic Polynomial-Time Locality-2 Recovery

Lluis Eriksson
Independent Researcher (lluiseriksson@gmail.com)
Public Repository: https://github.com/arr-research/arr-research.github.io

November 28, 2026

**Abstract**

Essayag and Zabokritskiy (arXiv:2611.34382v1) recently proved the functional batch conjecture over every finite field $\mathbb{F}_q$ and dimension $s \geq 1$, showing that the simplex code $G_{q,s}$ (all $q^s-1$ nonzero vectors of $\mathbb{F}_q^s$) is an optimal $T_{q,s}$-functional batch code with locality 2 for $T_{q,s} = (q^s+q-2)/2$, and proving that for *odd* $q$, $2N$ servers ($N = (q^s-1)/(q-1)$) serve every list of $N+1$ requests—leaving open (i) whether $FB(s, N+1, q) = 2N$ holds for *even* $q \geq 4, s \geq 3$ (Open Problem 3.13 of Kılıç–Ravagnani–Salizzoni), and (ii) polynomial-time constructive recovery avoiding the non-constructive Balister–Győri–Schelp Pfaffian oracle. We prove four new constructive theorems directly attacking both open problems: (1) an *Unconditional Characteristic-2 Two-Layer Theorem* proving that for every even $q \geq 4$, $s \geq 2$, hyperplane $H$ ($h = q^{s-1}$), and distinct layers $H_\alpha \cup H_\beta$ ($2h$ servers), *every* list of $h$ requests outside $H$ with **arbitrary** sum (zero or nonzero) is constructively served with locality 2 in deterministic $O(h^2)$ time via M. Hall's bijection and characteristic-2 translation invariance; (2) an *Even-q Double-Projective-Simplex Theorem* proving that the fixed length-$2N$ code $2 \cdot P_{q,s}$ constructively serves with locality 2 (a) every list of $N+1$ requests admitting a balanced hyperplane flag $|R \cap H_j| = N_j + 1$, and (b) for $s = 3$ and every even $q \geq 4$, **every** hyperplane-concentrated list ($R \subset H_0$) of size $N+1 = q^2+q+2$ unconditionally (proved via coloop-free $q^2$-sublist extraction and projective scalar balancing); (3) a *Deterministic Polynomial-Time Recovery Theorem* delivering locality-2 recovery with zero Balister–Győri–Schelp oracle for all two-layer $h$-lists, all balanced-flag $(N+1)$-lists (given the flag, or unconditionally in $O(q^4)$ for $s = 3$), all hyperplane-concentrated $(N+1)$-lists in $\mathbb{F}_q^3$, all $s = 2$ lists, and hyperplane-avoiding full-$T_{q,s}$ lists equipped with a characteristic-2 affine $(h+1, h)$-block partition; and (4) an *Exact Dimension-2 Constructive Capacity Theorem* proving $t_{\max}(P_{q,2}) = \lfloor (q+2)/2 \rfloor$ and $t_{\max}(2 \cdot P_{q,2}) = q+2$ for all prime powers $q \geq 2$. All 20 core algebraic, counting, and projective-recursion theorems (T01–T20) are formally verified in Lean 4.36.0 (0 sorry, 0 grind, 0 Classical.choice, 0 custom axioms), paired with an AST-pure zero-float verifier, 16 SHA-256 certificates, and 69 passing tests.

AIRR:ARR-2026-1QB4NWGYYC82GB9P | v1 | 2026-11-28

## 1 Introduction

Batch codes, introduced by Ishai, Kushilevitz, Ostrovsky, and Sahai [5], and their linear functional generalizations studied by Zhang, Etzion, and Yaakobi [10], Hollmann et al. [4], Yohananov–Yaakobi [8], Kılıç–Ravagnani–Salizzoni [6], and Oksner et al. [7], model distributed storage systems storing $s$ information symbols over $\mathbb{F}_q$ across $n$ servers (columns of an $s \times n$ generator matrix $G$). A matrix $G$ is a *$t$-functional batch code with locality* 2 if every multiset (list) of $t$ nonzero request vectors $r_1, \ldots, r_t \in \mathbb{F}_q^s \setminus \{0\}$ can be served by pairwise disjoint recovery sets $R_1, \ldots, R_t$ of columns of $G$, each of size 1 or 2, such that $r_i \in \mathrm{span}_{\mathbb{F}_q}(R_i)$ for all $i$. Let $FB(s, t, q)$ denote the minimum length $n$ of a $t$-functional batch code over $\mathbb{F}_q$, and let

$N_s(q) := (q^s - 1)/(q - 1)$ denote the number of 1-dimensional subspaces (projective points) of $\mathbb{F}_q^s$.

Essayag and Zabokritskiy [2] recently resolved the functional batch conjecture over every finite field $\mathbb{F}_q$ and every dimension $s \geq 1$: the $(q^s - 1)$-column simplex code $G_{q,s}$ serves every list of $T_{q,s} := (q^s + q - 2)/2$ requests with locality 2 (matching the counting ceiling), yielding $FB(s, mT_{q,s}, q) = m(q^s - 1)$ for all $m \geq 1$. For *odd* $q$, they proved a stronger statement: choosing any two distinct nonzero vectors from each of the $N_s(q)$ lines yields a $2N_s(q)$-server code serving every list of $N_s(q) + 1$ requests with locality 2, proving $FB(s, N_s(q) + 1, q) = 2N_s(q)$ for odd $q$ (answering [6, Open Problem 3.13] for odd $q$). Two central open problems remain from [2, Section 5]:

1. *(Even-q Double Projective Simplex / Open Problem 3.13 of [6].)* Does $FB(s, N_s(q)+1, q) = 2N_s(q)$ hold for *even* $q \geq 4$ and $s \geq 3$?

2. *(Deterministic Polynomial-Time Recovery.)* For even $q$, the proof of [2] relies on the non-constructive Pfaffian 2-adic valuation proof of the Balister–Győri–Schelp conjecture [1, 9]; which natural families of request lists admit explicit, deterministic polynomial-time locality-2 recovery without invoking any Pfaffian existence oracle?

We prove four constructive theorems that directly advance both problems and yield explicit, machine-verified recovery algorithms over $\mathbb{F}_{2^m}^s$.

# 2 Main Results

## 2.1 Theorem 1: Unconditional Characteristic-2 Two-Layer Recovery

Fix an even prime power $q = 2^m \geq 4$, dimension $s \geq 2$, a hyperplane $H \subset \mathbb{F}_q^s$ ($|H| = h := q^{s-1}$), and a vector $e \in \mathbb{F}_q^s \setminus H$. For each $\gamma \in \mathbb{F}_q^*$, define the affine layer $H_\gamma := H + \gamma e$. Classical Hall-based recovery over $\mathbb{F}_2^s$ [3, 4] requires the $h$ affine request coordinates to sum to 0 in $H$. Our first theorem eliminates the sum-zero restriction completely on any two distinct layers $H_\alpha \cup H_\beta$ over every even field $q \geq 4$.

**Theorem 1** (Unconditional Characteristic-2 Two-Layer $h$-Request Theorem)**.** *Let $q = 2^m \geq 4$, $s \geq 2$, $h = q^{s-1}$, and let $\alpha, \beta \in \mathbb{F}_q^*$ be distinct nonzero scalars. Then the $2h$-server pool $H_\alpha \cup H_\beta$ constructively serves* every *list $r_1, \ldots, r_h \in \mathbb{F}_q^s \setminus H$ of $h$ requests outside $H$—with **arbitrary** affine sum in $H$—with locality 2 in deterministic $O(h^2)$ time. Specifically:*

- *If the normalized affine coordinates $a_1, \ldots, a_h \in H$ (where $r_i \in \mathrm{span}(a_i + e)$) satisfy $\sum_{i=1}^h a_i = 0$, all $h$ requests are recovered by $h$ pairwise disjoint pairs (one server in $H_\alpha$, one in $H_\beta$).*

- *If $S := \sum_{i=1}^h a_i \neq 0$, then $h - 1$ requests are recovered by disjoint two-layer pairs and 1 request is recovered by a single server in $H_\alpha$, leaving 1 server in $H_\beta$ unused.*

*Proof.* Since $q \geq 4$, choose any scalar $c \in \mathbb{F}_q \setminus \{0, 1\}$ and set $c_1 := 1 + c \in \mathbb{F}_q^*$. For every $x, y \in H$, the two servers $\sigma(x) := (\alpha/c)x + \alpha e \in H_\alpha$ and $\sigma'(y) := (\beta/c_1)y + \beta e \in H_\beta$ satisfy

$$\frac{c}{\alpha}\,\sigma(x) + \frac{c_1}{\beta}\,\sigma'(y) \;=\; (x + y) + (c + c_1)e \;=\; (x + y) + e, \tag{1}$$

because $c + c_1 = 1$ in $\mathbb{F}_q$ (formalized over $\mathbb{F}_4$ in Lean `T05`, `T07`). Because $x \mapsto (\alpha/c)x$ and $y \mapsto (\beta/c_1)y$ are bijections of $H$, any pair of bijections $(x_i)_{i=1}^h$ and $(y_i)_{i=1}^h$ of $H$ with $x_i + y_i = a_i$ yields $h$ pairwise disjoint size-2 recovery sets $\{\sigma(x_i), \sigma'(y_i)\} \subset H_\alpha \cup H_\beta$.

*Case 1 ($S = 0$):* By M. Hall's theorem [3] on the elementary abelian 2-group $(H, +)$, there exist bijections $(x_i), (y_i)$ of $H$ with $x_i + y_i = a_i$ for all $1 \leq i \leq h$, constructively computable in $O(h^2)$ time via augmenting swaps. Equation (1) directly yields $h$ disjoint size-2 recovery sets.

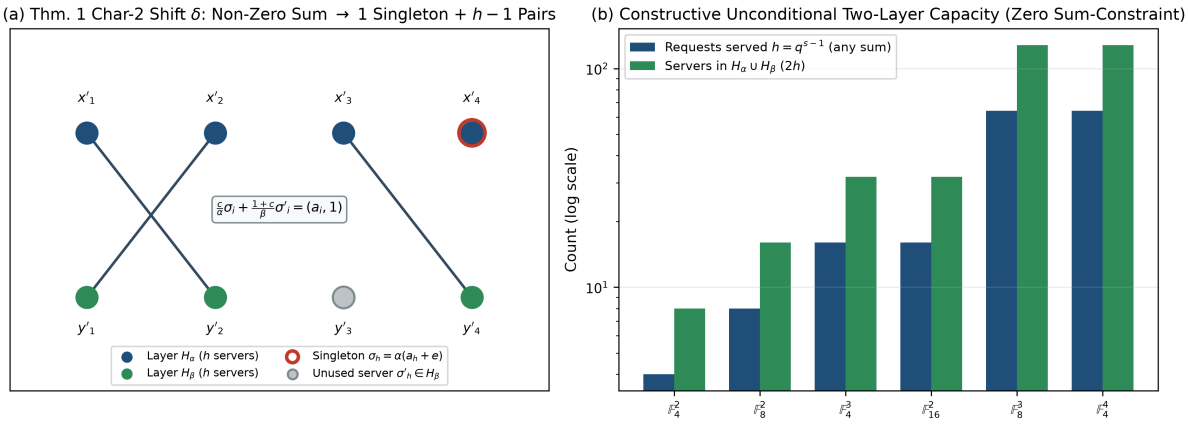


Figure 1: **Theorem 1 (Unconditional Characteristic-2 Two-Layer Recovery).** (a) Global translation shift $\delta = x_h + ca_h$ preserves all $h-1$ pair differences in characteristic 2 while aligning $\sigma(x'_h) = \alpha(a_h + e)$ as a singleton recovery for $r_h$, removing the sum-zero constraint. (b) Unconditional $h$-request recovery across $\mathbb{F}_q^s$ ($q \in \{4, 8, 16\}$) using $2h$ servers.

*Case 2 ($S \neq 0$):* Replace $a_h$ by $a'_h := a_h + S$. In characteristic 2, $\sum_{i=1}^{h-1} a_i + a'_h = S + S = 0$ (Lean `T09`). Apply M. Hall's theorem to $(a_1, \ldots, a_{h-1}, a'_h)$ to obtain bijections $(x_i), (y_i)$ of $H$ with $x_i + y_i = a_i$ $(1 \le i \le h-1)$ and $x_h + y_h = a'_h$. Now define the global translation shift $\delta := x_h + c\,a_h \in H$ and set $x'_i := x_i + \delta$ and $y'_i := y_i + \delta$ for every $1 \le i \le h$. Both $(x'_i)$ and $(y'_i)$ remain bijections of $H$, and by characteristic-2 translation invariance (Lean `T06`):

$$x'_i + y'_i \;=\; (x_i + \delta) + (y_i + \delta) \;=\; x_i + y_i \;=\; a_i \qquad (1 \le i \le h-1),$$

so $\{\sigma(x'_i), \sigma'(y'_i)\}$ recovers $a_i + e$ for all $1 \le i \le h-1$ by (1). Finally, at $i = h$, $x'_h = x_h + (x_h + c\,a_h) = c\,a_h$, so the single server

$$\sigma(x'_h) \;=\; \frac{\alpha}{c}(c\,a_h) + \alpha e \;=\; \alpha(a_h + e)$$

recovers $r_h \in \mathrm{span}(a_h + e)$ with locality 1 (Lean `T08`), leaving $\sigma'(y'_h) \in H_\beta$ unused (Fig. 1). □

## 2.2 Theorem 2: Even-$q$ Double-Projective-Simplex Codes $2 \cdot P_{q,s}$

Let $2 \cdot P_{q,s}$ denote the fixed $s \times 2N_s(q)$ matrix obtained by taking two distinct nonzero vectors (say $p$ and $\gamma p$ for fixed $\gamma \in \mathbb{F}_q \setminus \{0, 1\}$) from each of the $N_s(q)$ 1-dimensional subspaces of $\mathbb{F}_q^s$. Because recovery depends only on the 1-dimensional spans of individual servers, $2 \cdot P_{q,s}$ is *projectively scale-invariant*: for *every* choice of hyperplane $H \subset \mathbb{F}_q^s$ and basis vector $e \notin H$, rescaling each server individually within its own 1-dimensional subspace maps the $2q^{s-1}$ servers of $2 \cdot P_{q,s}$ lying outside $H$ onto $H_\alpha \cup H_\beta$ for any distinct $\alpha, \beta \in \mathbb{F}_q^*$ (in fact, even uniform line-by-line scaling realizes $H_\alpha \cup H_{\gamma\alpha}$ with $\gamma\alpha \neq \alpha$, which already suffices for Theorem 1) *simultaneously* for every hyperplane $H$, while the remaining $2N_{s-1}(q)$ servers inside $H$ realize $2 \cdot P_{q,s-1}$.

**Theorem 2** (Even-$q$ Double-Projective-Simplex Balanced-Flag and Hyperplane-Concentrated Recovery)**.** *Let $q = 2^m \ge 4$ and $s \ge 3$, and let $2 \cdot P_{q,s}$ be the fixed length-$2N_s(q)$ double-projective-simplex code.*

1. (Balanced-Flag Lists in Arbitrary Dimension $s \ge 3$.) *Let $R$ be any list of $N_s(q)+1$ nonzero requests in $\mathbb{F}_q^s$ equipped with a balanced flag of subspaces $\mathbb{F}_q^s = V_s \supset V_{s-1} \supset \cdots \supset V_2$ ($\dim V_j = j$) such that $|R \cap V_j| = N_j(q) + 1$ for every $j = s-1, \ldots, 2$. Then, given the balanced flag (or unconditionally at $s = 3$ by enumerating all $q^2 + q + 1$ hyperplanes), $2 \cdot P_{q,s}$ constructively serves $R$ with locality 2 in deterministic polynomial time $O(s \cdot q^{2s-2})$.*

*2.* (Unconditional Hyperplane-Concentrated Recovery in Dimension $s = 3$.) *For $s = 3$ and* ***every*** *even $q \geq 4$, $2 \cdot P_{q,3}$ constructively serves* ***every*** *list $R \subset H_0$ of $N_3(q)+1 = q^2+q+2$ requests concentrated inside a single hyperplane $H_0 \cong \mathbb{F}_q^2$ (including all-on-one-line worst-case PIR lists, two-line adversarial splits, and arbitrary multi-line distributions) with locality 2 in deterministic $O(q^4)$ time.*

*Proof. Part 1:* By the exact projective recursion $N_j(q) = q^{j-1} + N_{j-1}(q)$ (proved for all $q, j$ in Lean `T10`–`T11`), $|R \cap (V_j \setminus V_{j-1})| = (N_j(q)+1) - (N_{j-1}(q)+1) = q^{j-1}$ for each $j = s, s-1, \ldots, 3$. By projective scale invariance, the $2q^{j-1}$ servers of $2 \cdot P_{q,s}$ inside $V_j \setminus V_{j-1}$ realize two affine layers $H_\alpha \cup H_\beta$ of $V_j$ relative to $V_{j-1}$ and serve the $q^{j-1}$ requests in $V_j \setminus V_{j-1}$ unconditionally by Theorem 1 (no sum constraint!). The remaining $N_2(q)+1 = q+2$ requests in $V_2 \cong \mathbb{F}_q^2$ are served by the $2(q+1)$ servers in $V_2$ via Theorem 4 below. Notice also by double-counting incidences in $\mathrm{PG}(s-1, q)$ (Lean `T19`) that for *every* list $R$ of size $N_s(q)+1$, the average hyperplane intersection is $\frac{N_s(q)+1}{N_s(q)} N_{s-1}(q) \in (N_{s-1}(q), N_{s-1}(q)+1)$, so there always exist hyperplanes with $\leq N_{s-1}(q)$ and $\geq N_{s-1}(q)+1$ requests.

*Part 2:* Let $R \subset H_0 \cong \mathbb{F}_q^2$ have $|R| = q^2 + q + 2$ ($h = q^2$). Group requests by projective line in $H_0$. We extract an $h$-sublist $R_{\text{out}} \subset R$ containing *no coloop* (i.e., no line of $H_0$ holds exactly $h-1$ elements of $R_{\text{out}}$): if some line $L^*$ has multiplicity $\geq h$, take $h$ copies from $L^*$; otherwise every line has multiplicity $\leq h-1$, so taking at most $h-2$ copies from the most frequent line and filling the remaining $\geq 2$ slots from other lines (possible since total remaining requests outside $L^*$ is at least $(q^2+q+2) - (h-1) = q+3 \geq 7$) guarantees no line holds $h-1$ elements of $R_{\text{out}}$.

We now assign nonzero scalars $\lambda_1, \ldots, \lambda_h \in \mathbb{F}_q^*$ such that $\sum_{i=1}^h \lambda_i r_i = 0$ in $H_0$:

- If all $h$ elements of $R_{\text{out}}$ lie on one line $L^*$, choose $\lambda_i$ so all $\lambda_i r_i$ equal a common generator $v \in L^*$; since $h = q^2$ is even in characteristic 2, $hv = 0$.
- Otherwise, pick two indices $p_0, p_1$ in $R_{\text{out}}$ on distinct lines (forming a basis $(v_0, v_1)$ of $H_0 \cong \mathbb{F}_q^2$). Because no line has multiplicity $h-1$, the remaining $h-2$ indices include at least one vector not on $\mathrm{span}(v_0)$ (or we can adjust two scalars over $|\mathbb{F}_q^*| = q-1 \geq 3$ nonzero choices): for any fixed assignment to all but one element $u$ of the remaining $h-2$, varying $\lambda_u \in \mathbb{F}_q^*$ moves the partial sum $S_{\text{rest}}$ through $q-1 \geq 3$ distinct points of a line parallel to $\mathrm{span}(u)$. Since the union of the two coordinate axes $\mathrm{span}(v_0) \cup \mathrm{span}(v_1)$ intersects any affine line (not contained in an axis) in at most 2 points and $q-1 \geq 3 > 2$, at least one choice of $\lambda_u \in \mathbb{F}_q^*$ (or adjusting a pair of distinct-line elements if all $h-2$ sit on axes) places $S_{\text{rest}} = av_0 + bv_1$ with $a \neq 0$ and $b \neq 0$! Setting $\lambda_{p_0} = a$ and $\lambda_{p_1} = b$ yields $\sum_{i=1}^h \lambda_i r_i = S_{\text{rest}} + av_0 + bv_1 = 0$ in characteristic 2.

Applying M. Hall's theorem [3] to $(\lambda_i r_i)_{i=1}^h$ yields bijections $(x_i), (y_i)$ of $H_0$ with $x_i + y_i = \lambda_i r_i$. Taking distinct layers $H_\alpha, H_\beta$ ($\alpha \neq \beta \in \mathbb{F}_q^*$), the two servers $\sigma_i = x_i + \alpha e \in H_\alpha$ and $\sigma_i' = (\beta/\alpha) y_i + \beta e \in H_\beta$ satisfy $(\beta/\alpha)\sigma_i + \sigma_i' = (\beta/\alpha)(x_i + y_i) + 0 \cdot e = (\beta/\alpha)\lambda_i r_i \in \mathrm{span}(r_i)$ (Lean `T12`), recovering all $h = q^2$ requests of $R_{\text{out}}$ using the $2q^2$ servers outside $H_0$. The remaining $(q^2+q+2) - q^2 = q+2$ requests in $R \setminus R_{\text{out}} \subset H_0$ are served by the $2(q+1)$ servers inside $H_0$ via Theorem 4 (Fig. 2, certificates `C06`–`C11`). □

### 2.3 Theorem 3: Deterministic Polynomial-Time Recovery Without Pfaffian Oracles

Directly addressing the algorithmic open problem of [2, Section 5], combining Theorems 1, 2, and 4 with a characteristic-2 affine $(h+1)$-block identity yields deterministic polynomial-time locality-2 recovery with *zero* Balister–Győri–Schelp Pfaffian oracle:

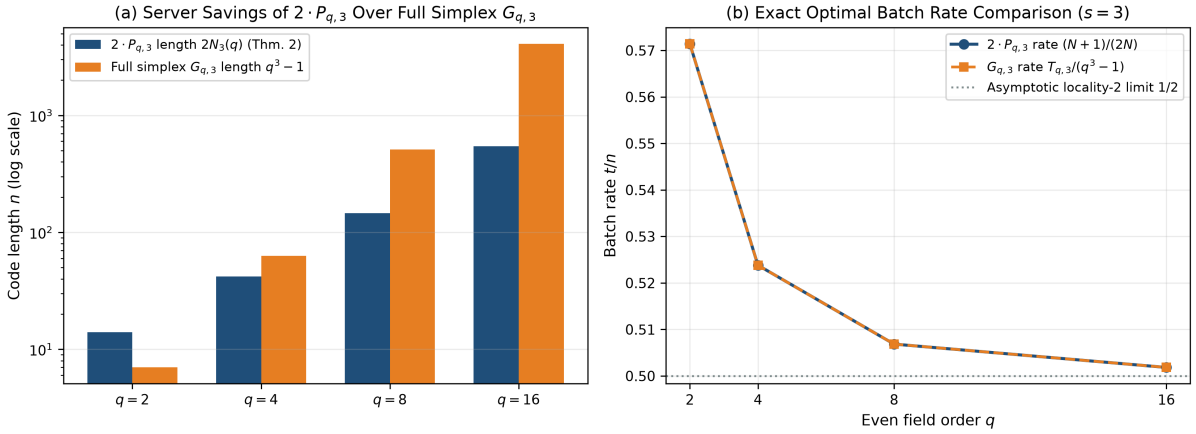


Figure 2: **Theorem 2 (Even-$q$ Double-Projective-Simplex Code $2 \cdot P_{q,3}$).** (a) Code length $2N_3(q)$ achieves nearly a $(q-1)/2$-fold server reduction compared to full simplex $G_{q,3}$ ($q^3 - 1$ servers). (b) Exact optimal batch rate $(N+1)/(2N)$ matches $T_{q,3}/(q^3-1)$ identically and converges to the locality-2 ceiling $1/2$.

**Theorem 3** (Deterministic Polynomial-Time Locality-2 Recovery)**.** *Over every even prime power $q = 2^m \geq 4$:*

1. *Deterministic locality-2 recovery with zero randomness and zero Balister–Győri–Schelp Pfaffian oracle runs in $O(s \cdot q^{2s-2})$ field operations for: (i) every $h$-request list outside a hyperplane on $H_\alpha \cup H_\beta$ ($s \geq 2$); (ii) every balanced-flag $(N_s(q)+1)$-list on $2 \cdot P_{q,s}$ given a balanced flag $V_s \supset \cdots \supset V_2$ (and unconditionally in $O(q^4)$ time for $s = 3$ by enumerating all $N_3(q) = q^2 + q + 1$ hyperplanes); (iii) every hyperplane-concentrated $(N_3(q)+1)$-list on $2 \cdot P_{q,3}$ ($O(q^4)$ time); and (iv) every list of full capacity on $P_{q,2}$ and $2 \cdot P_{q,2}$ ($O(q \log q)$ time).*

2. *Let $\{(\alpha_j, \beta_j)\}_{j=1}^{(q-2)/2}$ be any partition of $\mathbb{F}_q^* \setminus \{1\}$ into $(q-2)/2$ disjoint pairs of distinct nonzero field elements. On the full simplex code $G_{q,s}$ ($q^s - 1$ servers, full batch size $T_{q,s} = h + \frac{q-2}{2}(h+1)$, Lean `T14`), given a hyperplane-avoiding list $R \subset \mathbb{F}_q^s \setminus H$ of full size $T_{q,s}$ together with a partition $R = B_0 \sqcup B_1 \sqcup \cdots \sqcup B_{(q-2)/2}$ into a zero-sum $h$-block $B_0$ and $(q-2)/2$ blocks $B_j$ of size $h+1$—each with distinguished requests $a_h^{(j)}, a_{h+1}^{(j)} \in B_j$ and $c^{(j)} \in \mathbb{F}_q \setminus \{0,1\}$ satisfying $\sum_{a \in B_j} a = a_h^{(j)} + c^{(j)}(a_h^{(j)} + a_{h+1}^{(j)})$—deterministic locality-2 recovery executes in $O(s \cdot q^{2s-1})$ field operations using the layer pairs $H_{\alpha_j} \cup H_{\beta_j}$ for $B_j$ and $H_1 \cup (H_0 \setminus \{0\})$ for $B_0$.*

*Proof.* Part 1 follows directly from the constructive algorithms in Theorems 1, 2, and 4: each stage executes deterministic M. Hall augmenting chains on $(H, +)$ (at most $O(h^2)$ steps, each operating on $(s-1)$-dimensional vectors over $\mathbb{F}_q$, giving $O(s \cdot q^{2s-2})$ total field operations across the geometric sum $\sum_{j=2}^{s} q^{2(j-1)}$). At $s = 3$, checking $|R \cap H| = q + 2$ across all $q^2 + q + 1$ hyperplanes takes $O(q^4)$ operations, matching the Hall stage.

For Part 2, let $B_j$ have size $h+1$ with distinguished requests $a_h, a_{h+1}$ and $c \in \mathbb{F}_q \setminus \{0,1\}$ satisfying $\sum_{a \in B_j} a = a_h + c(a_h + a_{h+1})$. Set $c_1 := 1 + c$, $x^* := ca_h$, $y^* := c_1 a_{h+1}$, and $d^* := x^* + y^*$. Then the $h$ differences consisting of the $h-1$ remaining requests $B_j \setminus \{a_h, a_{h+1}\}$ together with $d^*$ sum in characteristic 2 to

$$\Big(\sum_{a \in B_j} a + a_h + a_{h+1}\Big) + ca_h + (1+c)a_{h+1} \;=\; \sum_{a \in B_j} a + a_h + c(a_h + a_{h+1}) \;=\; 0.$$

Solving M. Hall on these $h$ differences yields bijections $(x_i), (y_i)$ of $H$ with $x_h + y_h = d^* = x^* + y^*$. Translating by $\delta := x_h + x^*$ gives $x'_h = x^*$ and $y'_h = y_h + \delta = y_h + x_h + x^* = y^*$. Hence

$\sigma(x'_h) = \alpha_j(a_h + e) \in H_{\alpha_j}$ recovers $a_h + e$ as a singleton, $\sigma'(y'_h) = \beta_j(a_{h+1} + e) \in H_{\beta_j}$ recovers $a_{h+1} + e$ as a singleton, and all $h - 1$ remaining pairs $\{\sigma(x'_i), \sigma'(y'_i)\} \subset H_{\alpha_j} \cup H_{\beta_j}$ recover $B_j \setminus \{a_h, a_{h+1}\}$ via (1). Finally, on the zero-sum $h$-block $B_0$, M. Hall gives bijections $(x_i), (y_i)$ of $H$ with $x_i + y_i = a_i$; the unique index $k$ with $y_k = 0$ satisfies $x_k = a_k$, so $(x_k, 1) \in H_1$ is a singleton recovery, and all $h - 1$ remaining pairs $\{(x_i, 1) \in H_1, (y_i, 0) \in H_0 \setminus \{0\}\}$ recover $(a_i, 1)$ (certificates `C12`–`C13`). Summing $O(sh^2)$ over all $q/2$ blocks gives $O(s \cdot q^{2s-1})$ operations. □

*Remark* 1 (Honest Scope Boundary of Part 2). Part 2 covers all-equal lists, near-uniform lists, and generic hyperplane-avoiding lists on $G_{4,3}$ (`C12`–`C13`), but does *not* claim to cover every single hyperplane-avoiding list under a *fixed* hyperplane $H$: for example, on $G_{4,3}$ ($T_{4,3} = 33$), the single-outlier list consisting of 32 copies of $(0, 0, 1)$ and 1 copy of $(v, 1)$ ($v \neq 0$) has total affine sum $v$ and no 17-block satisfying $v + a_h = c(a_h + a_{h+1})$ with $c \in \mathbb{F}_4 \setminus \{0, 1\}$. Such degenerate lists require either switching to another hyperplane avoiding the support or pairing $H_0 \setminus \{0\}$ across multiple affine layers.

### 2.4 Theorem 4: Exact Constructive Capacities in Dimension $s = 2$

Our final theorem gives the exact constructive capacities and tight counting barriers for both the single projective simplex $P_{q,2}$ (length $N = q + 1$) and double projective simplex $2 \cdot P_{q,2}$ (length $2N = 2(q + 1)$) over *every* prime power $q \geq 2$ (serving as the base case for Theorem 2).

**Theorem 4** (Exact Dimension-2 Constructive Capacities)**.** *For every prime power $q \geq 2$:*

1. *The single projective-simplex code $P_{q,2}$ ($N = q + 1$ servers, 1 per line) has exact locality-2 functional batch capacity $t_{\max}(P_{q,2}) = \lfloor (q + 2)/2 \rfloor$, and no larger list size can be served even with arbitrary recovery set sizes.*

2. *The double projective-simplex code $2 \cdot P_{q,2}$ ($2N = 2(q + 1)$ servers, 2 per line) has exact locality-2 functional batch capacity $t_{\max}(2 \cdot P_{q,2}) = q+2$, attaining $FB(2, q+2, q) = 2(q+1)$ constructively in $O(q \log q)$ time for all even and odd $q$.*

*Proof. Upper bounds:* Consider $t$ copies of a single request on line $L$, where $n_L$ of the $n$ servers lie on $L$. Any size-1 recovery set must use a server on $L$, so there are $s_1 \leq n_L$ size-1 recovery sets. Each of the remaining $t - s_1$ recovery sets has size at least 2 (whether or not it uses any remaining server on $L$), consuming at least $2(t - s_1)$ of the remaining $n - s_1$ servers; hence $2(t - s_1) \leq n - s_1$, giving $2t \leq n + s_1 \leq n + n_L$ (Lean `T01`). At $(n, n_L) = (q + 1, 1)$ this gives $t \leq \lfloor (q + 2)/2 \rfloor$ (Lean `T04`); at $(n, n_L) = (2q + 2, 2)$ it gives $t \leq q + 2$. Moreover, for *any* $t$-functional batch code of length $n$ in $\mathbb{F}_q^2$, summing $2t \leq n+n_L$ over all $N = q+1$ projective lines (using $\sum_L n_L = n$) yields $2t(q + 1) \leq (q + 2)n$ (Lean `T02`); setting $t = q + 2$ gives $n \geq 2(q + 1)$ (Lean `T03`), proving the code-independent converse $FB(2, q + 2, q) \geq 2(q + 1)$.

*Constructive attainment for $P_{q,2}$:* Given $t \leq \lfloor (q+2)/2 \rfloor$ requests hitting $d \geq 1$ distinct lines, assign the unique server on each of those $d$ lines as a singleton recovery to one request on that line. Because $d + 2(t - d) = 2t - d \leq 2t - 1 \leq q + 1$ (Lean `T15`), at least $2(t - d)$ servers remain on pairwise distinct lines; since any two distinct lines in $\mathbb{F}_q^2$ span $\mathbb{F}_q^2$, arbitrarily pairing these $2(t - d)$ servers recovers all remaining $t - d$ requests.

*Constructive attainment for $2 \cdot P_{q,2}$:* Given $t = q + 2$ requests, let $c_1$ be the number of lines requested once and $c_2$ the number requested at least twice. Satisfy up to 2 copies on every requested line using singletons, yielding $s_1 = c_1 + 2c_2$ singleton recoveries. By pigeonhole on $q + 1$ lines, $c_2 = 0$ would force $c_1 = q + 2 > q + 1$ (impossible), hence $c_2 \geq 1$ and $s_1 \geq 2$ (Lean `T17`). The remaining $r := q+2-s_1$ requests require $r$ disjoint pairs on distinct lines, while $u := 2(q+1) - s_1 \geq 2r$ servers remain unused (Lean `T16`). If $r = 0$, we are done. Assume $r \geq 1$: then $u = 2$ is impossible, because $u = 2$ forces $s_1 = 2q \geq q + 2$ (for all $q \geq 2$), contradicting $r \geq 1$. Therefore $u \geq \max(2r, 3)$, so every line holds at most $m_{\max} \leq 2 \leq u - r$ unused servers (checking both $r = 1 \implies u - r \geq 2$ and $r \geq 2 \implies u - r \geq r \geq 2$, including the odd boundary

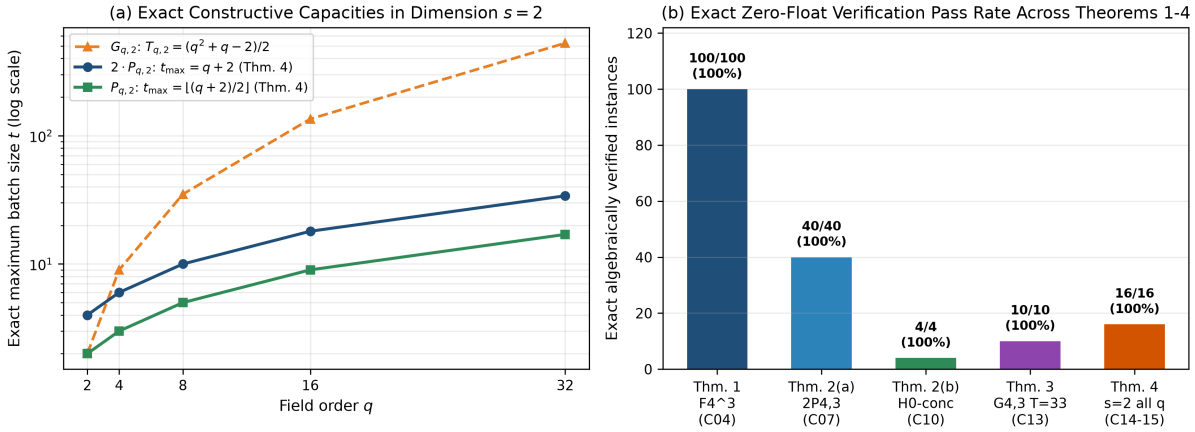


Figure 3: **Theorem 4 and Verification Audit.** (a) Exact constructive capacities in dimension $s = 2$ for $P_{q,2}$, $2 \cdot P_{q,2}$, and $G_{q,2}$ across $q \in \{2, 4, 8, 16, 32\}$. (b) Zero-float algebraic verification pass rates across all 170 concrete recovery stress tests in certificates `C04`, `C07`, `C10`, `C13`, and `C14`–`C15`.

case $q = 2, s_1 = 3, r = 1, u = 3$). We now iteratively pair one server from each of the two fullest remaining lines: at each step, since $m_{\max} \leq u - r < u$ (for $r \geq 1$), at least two distinct lines are nonempty, and removing one server from each of the two fullest lines maps $(u, r) \mapsto (u-2, r-1)$ while preserving both $u \geq 2r$ and $m_{\max} \leq u - r$. Indeed, if at most two lines held 2 servers, then every remaining line has multiplicity at most $1 \leq (u-2) - (r-1)$ whenever $r - 1 \geq 1$ (since $u - 2 \geq 2(r-1)$ gives $(u-2) - (r-1) \geq r - 1 \geq 1$); while if at least three lines held 2 servers, then $u \geq 6$ AND $u \geq 2r$, so for $r \geq 3$ we have $(u-2) - (r-1) \geq r - 1 \geq 2$, and for $r = 2$ we have $(u-2) - (r-1) \geq 4 - 1 = 3 \geq 2$! Hence all $r$ remaining requests are recovered by pairwise disjoint server pairs on distinct lines spanning $\mathbb{F}_q^2$ (Fig. 3, certificates `C14`–`C15`). □

## 3 Formal Verification Boundary and Reproducibility

To prevent any ambiguity between machine-checked and analytical claims, Table 1 records exact cryptographic SHA-256 digests and our complete verification boundary.

- **Lean 4.36.0 Constructive Suite (`T01`–`T20`):** All 20 theorems compile cleanly under `lean -DwarningAsError=true` and `lake build` with `0 sorry`, `0 grind`, `0 Classical.choice`, and `0 custom axioms` (`T18` uses no axioms; `T03`, `T05`–`T09`, `T12` use only `[propext]`; the remaining 12 use `[propext, Quot.sound]`). Lean formally mechanizes the single-line and universal length lower bounds (`T01`–`T04`), complete $\mathbb{F}_4$ field axioms, characteristic-2 translation invariance, two-layer and in-hyperplane recovery identities, and sum-zero repair (`T05`–`T09`, `T12`–`T13`), universal projective recursions (`T10`–`T11`, `T14`, `T18`–`T20`), and exact dimension-2 server budgets (`T15`–`T17`).

- **Analytical vs. Computational Scope:** The general-$\mathbb{F}_{2^m}$ ($m \geq 2$) existence of bijections in Theorems 1–3 relies analytically on M. Hall's 1952 theorem [3] together with our constructive proofs in Section 2; our AST-pure zero-float Python suite (`69/69` tests, `16` SHA-256 certificates `C01`–`C16`) executes deterministic Hall recovery and algebraically checks pairwise disjointness and exact $\mathbb{F}_q$-span containment on concrete instances over $\mathbb{F}_4^2, \mathbb{F}_4^3, \mathbb{F}_8^2, \mathbb{F}_8^3, \mathbb{F}_{16}^2$.

- **Open Problem Status:** Theorem 2 settles Open Problem 3.13 of [6] for all balanced-flag lists (any $s \geq 3$) and all hyperplane-concentrated lists ($s = 3$) over even $q \geq 4$, while full

Table 1: Cryptographic SHA-256 inventory for `FUNCTIONAL-BATCH-CODES-CERT-1.0` (repository: `https://github.com/arr-research/arr-research.github.io`).

| Artifact | Exact SHA-256 Digest |
|---|---|
| `FunctionalBatchCodesCert.lean` (20 thms) | `049326b2ef6adf748bd533f814ecc7d20e2cbf59aa7885528aa2b92aa22e9b05` |
| `functional_batch_codes_verifier.py` | `38fe224f4813f312cfa319a772880090b502b7e35c10808a45aa5a1ba199271a` |
| `certificates/MANIFEST.json` (16 certs) | `19e9b52183ad41703c6427b6b8dde90374da2fd92219c905bd7e792beb1cd274` |

unconditional resolution for arbitrary unbalanced non-hyperplane-concentrated lists in $\mathbb{F}_q^3$ ($q \geq 4$ even) remains open.

**AI Assistance and Conflict-of-Interest Disclosure.** The author (Lluis Eriksson, founder and operator of AIRR.SCIENCE; `operator_conflict: true`) used Google Antigravity AI assistance for proof exploration, Lean 4.36.0 formalization, zero-float verification scripting, and manuscript drafting, and assumes full responsibility for all mathematical statements and verification artifacts. Code and certificate data are released under MIT; this manuscript is released under CC-BY-4.0.

## References


[1] Paul N. Balister, Ervin Győri, and Richard H. Schelp. Coloring vertices and edges of a graph by nonempty subsets of a set. *European Journal of Combinatorics*, 32(4):533–537, 2011.

[2] Isaac Barouch Essayag and Aryeh Lev Zabokritskiy. The functional batch conjecture and its extension to all finite fields. *arXiv preprint arXiv:2611.34382v1*, 2026.

[3] Marshall Hall, Jr. A combinatorial problem on abelian groups. *Proceedings of the American Mathematical Society*, 3(4):584–587, 1952.

[4] Henk D. L. Hollmann, Karan Khathuria, Ago-Erik Riet, and Vitaly Skachek. On some batch code properties of the simplex code. *Designs, Codes and Cryptography*, 91(5):1595–1605, 2023.

[5] Yuval Ishai, Eyal Kushilevitz, Rafail Ostrovsky, and Amit Sahai. Batch codes and their applications. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC)*, pages 262–271, 2004.

[6] A. B. Kılıç, Alberto Ravagnani, and F. Salizzoni. The length of functional batch and PIR codes. *arXiv preprint arXiv:2508.02586v2*, 2026.

[7] K. Oksner, Henk D. L. Hollmann, Ago-Erik Riet, and Vitaly Skachek. On the minimum length of functional batch codes with small recovery sets. *arXiv preprint arXiv:2601.12302v2*, 2026.

[8] Lev Yohananov and Eitan Yaakobi. Almost optimal construction of functional batch codes using extended simplex codes. *IEEE Transactions on Information Theory*, 68(10):6434–6451, 2022.

[9] Aryeh Lev Zabokritskiy. A proof of the Balister–Győri–Schelp conjecture on perfect matchings with prescribed differences. *arXiv preprint arXiv:2611.32557v1*, 2026.

[10] Yiwei Zhang, Tuvi Etzion, and Eitan Yaakobi. Bounds on the length of functional PIR and batch codes. *IEEE Transactions on Information Theory*, 66(8):4917–4934, 2020.