

SMALL EXTENDED FORMULATIONS FOR GROUP-CONSTRAINED FLOWS AND COGRAPHIC INTEGER HULLS

AVINASH BHARDWAJ

ABSTRACT. We construct small extended formulations for nonnegative integer flows whose arc labels have a prescribed sum in a finite abelian group. For circulations, the formulation is a single integral network that combines cycles even when they lie in different components. This yields an $O(\Delta n^2)$ bound for strictly Δ -modular cographic integer hulls of translated cones, improving the previous $O(n^\Delta)$ bound for fixed Δ . We also show that polynomial dependence on Δ is necessary, up to the exponent. For flows with prescribed supplies and demands, we obtain two further bounds. The first is exponential in the number of supply and demand vertices and polynomial in the binary encoding length of their balances. The second is polynomial in graph size and group order whenever the number of supply vertices is fixed; its inequality count is independent of the demand magnitudes.

1. INTRODUCTION

Paat, Walsh, and Xu [27] proved that the integer hull of a strictly Δ -modular translated cone in $\mathbb{R}^n$ with a cographic row matroid has an extended formulation of size $O(n^\Delta)$ for fixed Δ . We improve their result and provide a bound of $O(\Delta n^2)$. The argument constructs a single integral network extension for group-constrained circulations. For prescribed balances, we obtain bounds exponential in the number of terminals and polynomial in binary balance length. A second construction gives polynomial-size formulations whenever the number of supply vertices is fixed, with no restriction on the number or magnitude of the demands.

For an integer matrix $A \in \mathbb{Z}^{m \times n}$ of column rank n , write

$$P(A, b) = \{x \in \mathbb{R}^n : Ax \leq b\}, \quad P_I(A, b) = \text{conv}(P(A, b) \cap \mathbb{Z}^n).$$

We write $\text{xc}(P)$ for the minimum number of inequalities in an extended formulation of P ; equations and affine projection maps are free. The matrix A is *strictly Δ -modular* if every nonzero maximal minor has absolute value Δ . Our structural assumption is that the row matroid $M(A^\top)$ is cographic. If $b \in \text{col}(A)$, all inequalities are tight at a common point and $P(A, b)$ is a translated cone.

Theorem 1. *Let $A \in \mathbb{Z}^{m \times n}$ have rank $n \geq 1$, and let $b \in \mathbb{Z}^m \cap \text{col}(A)$. If A is strictly Δ -modular and $M(A^\top)$ is cographic, then*

$$(1) \quad \text{xc}(P_I(A, b)) \leq \Delta m(m - n + 1).$$

After deleting repeated inequalities and zero rows, one obtains

$$(2) \quad \text{xc}(P_I(A, b)) \leq 36\Delta n^2.$$

These formulations can be constructed in time polynomial in Δ and the binary encoding length of (A, b) .

For $\Delta \geq 3$, this improves the dimension dependence in [27]. For $\Delta = 2$, it retains the quadratic bound; for $\Delta = 1$, the polyhedron is already integral and the reduced system itself has $O(n)$ inequalities. The construction time is polynomial in the numerical value of Δ .

Date: 4 October 2026.

2020 Mathematics Subject Classification. Primary 90C57; Secondary 90C10, 05C21, 05B35.

Key words and phrases. Extension complexity, integer hull, finite abelian group, network flow, cographic matroid, bounded subdeterminants.

The dependence on the two parameters has corresponding lower bounds. For every Δ , the choice $A = \text{diag}(\Delta, 1, \dots, 1)$ and $b = 0$ gives the negative orthant, of extension complexity n . Proposition 36, using the matching lower bound of Jia, Svensson, and Yuan [18], gives a family with $n = O((\log \Delta)^2)$ and extension complexity $\Delta^{\Omega(1)}$. Thus no uniform bound $\Delta^{o(1)} \text{poly}(n)$ is possible: polynomial dependence on Δ is necessary, up to the exponent.

The main step is a formulation for group-constrained circulations. Let $G = (V, E)$ be a directed multigraph with $V \neq \emptyset$ and node-arc incidence matrix D . Let Γ be a finite abelian group of order q , let $\gamma_e \in \Gamma$ label each arc, and fix $g \in \Gamma$. For $d \in \mathbb{Z}^V$ with $\sum_v d_v = 0$, define

$$(3) \quad S_{d,g} = \left\{ y \in \mathbb{Z}_{\geq 0}^E : Dy = d, \sum_{e \in E} y_e \gamma_e = g \right\}, \quad Q_{d,g} = \text{conv } S_{d,g}.$$

There are no arc capacities. For zero balances, abbreviate $S_g = S_{0,g}$ and $Q_g = Q_{0,g}$.

Theorem 2. *Suppose $V \neq \emptyset$ and $|\Gamma| = q$. There is a directed multigraph K with $q|V||E|$ arcs, a coordinate-summing map $\pi : \mathbb{R}^{E(K)} \rightarrow \mathbb{R}^E$, and distinguished vertices h_a , $a \in \Gamma$, such that*

$$(4) \quad Q_g = \pi(F_g), \quad F_g = \{z \in \mathbb{R}_{\geq 0}^{E(K)} : Bz = \mathbf{e}_{h_g} - \mathbf{e}_{h_0}\},$$

where B is the incidence matrix of K . Moreover,

$$(5) \quad \pi(F_g \cap \mathbb{Z}^{E(K)}) = S_g.$$

In particular, $\text{xc}(Q_g) \leq q|V||E|$.

A nonnegative integer circulation decomposes into directed cycles, but those cycles need not meet. A group-expanded graph records the label of a walk; it does not, by itself, combine cycles at unrelated locations. Our auxiliary graph has one copy for each possible base vertex. Copies meet only at vertices that represent a completed closed walk with a given group label. This permits successive choices of bases while preserving circulation conservation under projection.

Nonzero balances introduce an additional choice: which supply units serve which demand units. Writing $k = \frac{1}{2}\|d\|_1$ for total supply, we construct an integral network extension with at most

$$q(|V| + k2^{k-1})|E|$$

inequalities, with $k2^{k-1} = 0$ at $k = 0$; see Theorem 11. The network records completed assignments by a subset of demand tokens. For $q = 2$, the odd-red bipartite matching family of Jia, Svensson, and Yuan [18] has extension complexity $2^{\Omega(k)}$, so exponential dependence on total supply is unavoidable in general.

Total supply can be exponentially larger than its encoding length. We obtain binary-length bounds by exploiting the distribution of supplies and demands. When there is a single source, paths can be grouped by their destination. Group convolution and binary doubling then give an extended formulation polynomial in q , graph size, and binary demand length, even with arbitrarily many sinks; the exact bound is in Theorem 15. A second construction below removes the demand-magnitude dependence at the cost of a larger dependence on group parameters; Corollary 30 compares the two bounds. Reversing the arcs gives the corresponding single-sink results. More generally, we can bound the total number of supply and demand vertices.

Call a vertex with $d_v \neq 0$ a *terminal*. Let s and p be the numbers of supply and demand vertices, respectively, and put

$$t = s + p, \quad h = \min(s, p), \quad M = \max(1, \|d\|_\infty), \quad L_t = \lceil \log_2 M \rceil.$$

We use the convention $h2^{h-1} = 0$ when $h = 0$.

Theorem 3. *For every d and g ,*

$$(6) \quad \text{xc}(Q_{d,g}) \leq 2^t q(|V| + h2^{h-1})|E| + 4^t q^2 L_t.$$

The formulation is explicit, also for empty or unbounded hulls. Its size and construction time are bounded by a function of t, q times a polynomial in graph size and binary balance length. In particular,

$$\text{xc}(Q_{d,g}) \leq 2^{O(t)} q^2 (|V||E| + \log(1 + \|d\|_\infty)).$$

The proof splits an integer flow into two integer flows whose balances round half the original balance. At each depth, at most 2^t distinct balance vectors occur. A shared formulation represents each state once, and after logarithmically many depths the remaining balances are in $\{-1, 0, 1\}^V$. The odd-red matching family has $t = 2k$ and unit balances, so its known lower bound makes exponential dependence on t necessary up to constants in the exponent and polynomial factors. Through the slack equation $Dy = Db$, these results also give cographic bounds for arbitrary integer right-hand sides.

A different decomposition permits arbitrarily many sinks. Let ρ denote the exponent of Γ , so $\rho\alpha = 0$ for every $\alpha \in \Gamma$ and $\rho \leq q$.

Theorem 4. *Suppose d has $s \geq 1$ supply vertices and $p \geq 1$ demand vertices. Then*

$$(7) \quad \text{xc}(Q_{d,g}) \leq spq(|V| + 1)|E| + O_s(p^{2s-1}q^{2s+1}\rho^s),$$

where the implied constant depends only on s . In particular, for two supply vertices,

$$(8) \quad \text{xc}(Q_{d,g}) \leq 2pq(|V| + 1)|E| + O(p^3q^5\rho^2) = O(pq(|V| + 1)|E| + p^3q^7).$$

For fixed s , the formulation can be constructed in time polynomial in q , graph size, and binary balance length. The inequality count is independent of the balance magnitudes. The same conclusions hold with supply and demand reversed.

The proof first replaces the original flow by a transportation array of one-unit source–sink flows, indexed by their group labels. From each column we extract a multiple of ρ , leaving a residual demand smaller than ρsq . The extracted array has label zero and is handled by ordinary transportation constraints. The remaining group constraint is handled by a counting network. For fixed s , both the number of residual supply vectors and the size of each counting network are polynomial.

Table 1 compares the flow bounds. They exploit different features of the balance vector: total supply counts individual path units, terminal count controls the number of rounded balance states, and source count controls the dimension of the residual supply vectors. A large demand at one vertex increases the first parameter but not the other two. Before extracting the bounded residual, the same counting network gives a bound polynomial in numerical total supply for fixed source count (Proposition 25).

The circulation formulation is the point of contact with the closest predecessors. Conforti, Fiorini, Huynh, and Weltge [8, Lemma 13] give an $O(|V||E|)$ -size formulation for odd-parity nonnegative circulations, using a doubled graph and a disjunction over base vertices. Paat, Walsh, and Xu [27, Lemma 6] extend this to general lattice congruences with a bound $O(q^{q+1}|E||V|^{q-1})$. Their proof enumerates tuples of bases for the contributing cycles. Theorem 2 replaces that enumeration by one network. Our cographic application uses their slack-space reduction.

Group-state networks originate in the finite-group relaxations of Gomory [14]. Köppe, Louveaux, Weismantel, and Wolsey [19, Proposition 4] give a group-state path formulation for corner polyhedra. When G has one vertex and all arcs are loops, Theorem 2 specializes to that formulation. The additional issue for a general graph is conservation at every original vertex. Gribanov, Shumilov, Malyshev, and Pardalos [15] relate canonical-form integer programs to standard-form systems with an additional finite-group constraint, providing a general setting for slack-space reductions of this kind. Fiorini and Pashkovich [12] study flow-based extensions. Our shared formulations for binary demands use the connection between recursive computations and extended formulations [21],

Balance regime	Upper bound on $\text{xc}(Q_{d,g})$
Zero balance	$q V E $ (Theorem 2)
Total supply k	$q(V + k2^{k-1}) E $ (Theorem 11)
s sources, total supply k	$spq(V + 1) E + sq^2 \binom{k+s-1}{s}$ (Proposition 25)
One source, p sinks	$pq(V + 1) E + \min\{q^2L, pq + \rho pq^3\}$ (Corollary 30)
t terminals	$2^t q(V + h2^{h-1}) E + 4^t q^2 L_t$ (Theorem 3)
s sources, p sinks	$spq(V + 1) E + O_s(p^{2s-1}q^{2s+1}\rho^s)$ (Theorem 4)

TABLE 1. The main flow bounds. Here $k = \frac{1}{2}\|d\|_1$, and L is the binary-doubling count in Theorem 15; t, h, L_t, ρ are defined above. The one-source row takes the better of a logarithmic-demand bound and a demand-independent bound. The fixed-source bound is also independent of demand magnitudes.

together with disjunctive programming [3]. Hertrich and Loho [16, Theorem 3.1 and Proposition 3.3] develop the corresponding support-function and circuit framework. Section 4.2 gives a formulation for the empty and unbounded residue hulls arising here.

Remark 20 combines the bounded-cycle decomposition of [27] with shared formulations to give a second circulation bound, $q|V||E| + q|V| + O(q^2 \log q)$. The direct network of Theorem 2 removes the additional union and convolution terms and gives an integral extension with equality on integer points.

There is also a close connection with homology-constrained circulations. The formulations of Conforti, Fiorini, Huynh, Joret, and Weltge [7, Section 12.3] use covering graphs and sums of closed-walk polyhedra. Morell, Seidel, and Weltge [22] study integer circulations in a prescribed homology class. Their integer homology coordinates impose additional equalities alongside the finite-group constraints.

Algorithmic work on modular integer programs provides broader context. Artmann, Weismantel, and Zenklusen [2] solve bimodular integer programs in strongly polynomial time. Fiorini, Joret, Weltge, and Yuditsky [13] give strongly polynomial algorithms for totally Δ -modular systems with at most two nonzero entries per row or per column, for fixed Δ . Nägele, Santiago, and Zenklusen [23], and Nägele, Nöbel, Santiago, and Zenklusen [24], develop congruency- and group-constrained totally unimodular optimization, including randomized feasibility results beyond the bimodular case. Their circulation subproblems may have capacities. Our path partitions and shared formulations use the absence of capacities essentially. Aprile et al. [1] study nearly totally unimodular systems with bounded square subdeterminants and a different cographic matrix convention: in their transpose-network setting, the constraint rows represent a graphic matroid. This differs from the cographic row-matroid hypothesis here.

The transportation reduction in Section 6 also connects to block-structured integer programming. Each sink contributes a block of variables with one column-sum equation; the source sums and group constraint couple the blocks. This is the local-global structure studied in n -fold integer programming [10]. Cslovjcek et al. [9] decompose large right-hand sides into bounded vectors and encode repeated pieces by their multiplicities, obtaining algorithms through mixed-integer formulations with few integer variables. Lassota and Ligthart [20] extend this approach to affine decompositions for 4-block systems. Jansen et al. [17] use doubling and bounded residual problems in algorithms for combinatorial n -fold systems. In our transportation hull, multiples of the group

exponent have label zero and admit an ordinary integral transportation formulation. Enumerating only the residual source sums gives the explicit linear-extension bound of Theorem 4, polynomial in p, q for fixed s and independent of the demand magnitudes.

Eisenbrand and Rothvoss [11] obtain a different form of independence from the right-hand side. Their integer-hull descriptions have fixed left-hand sides and right-hand sides that depend affinely on b within prescribed lattice residue classes. Their bounds depend on the ambient dimension and the coefficient bound. Here the ambient dimension grows with the graph and the number of sinks, while the exponent in Theorem 4 depends only on the number of sources.

Finally, the lower bounds of Cevallos, Weltge, and Zenklusen [6] and Jia, Svensson, and Yuan [18] rule out blanket conclusions from bounded determinants or fixed group order. Jia et al. supply the obstructions for total supply and terminal count used here. Nägele, Nöbel, and Zenklusen [25] subsequently establish large facet coefficients for odd-red matching hulls and refute a proposed original-space description.

1.1. Notation and organization. Groups are written additively with identity 0; integer multiples of group elements mean repeated addition. The incidence column of (u, v) is $\mathbf{e}_v - \mathbf{e}_u$, so negative balances are supplies. Loops, parallel arcs, and disconnected graphs are allowed. A directed cycle may consist of a loop. For a positive integer a , $\text{bits}(a)$ denotes the number of ones in its binary expansion. We write χ^W for the arc-count vector of a directed walk W , and

$$C = \{y \in \mathbb{R}_{\geq 0}^E : Dy = 0\}$$

for the ordinary circulation cone. A polyhedron is *integral* if it is the convex hull of its integer points. Empty sets are permitted and can be represented by an inconsistent equation. Standard facts on integral flows and totally unimodular matrices are used as in [28]; matroid conventions follow [26].

Group elements are assumed to admit enumeration and arithmetic in time polynomial in q and their encoding length, for example through invariant factors. The construction bounds use the numerical group order q . Section 2 proves the circulation theorem. Sections 3 and 4 treat prescribed balances and shared formulations for binary demands. Section 5 proves the terminal-count bound. Section 6 develops the transportation counting network, its two-source unit-demand specialization, and the demand-independent bounds. Section 7 gives the cographic consequences, and Section 8 concludes with the scope of the results and further questions.

2. THE NETWORK EXTENSION

2.1. Construction. The auxiliary graph records two pieces of information during a closed walk: its chosen base vertex and its accumulated label. It remembers the base until the walk returns there; only then may the next walk choose a different base. The common vertices below represent these moments of completion.

For each $r \in V$, start with a copy of the group expansion of G , whose vertices are (r, u, a) for $u \in V$ and $a \in \Gamma$. Identify all vertices (r, r, a) , as r varies, with one common vertex h_a . Thus define

$$\nu_r(u, a) = \begin{cases} h_a, & u = r, \\ (r, u, a), & u \neq r. \end{cases}$$

For every triple $(r, e, a) \in V \times E \times \Gamma$, where $e = (u, w)$, place an arc

$$(9) \quad \nu_r(u, a) \longrightarrow \nu_r(w, a + \gamma_e).$$

Arcs with identical endpoints remain distinct and retain their triple indices. Denote the resulting multigraph by K , and its incidence matrix by B . Define

$$(10) \quad (\pi z)_e = \sum_{r \in V} \sum_{a \in \Gamma} z_{r,e,a}.$$

There are q common vertices and $q|V|(|V| - 1)$ other vertices. Thus K has $q(1 + |V|(|V| - 1))$ vertices. There is one arc for each triple, giving the $q|V||E|$ inequalities in Theorem 2.

Within the copy based at r , a walk that leaves a common vertex can next reach a common vertex only by returning to r in the original graph. The original vertex may then change, because the next excursion may use a different base copy. The group coordinate does not change at this switch.

2.2. Exactness on integer points.

Lemma 5. *If $z \in \mathbb{Z}_{\geq 0}^{E(K)}$ and $Bz = \mathbf{e}_{h_g} - \mathbf{e}_{h_0}$, then $\pi z \in S_g$.*

Proof. Fix $r \in V$ and set $y_e^r = \sum_{a \in \Gamma} z_{r,e,a}$. For each $u \neq r$, sum the balances at the vertices (r, u, a) over all $a \in \Gamma$. The right-hand sides are zero, so $(Dy^r)_u = 0$. Since the entries of Dy^r sum to zero, $(Dy^r)_r = 0$ as well. Summing over r gives $D\pi z = 0$.

Define a group-valued potential on K by

$$\phi(h_a) = a, \quad \phi(r, u, a) = a.$$

The potential difference across the arc (r, e, a) equals γ_e . Since z is integral, summation in Γ is legitimate, and the incidence equations give

$$\begin{aligned} \sum_{e \in E} (\pi z)_e \gamma_e &= \sum_{\alpha \in E(K)} z_\alpha (\phi(\text{head } \alpha) - \phi(\text{tail } \alpha)) \\ &= \sum_{v \in V(K)} (Bz)_v \phi(v) = g. \end{aligned}$$

Nonnegativity and integrality are preserved by (10). $\square$

Lemma 6. *Every $y \in S_g$ is the projection of a nonnegative integer vector z with $Bz = \mathbf{e}_{h_g} - \mathbf{e}_{h_0}$.*

Proof. Decompose y into directed cycles, with repetitions:

$$y = \chi^{C_1} + \cdots + \chi^{C_k}.$$

Such a decomposition exists by repeatedly following positive-flow arcs until a vertex repeats, subtracting the resulting directed cycle, and continuing. Choose a vertex r_i on each C_i . Starting at h_0 , traverse C_1 in the copy based at r_1 . This ends at the common vertex whose group coordinate is the label sum of C_1 . Traverse C_2 in the copy based at r_2 , starting at that common vertex, and continue. The concatenation ends at h_g . Its arc-count vector is a nonnegative integer z satisfying the desired balance equations and projecting to y . If $g = 0$, the concatenation is closed. For $y = 0$, take $z = 0$. $\square$

Proof of Theorem 2. Lemmas 5 and 6 prove (5). The matrix B is totally unimodular, and its right-hand side is integral. Hence

$$(11) \quad F_g = \text{conv}(F_g \cap \mathbb{Z}^{E(K)}).$$

This assertion applies to the unbounded polyhedron as well: its vertices are integral and its recession cone is generated by integer directed-cycle vectors. A nonnegative real combination of these ray generators can be included in a convex combination of integer points by rounding each of its finitely many coefficients up and down. If F_g is empty, (11) is immediate. Therefore

$$\pi(F_g) = \text{conv } \pi(F_g \cap \mathbb{Z}^{E(K)}) = \text{conv } S_g = Q_g.$$

The only inequalities are the $q|V||E|$ nonnegativity constraints. $\square$

Example 7 (Disconnected contributions). Take two vertices r, s , with one loop e_r at r and one loop e_s at s . Let $\Gamma = (\mathbb{Z}/2\mathbb{Z})^2$, label the loops by $(1, 0)$ and $(0, 1)$, and set $g = (1, 1)$. Then $(y_{e_r}, y_{e_s}) = (1, 1)$ belongs to S_g . In a single ordinary group expansion, there is no walk from $(u, 0)$ to (u, g) for either base u : each component generates only one of the two coordinate subgroups. Thus

even the union of these two single-base flow polyhedra is empty. In K , the first loop gives an arc from $h_{(0,0)}$ to $h_{(1,0)}$, and the second gives an arc from $h_{(1,0)}$ to $h_{(1,1)}$. Their concatenation gives the required lift.

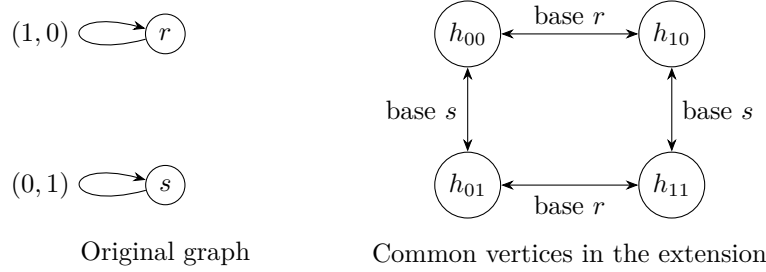


FIGURE 1. The two disconnected loops in Example 7 combine to label $(1,1)$. The path h_{00}, h_{10}, h_{11} uses each loop once. Only the common vertices and arcs between them are drawn; each double arrow denotes two directed arcs.

The example isolates the issue addressed by the construction. Allowing unrestricted switches between original vertices in a single group expansion would lose conservation. Remembering the base during each excursion allows the switch exactly when the original walk closes.

Corollary 8. *Let $R \subseteq V$ meet every directed cycle of G . Theorem 2 holds with only the copies based at vertices of R , giving an integral network extension with $q|R||E|$ inequalities.*

Proof. Keep the common vertices h_a and restrict the arc indices (r, e, a) to $r \in R$. In each retained copy, the proof of Lemma 5 still gives circulation conservation, and the group-potential argument is unchanged. Every directed cycle has a vertex in R , so its lift in Lemma 6 can use a retained copy. Total unimodularity then proves the hull equality as before. If $R = \emptyset$, the graph is acyclic and its only nonnegative circulation is zero; the edgeless auxiliary network represents $Q_0 = \{0\}$ and $Q_g = \emptyset$ for $g \neq 0$. $\square$

A set R with this property is a directed feedback vertex set. The construction is explicit when such a set is supplied; using its minimum cardinality in the size bound does not require an algorithm for finding a minimum set.

2.3. Recession and optimization.

Proposition 9. *One has $Q_0 = C$. For every g such that $S_g \neq \emptyset$, one has $\text{rec}(Q_g) = C$.*

Proof. The cone C is generated by directed-cycle vectors. For each cycle C_i , the vector $q\chi^{C_i}$ belongs to S_0 . Arbitrary nonnegative real combinations of these vectors are convex combinations of nonnegative integer combinations, by independent rounding of their coefficients. Hence $C \subseteq \text{conv } S_0$, and the reverse inclusion follows from the defining equations.

For the second claim, $S_g + S_0 \subseteq S_g$ implies $Q_g + C \subseteq Q_g$. Thus $C \subseteq \text{rec}(Q_g)$. Conversely, $Q_g \subseteq C$, so any recession direction of nonempty Q_g is nonnegative and satisfies the homogeneous circulation equations. This proves the reverse inclusion. $\square$

Corollary 10. *For rational arc costs c , minimizing $c^\top y$ over S_g can be solved deterministically in time polynomial in q , $|V| + |E|$, and the encoding length of the input, assuming the group is given in a form supporting enumeration and arithmetic in polynomial time in q and its encoding length.*

Proof. Give auxiliary arc (r, e, a) cost c_e . For $g \neq 0$, feasibility of F_g is equivalent to reachability from h_0 to h_g . If feasible and K has a negative-cost directed cycle anywhere, the objective is unbounded, since its circulation can be added independently to a feasible flow. Otherwise a shortest

h_0 - h_g path is optimal. For $g = 0$, zero flow is feasible and is optimal unless a negative-cost cycle exists. Reachability, negative-cycle detection, and shortest paths have deterministic polynomial algorithms. Integral optimal auxiliary flows project to elements of S_g by Lemma 5. $\square$

The absence of capacities is essential to this argument: auxiliary cycles can be added without restricting the path choice. Congruency-constrained optimization with capacities requires different techniques; see [23].

3. PRESCRIBED BALANCES

For $d \in \mathbb{Z}^V$ with zero coordinate sum, let $k = \frac{1}{2}\|d\|_1$. The circulation construction extends by keeping track of the k path units required by this balance. As before, set $k2^{k-1} = 0$ when $k = 0$.

Theorem 11. *For every d and g , the polyhedron $Q_{d,g}$ has an integral network extension with at most*

$$(12) \quad q(|V| + k2^{k-1})|E|$$

inequalities. The construction is explicit, and applies also when $Q_{d,g}$ is empty or unbounded.

An integral d -flow decomposes into k simple directed paths from supply vertices to demand vertices, with the prescribed endpoint multiplicities, and a multiset of simple directed cycles. To see this, append a supersource sending $-d_v$ units to each supply vertex and a supersink receiving d_v units from each demand vertex, and decompose the resulting integral flow.

Theorem 11 joins path blocks in a fixed supply order. A subset of demand tokens records which endpoints have been served. After all tokens are served, the circulation construction accounts for any remaining cycles.

3.1. Construction. List supply units in a fixed order as $s_1, \dots, s_k$, repeating a vertex v exactly $-d_v$ times. List demand units as $t_1, \dots, t_k$, repeating v exactly d_v times. Demand units are distinguishable tokens even when their vertices coincide. A supply vertex never equals a demand vertex.

For each $S \subseteq [k]$ and $a \in \Gamma$, create a hub $h_{S,a}$. For each $S \subsetneq [k]$ and $j \notin S$, let $i = |S| + 1$ and create a copy of G with arcs entering s_i or leaving t_j deleted. Expand this copy by Γ : an original arc $e = (u, v)$ yields arcs $(u, a) \rightarrow (v, a + \gamma_e)$ for all $a \in \Gamma$. Identify its vertex (s_i, a) with $h_{S,a}$ and its vertex (t_j, a) with $h_{S \cup \{j\}, a}$. All other vertices are private to this copy. We call this a path block. Deleting the indicated arcs prevents a block from being traversed backwards between layers. Every simple s_i - t_j path of G remains present.

At the final hubs, attach the circulation module from Section 2. Explicitly, for every $r \in V$ take a group-expanded copy of G and identify (r, a) in that copy with $h_{[k], a}$. Other vertices are private to the copy indexed by r . These copies permit a directed cycle to be traversed starting and ending at hubs, while its original base vertex may change between consecutive cycles.

Denote the resulting graph by K_d , its incidence matrix by B_d , and let π sum all auxiliary arc variables corresponding to each original arc. We claim that

$$(13) \quad Q_{d,g} = \pi\{z \in \mathbb{R}_{\geq 0}^{E(K_d)} : B_d z = \mathbf{e}_{h_{[k],g}} - \mathbf{e}_{h_{\emptyset,0}}\}.$$

There are $\sum_{S \subseteq [k]} (k - |S|) = k2^{k-1}$ path blocks, each with at most $q|E|$ arcs, and the circulation module has $q|V||E|$ arcs.

3.2. Proof of exactness.

Proof of Theorem 11. First suppose $k > 0$ and z is an integral feasible auxiliary flow. Decompose it into one directed path from $h_{\emptyset,0}$ to $h_{[k],g}$ and directed cycles. On reaching a hub in layer S , a path can leave through a path block whose source is $s_{|S|+1}$. It can leave that block only at a hub in layer $S \cup \{j\}$; it cannot return to a source hub because all arcs entering that source were deleted. Thus

the auxiliary path completes exactly k blocks, uses each demand token once, and then stays in the final circulation module. Its projection before the final module has balance d .

No auxiliary cycle can move between different subset layers. A cycle inside a path block projects to a circulation. In the final module, an excursion from one hub to another is a closed walk based at the indexing vertex of its copy, and hence projects to a circulation. Cycles entirely avoiding hubs also project to circulations. This proves $D\pi z = d$.

Each auxiliary node has an unambiguous group coordinate $\phi(w)$. Across an auxiliary arc corresponding to e , the change of this coordinate is γ_e . Since z is integral, telescoping in Γ gives

$$\sum_e (\pi z)_e \gamma_e = \sum_{w \in V(K_d)} (B_d z)_w \phi(w) = g.$$

Thus $\pi z \in S_{d,g}$.

Conversely, decompose $y \in S_{d,g}$ into k simple supply–demand paths and directed cycles. Order the paths by the fixed supply list and assign their endpoints to the distinguishable demand tokens. Starting at $h_{\emptyset,0}$, lift the paths through the corresponding successive path blocks, carrying their accumulated group coordinate. At the final layer, lift each remaining cycle through a copy based at one of its vertices. Concatenate these cycle traversals at the final hubs. The resulting walk ends at $h_{[k],g}$, and its arc-count vector projects exactly to y .

For $k = 0$, there are no path blocks. The same argument in the final circulation module proves both inclusions, including the case $g = 0$ where the auxiliary right-hand side vanishes.

Finally, B_d is a node–arc incidence matrix, so the flow polyhedron in (13) is integral by total unimodularity. Its recession cone is generated by integral directed cycles, so the integrality assertion also holds for the unbounded polyhedron. Taking convex hulls proves (13). Counting nonnegativity inequalities gives (12). $\square$

Corollary 12. *Suppose the positive demands are $\ell_1, \dots, \ell_p$, with total $k > 0$, and put*

$$B(\ell) = \sum_{j=1}^p \ell_j \prod_{i \neq j} (\ell_i + 1).$$

Then $Q_{d,g}$ has an integral network extension with at most

$$q(|V| + B(\ell))|E|$$

inequalities. The term $|V|$ may be replaced by $|R|$ for any directed feedback vertex set R .

Proof. Replace a subset of demand tokens by a count vector $x \in \prod_{j=1}^p \{0, \dots, \ell_j\}$. At a state x with $|x| := \sum_j x_j < k$, the next source is $s_{|x|+1}$. For every j with $x_j < \ell_j$, attach the path block to sink t_j , from hubs indexed by x to hubs indexed by $x + \mathbf{e}_j$. For fixed j , there are $\ell_j \prod_{i \neq j} (\ell_i + 1)$ such blocks. Every auxiliary path uses each sink the required number of times; conversely, a path decomposition of any integer d -flow determines these successive counts. The conservation, group-label, and integrality arguments of Theorem 11 therefore apply. The final assertion uses Corollary 8 for the circulation module. $\square$

For one sink, $B(\ell) = k$; for k unit-demand sinks, $B(\ell) = k2^{k-1}$. More generally $B(\ell) \leq k2^{k-1}$, since $\ell_i + 1 \leq 2^{\ell_i}$ for positive integers ℓ_i . Reversing arcs gives the analogous bound using supply multiplicities.

Lemma 13. *Every nonempty $Q_{d,g}$ has recession cone C .*

Proof. By Proposition 9, $Q_{0,0} = C$. Since $S_{d,g} + S_{0,0} \subseteq S_{d,g}$, we have $Q_{d,g} + C = Q_{d,g}$. The reverse containment for recession directions follows from $Q_{d,g} \subseteq \{y \geq 0 : Dy = d\}$. $\square$

3.3. Dependence on total supply.

Corollary 14 (Consequence of Jia–Svensson–Yuan [18]). *Even for $q = 2$, there is no universal bound of the form $2^{o(k)}(|V| + |E|)^{O(1)}$ on $\text{xc}(Q_{d,g})$.*

Proof. Use a complete bipartite graph with k vertices on each side and two parallel arcs, one red and one blue, for every left–right pair. Orient all arcs from left to right. Assign balance -1 to every left vertex and 1 to every right vertex. Nonnegative integral flows are exactly incidence vectors of perfect matchings. Label red arcs 1 , blue arcs 0 in $\mathbb{Z}/2\mathbb{Z}$, and set $g = 1$. Thus $Q_{d,1}$ is exactly the odd-red perfect matching polytope on the doubled complete bipartite graph G_k used in [18]. Their definition gives one edge of each color for every left–right pair, so the identification is direct. Theorem 1.2 of [18] gives extension complexity $2^{\Omega(k)}$. This graph has $2k$ vertices and $2k^2$ arcs, establishing the claim. $\square$

Theorem 11 and Corollary 14 determine the worst-case dependence on total supply up to constants in the exponential exponent and polynomial factors.

4. SHARED FORMULATIONS AND SINGLE-SOURCE BINARY DEMANDS

Theorem 15. *Suppose d has exactly one negative coordinate, at s , and has positive coordinates $\ell_j = d_{t_j}$ at p distinct demand vertices $t_1, \dots, t_p$. Put*

$$L = \sum_{j=1}^p (\lfloor \log_2 \ell_j \rfloor + \text{bits}(\ell_j) - 1) + (p - 1).$$

Then, for every $g \in \Gamma$,

$$(14) \quad \text{xc}(Q_{d,g}) \leq pq(|V| + 1)|E| + q^2 L.$$

In particular this is polynomial in q , graph size, and binary demand length. The same conclusion holds when d has exactly one positive coordinate, by reversing all arcs.

For a single source, the path endpoints specify how many one-unit flow problems must be combined. The proof has two ingredients: group convolution expresses that combination, and a shared formulation reuses the factors created by binary doubling. The common recession cone permits this sharing even when some residue classes are empty.

4.1. Decomposition by demand and residue. Fix distinct vertices s, t and define, for $r \geq 1$ and $a \in \Gamma$,

$$R_{r,a}^{s,t} = Q_{r(\mathbf{e}_t - \mathbf{e}_s), a}.$$

Lemma 16. *For positive integers r, s' and every $g \in \Gamma$,*

$$(15) \quad R_{r+s',g}^{s,t} = \text{conv} \bigcup_{a \in \Gamma} (R_{r,a}^{s,t} + R_{s',g-a}^{s,t}).$$

Proof. Decompose an integral $(r + s')$ -unit flow into $r + s'$ simple s – t paths and cycles. Assign r paths to one part, s' paths to the other, and all cycles to the first part. If the first part has group sum a , the second has group sum $g - a$. Conversely, adding any two such integral flows gives a feasible $(r + s')$ -unit flow. Taking convex hulls and using $\text{conv}(X + Y) = \text{conv } X + \text{conv } Y$ proves the formula. $\square$

For every a , Theorem 11 with $k = 1$ supplies a leaf formulation

$$R_{1,a}^{s,t} = \pi\{z \geq 0 : Bz = f_a\}, \quad f_a = \mathbf{e}_{h_{\{1\},a}} - \mathbf{e}_{h_{\emptyset,0}},$$

using at most $q(|V| + 1)|E|$ inequalities. Its homogenization is $z \geq 0, Bz = \tau f_a$. For $\tau > 0$ its projection is $\tau R_{1,a}^{s,t}$. At $\tau = 0$ the projection is exactly C : every auxiliary circulation projects to

an original circulation, and the final module with zero right-hand side projects onto $Q_{0,0} = C$ by Theorem 2 and Proposition 9. Thus the fiber requirement (16) below holds even when $R_{1,a}^{s,t}$ is empty.

4.2. A formulation that shares subcomputations. We use one formulation for each distinct subproblem, even if it occurs in several branches. A node carries a nonnegative load recording its total use: a union distributes that load among its children, while a sum sends it to both children. The resulting conservation equations allow repeated uses to be aggregated. We apply the dynamic-programming and monotone-circuit construction [16, 21] to polyhedra with a common recession cone.

Example 17. Fix a source–sink pair and let $\Gamma = \mathbb{Z}/2\mathbb{Z}$. Abbreviate the one-unit residue hulls by $P_0 = R_{1,0}^{s,t}$ and $P_1 = R_{1,1}^{s,t}$, and suppose both are nonempty. The two-unit hulls are

$$U_0 = \text{conv}((P_0 + P_0) \cup (P_1 + P_1)), \quad U_1 = P_0 + P_1.$$

For four units with odd label, the two residue assignments give the same sum, so

$$R_{4,1}^{s,t} = U_0 + U_1.$$

An expression tree gives separate copies of P_0 for its two occurrences in U_0 and its occurrence in U_1 , and likewise for P_1 . We can instead keep one homogeneous formulation for each P_i .

Give the two branches of U_0 weights $\lambda_0, \lambda_1 \geq 0$ with $\lambda_0 + \lambda_1 = 1$. Its first branch uses P_0 twice, while its second uses P_1 twice. The sum U_1 uses each once. Thus the total loads on the two leaves are

$$\tau_0 = 2\lambda_0 + 1, \quad \tau_1 = 2\lambda_1 + 1.$$

Impose one homogeneous leaf formulation with projected point $x_i \in \tau_i P_i$ for each i , and output $x = x_0 + x_1$. For fixed λ , this projects onto

$$(2\lambda_0 + 1)P_0 + (2\lambda_1 + 1)P_1 = 2\lambda_0 P_0 + 2\lambda_1 P_1 + P_0 + P_1.$$

Taking all λ gives exactly $U_0 + U_1$. The equality uses convexity: for nonnegative weights, the sum of weighted copies of a nonempty convex set is the same set scaled by the sum of the weights. Separate occurrences may therefore choose different points; their weighted average is represented by the single shared copy.

This example illustrates the two load rules: a convex union splits its load, and a Minkowski sum passes its load to each summand. Repeated use adds loads, with multiplicity. The following lemma applies these rules throughout an acyclic computation and also handles empty leaves and zero loads.

Consider a finite acyclic computation with a distinguished root. Its leaves represent polyhedra $P_\ell \subseteq \mathbb{R}^E$. Each internal node is either a sum node, representing $P_u + P_w$, or a union node, representing $\text{conv} \bigcup_{u \in \mathcal{C}(v)} P_u$. A child may be used by several nodes; a sum may use the same child twice. Empty polyhedra are allowed. Assume every nonempty leaf has recession cone C . For every leaf assume we have a polyhedral cone in variables (z_ℓ, τ) , with a linear projection $x_\ell = \pi_\ell z_\ell$, whose projected fibers are

$$(16) \quad \{x_\ell : (z_\ell, \tau) \text{ is feasible}\} = \begin{cases} \tau P_\ell, & \tau > 0, \\ C, & \tau = 0. \end{cases}$$

The zero fiber is specified also for an empty leaf.

Lemma 18 (Explicit shared formulation). *The root polyhedron has an extended formulation with one copy of each leaf cone and one nonnegativity inequality for every union–child incidence. All remaining constraints are equations.*

Proof. Load equations. Orient computation edges from parents to children. Give every node v a free real variable τ_v . For every union node v and child u , introduce $\lambda_{v,u} \geq 0$ and impose

$$(17) \quad \sum_{u \in \mathcal{C}(v)} \lambda_{v,u} = \tau_v.$$

For each node u , impose the load equation

$$(18) \quad \tau_u = \mathbf{1}_{u=\text{root}} + \sum_{v: v \text{ union parent of } u} \lambda_{v,u} + \sum_{v: v \text{ sum parent of } u} m_{v,u} \tau_v,$$

where $m_{v,u}$ is the multiplicity of u among the two children of v . The root has no parents. Nodes outside its dependency graph may be retained: their loads are zero. Topological induction from parents to children shows that every $\tau_u \geq 0$; no separate inequalities for these variables are needed.

At a leaf, impose its conic formulation with parameter τ_ℓ . The output is $x = \sum_\ell x_\ell$.

Recovering a point at the root. For each leaf of positive load, let $p_\ell = x_\ell / \tau_\ell \in P_\ell$. Inductively from children to parents, define a point at each positive-load node: at a sum node take $p_v = p_u + p_w$; at a union node take $p_v = \sum_{u: \lambda_{v,u} > 0} (\lambda_{v,u} / \tau_v) p_u$. Every child used in these formulas has positive load. Therefore $p_v \in P_v$, and in particular p_{root} exists and belongs to the root polyhedron.

Multiply these identities by τ_v and sum over the internal nodes of positive load. By (18), the coefficient of p_u on the right is τ_u for each nonroot node. The internal-node terms cancel, leaving

$$p_{\text{root}} = \sum_{\ell: \tau_\ell > 0} \tau_\ell p_\ell.$$

Zero loads and empty classes. At zero-load leaves, $x_\ell \in C$. Every nonempty polyhedron represented in the computation has recession cone C . Indeed, sums and convex hulls of finite unions preserve this property when all nonempty operands have that recession cone: write each operand as a bounded polytope plus C and apply the operation to the bounded parts. In particular, the convex hull of their union is a bounded polytope plus the closed cone C , and is therefore closed. Consequently $x = p_{\text{root}} + \sum_{\ell: \tau_\ell = 0} x_\ell$ belongs to the root polyhedron. This argument also proves infeasibility of the formulation when the root polyhedron is empty.

Lifting a root point. Conversely, take a point in the root polyhedron and temporarily unroll the acyclic computation into a finite expression tree. At a sum node, represent the point as a sum of child points; at a union node, represent it as a finite convex combination of points in the children, merging terms that use the same convex child. Propagate the resulting nonnegative weights down the tree, copying the weight to both children at a sum. Lift each positive-weight leaf point into its homogeneous cone. Merge repeated occurrences of the same original leaf by summing their lifts, which is valid because its formulation is a convex cone. Set unused leaf lifts to zero. Aggregate the node weights and union weights in the same way. They satisfy (17)–(18), and their projected leaf sum equals the original point. This proves the reverse inclusion and the claimed size. $\square$

4.3. Sharing residue leaves within a network. For every fixed balance b of total supply k , Theorem 11 constructs a network with incidence matrix B_b , one source u_b , and target vertices $v_{b,a}$ for $a \in \Gamma$, such that

$$(19) \quad Q_{b,a} = \pi_b \{z \geq 0 : B_b z = \mathbf{e}_{v_{b,a}} - \mathbf{e}_{u_b}\}.$$

It has at most $q(|V| + k2^{k-1})|E|$ arcs, with the second term zero for $k = 0$. The network is the same for every target a . Its nonnegative circulation cone projects exactly to C . In the case $b = 0$, $u_b = v_{b,0}$ and $Q_{0,0} = C$. Every nonempty $Q_{b,a}$ has recession cone C .

Lemma 19. *Let $\tau_a \geq 0$ for $a \in \Gamma$. The projection of*

$$(20) \quad z \geq 0, \quad B_b z = \sum_{a \in \Gamma} \tau_a (\mathbf{e}_{v_{b,a}} - \mathbf{e}_{u_b})$$

is

$$C + \sum_{a: \tau_a > 0} \tau_a Q_{b,a}.$$

A positive load on an empty class makes the system infeasible; an empty sum denotes $\{0\}$.

Proof. For one inclusion, choose a lift of each positive-load summand in (19), scale it by τ_a , and add the lifts and a lift of the recession vector.

Conversely, decompose a real nonnegative flow satisfying (20) into source–target paths and directed cycles. The total weight of paths ending at $v_{b,a}$ is τ_a , except when $v_{b,a} = u_b$. Collecting and dividing these paths by τ_a gives a feasible unit-flow lift for the corresponding class. Directed cycles project to C . In the exceptional case $b = 0, a = 0$, the relevant class is already C ; no path is required for its load. If all loads are zero, the system is the auxiliary circulation cone and projects to C . $\square$

In Lemma 18, leaves corresponding to the same balance can consequently use one network (20), with their node loads as its right-hand-side coefficients. The output is the sum of these grouped projections. Lemma 19 decomposes each grouped projection into the separate positive-load leaf contributions and a vector in C . Assign any such vector to one of the original leaf fibers; a positive-load fiber absorbs C , and a zero-load fiber is C . Conversely, summing the separate network lifts gives (20). Thus the grouping preserves exactly the formulation of Lemma 18, including empty classes and zero loads. For all q residue leaves of a fixed balance, it replaces q network copies by one.

4.4. Binary doubling and the size bound.

Proof of Theorem 15. For a fixed pair (s, t_j) , keep all q residues as a vector of polyhedra. Equation (15) is a convolution of two such vectors. It uses q^2 binary sum nodes and q union nodes, each with q children. Binary exponentiation constructs the vector for ℓ_j units using $\lfloor \log_2 \ell_j \rfloor + \text{bits}(\ell_j) - 1$ convolutions. At a doubling, reuse the same vector twice; do not copy the computation below it.

An integral flow with balance d decomposes into ℓ_j paths from s to t_j for each j , plus cycles. Partition those paths by their endpoint and assign all cycles to one part. It follows that the vector $(Q_{d,g})_{g \in \Gamma}$ is the group convolution of the p vectors $(R_{\ell_j,a}^{s,t_j})_{a \in \Gamma}$, requiring $p - 1$ further convolutions. This equality includes every residue class, including empty ones.

There are pq residue leaves. Lemma 19 groups them into p networks, each with at most $q(|V| + 1)|E|$ inequalities. The L convolutions contribute $q^2 L$ union–child incidences. Lemma 18, with this grouping, proves (14). Reversing all arcs changes D and d to $-D$ and $-d$ while leaving the variables, labels, and target unchanged, giving the single-sink case. $\square$

The coefficients in the load equations are 0, 1, 2 apart from signs; large demands are represented by repeated doubling, not by that many variables. The number of equations and variables is also polynomial in the same parameters.

For example, a demand of 13 uses the sequence 1, 2, 3, 6, 12, 13: double, add one, double, double and add one. This uses five convolutions, as counted by $\lfloor \log_2 13 \rfloor + \text{bits}(13) - 1 = 5$. Each residue vector is constructed once and may be used by more than one subsequent operation. The bound therefore counts the computation graph, rather than its much larger expansion into a tree.

Remark 20 (Comparison with the direct circulation network). The bounded-cycle argument of [27, Lemma 6] also yields polynomial dependence on group order when combined with sharing. In the ordinary group expansion, let W_a be the convex hull, over bases $r \in V$, of the projections of unit flows from $(r, 0)$ to (r, a) . Each nonempty fiber has recession cone C , and its homogenized zero fiber projects to C : repeat each original cycle q times in the cover and scale by $1/q$. In particular $W_0 = C$. Write $\star$ for group convolution using convex unions and Minkowski sums. For $q \geq 2$,

$$(21) \quad Q_g = (W^{\star(q-1)})_g.$$

Indeed, repeatedly removing a nonempty zero-sum subcollection from a cycle decomposition leaves at most $q - 1$ cycles; such a subcollection exists among any q cycles by repeated partial sums. Pad the remaining factors by $0 \in W_0$, and absorb the removed circulation into their common recession cone. The reverse inclusion follows by adding integer circulations and taking convex hulls.

For each base, the argument of Lemma 19 groups all residue fibers into one $q|E|$ -arc network. The unions over bases cost $q|V|$ additional inequalities. Lemma 18 and binary exponentiation in (21) therefore give

$$\text{xc}(Q_g) \leq q|V||E| + q|V| + q^2(\lfloor \log_2(q-1) \rfloor + \text{bits}(q-1) - 1).$$

For $q = 1$, the hull is C . Theorem 2 avoids these additional union and convolution terms and retains an integral network extension.

5. A BOUNDED NUMBER OF TERMINALS

5.1. Splitting an integer flow into rounded halves. Retain t, h, M, L_t from Theorem 3. For an integer zero-sum vector b , define

$$H(b) = \{c \in \mathbb{Z}^V : \sum_v c_v = 0, \quad \lfloor b_v/2 \rfloor \leq c_v \leq \lceil b_v/2 \rceil \quad (v \in V)\}.$$

Lemma 21. *Every $y \in \mathbb{Z}_{\geq 0}^E$ with $Dy = b$ has a decomposition $y = z + (y - z)$ into nonnegative integer vectors such that $Dz \in H(b)$ and $D(y - z) \in H(b)$.*

Proof. The bounded polyhedron

$$\{z \in \mathbb{R}^E : 0 \leq z \leq y, \quad \lfloor b/2 \rfloor \leq Dz \leq \lceil b/2 \rceil\}$$

contains $y/2$. Its constraint matrix consists of signed copies of rows of D and of the identity, and is totally unimodular. Its right-hand side is integral, so it contains an integer point z . Put $c = Dz$. Incidence balances sum to zero, giving $c \in H(b)$. Also $b - c \in H(b)$, and $y - z$ is nonnegative and integral. $\square$

The argument uses standard TU rounding [28].

Corollary 22. *For every b and g ,*

$$(22) \quad Q_{b,g} = \text{conv} \bigcup_{c \in H(b)} \bigcup_{a \in \Gamma} (Q_{c,a} + Q_{b-c,g-a}).$$

Proof. For an integer point on the left, apply Lemma 21 and take a to be the group label of z . Conversely, any sum of integer points in the two indicated classes has balance b and label g . Taking convex hulls, using $\text{conv}(X + Y) = \text{conv} X + \text{conv} Y$, proves the equality. Empty classes cause no exception. $\square$

An arbitrary choice of $c \in H(b)$ need not work for a given flow. For example, with only the arcs $s_1 t_1, s_2 t_2$ and balances $(-1, -1, 1, 1)$, the flow using both arcs cannot split off a unit from s_1 to t_2 . The recurrence retains all rounded choices.

5.2. Counting distinct balance states. At depth j of repeated splitting, every balance coordinate belongs to

$$(23) \quad \left\{ \left\lfloor \frac{d_v}{2^j} \right\rfloor, \left\lceil \frac{d_v}{2^j} \right\rceil \right\}.$$

To verify this, for a real number x , taking the floor or ceiling of either $\lfloor x \rfloor / 2$ or $\lceil x \rceil / 2$ gives either $\lfloor x/2 \rfloor$ or $\lceil x/2 \rceil$. Apply this observation coordinatewise by induction. A zero coordinate stays zero, and a nonzero coordinate can only retain its sign or become zero.

Consequently each depth contains at most 2^t balance vectors. After L_t depths every coordinate is in $\{-1, 0, 1\}$. Such a leaf balance has at most h supply units. Leaves may be terminated earlier

when they already have this form. Across all depths there are still at most 2^t distinct leaf balances: each is zero or the original sign at every terminal.

For a nonleaf balance b , equation (22) has at most $2^t q$ sum terms for each target, and there are q targets. Thus at most $4^t q^2$ union–child incidences are needed per depth. Repeated occurrences of the same balance and residue are represented once. This identification preserves acyclicity: at a nonleaf state, $\|b\|_\infty \geq 2$, and each child balance has maximum absolute coordinate at most $\lceil \|b\|_\infty / 2 \rceil < \|b\|_\infty$. There are at most $L_t 4^t q^2$ union–child incidences overall.

The size bound counts this acyclic graph of distinct states. Expanding the recursion into a tree would duplicate the formulations below repeated states.

Proof of Theorem 3. Apply Lemmas 18 and 19 to recurrence (22). At most 2^t leaf balances occur. Each has total supply at most h , so their joint networks require at most $2^t q(|V| + h2^{h-1})|E|$ nonnegativity inequalities. The nonnegative union variables number at most $L_t 4^t q^2$. All other constraints are equations. This proves (6). The graph sizes and the computation size also bound the number of variables and equations. Group arithmetic and enumeration are assumed polynomial in q and their encoding length. The balance coordinates are obtained by binary division of input integers, so their computation has polynomial bit complexity. $\square$

The exponential terminal dependence is necessary in the worst case: Corollary 14 uses $t = 2k$ and $M = 1$. It excludes a bound $2^{o(t)}$ times a polynomial in graph size and binary balance length, already for $q = 2$.

6. A FIXED NUMBER OF SOURCES

We now allow arbitrarily many sinks with arbitrary integer demands. First, replace paths by their source, sink, and group label, obtaining a labeled transportation array. A counting network for these arrays already gives a bound polynomial in total supply for a fixed number of sources. To remove the dependence on total supply, extract multiples of the group exponent so that every residual column sum is bounded. Enumerate the residual source sums and apply the same counting network to each residual hull. The extracted part has label zero and is represented by ordinary transportation constraints.

6.1. Reduction to labeled transportation. Suppose the total supply is positive. Write the sources as s_i with supplies a_i , $1 \leq i \leq s$, and the sinks as t_j with demands b_j , $1 \leq j \leq p$. Let

$$I = \{(i, j, \alpha) : Q_{\mathbf{e}_{t_j} - \mathbf{e}_{s_i}, \alpha} \neq \emptyset\}.$$

For each available triple let $x_{ij\alpha}$ count one-unit flows of that type. Define T_g as the convex hull of the nonnegative integer arrays on I satisfying

$$\sum_{j, \alpha: (i, j, \alpha) \in I} x_{ij\alpha} = a_i, \quad \sum_{i, \alpha: (i, j, \alpha) \in I} x_{ij\alpha} = b_j, \quad \sum_{(i, j, \alpha) \in I} \alpha x_{ij\alpha} = g \quad \text{in } \Gamma.$$

Proposition 23. *With this notation,*

$$\text{xc}(Q_{d,g}) \leq \text{xc}(T_g) + spq(|V| + 1)|E|.$$

Proof. Take an extended formulation of T_g . For each source–sink pair attach the joint network of Lemma 19, using $x_{ij\alpha}$ as its residue loads and setting unavailable loads to zero. Sum all projected network flows for the output. There are at most sp networks, each with at most $q(|V| + 1)|E|$ inequalities.

To prove exactness, first decompose an integer d -flow into unit supply–demand paths and cycles. Assign all cycles to one path, which exists since total supply is positive. The resulting one-unit flows give an integer array in T_g and feasible network lifts. Convexification proves one inclusion.

Conversely, write a feasible real array as $x = \sum_r \theta_r x^r$, where x^r are integer arrays in the defining set of T_g and $\theta_r > 0$ sum to one. For each positive coordinate $x_{ij\alpha}$, Lemma 19 writes its contribution as $x_{ij\alpha} u_{ij\alpha}$ for some $u_{ij\alpha} \in Q_{\mathbf{e}_{t_j} - \mathbf{e}_{s_i}, \alpha}$, plus an overall vector in C . If $x_{ij\alpha} = 0$, all $x_{ij\alpha}^r = 0$ as well.

For each r , the sum $\sum_{ij\alpha} x_{ij\alpha}^r u_{ij\alpha}$ belongs to $Q_{d,g}$: expand each $u_{ij\alpha}$ as a convex combination of integer one-unit flows and use the integer multiplicities.

Averaging these sums with weights θ_r gives the projected output apart from its vector in C , which is absorbed by Lemma 13. This also proves that an empty target hull has no feasible lift. $\square$

6.2. A counting network for labeled transportation. For the remainder of this section, let $I \subseteq [s] \times [p] \times \Gamma$ be any set of available triples, and let $T_g(a, b)$ denote the integer hull defined by the transportation equations above, with nonnegative integer margins $a \in \mathbb{Z}_{\geq 0}^s$ and $b \in \mathbb{Z}_{\geq 0}^p$ of equal total k . The following network counts the units assigned to each source and remembers their accumulated group label.

Lemma 24. *For every such pair of margins,*

$$(24) \quad \text{xc}(T_g(a, b)) \leq sq^2 \binom{k+s-1}{s}.$$

At $k = 0$, the right side is interpreted as zero.

Proof. List k demand tokens in a fixed order, assigning b_j of them to column j . At level ℓ , create states (v, η) for every $v \in \mathbb{Z}_{\geq 0}^s$ with $\sum_i v_i = \ell$ and every $\eta \in \Gamma$. The vector v records the numbers of tokens served by the sources, and η records the accumulated label. If token $\ell + 1$ belongs to column j , each available triple (i, j, α) gives an arc

$$(v, \eta) \longrightarrow (v + \mathbf{e}_i, \eta + \alpha).$$

Project this arc to the unit vector of coordinate (i, j, α) . A path from $(0, 0)$ at level zero to (a, g) at level k specifies an integer array with the required margins and label. Conversely, every such array assigns a source and a label to each token and therefore gives a path. The unit-flow polytope in this acyclic graph is integral, so its projection is $T_g(a, b)$.

There are $q \binom{\ell+s-1}{s-1}$ states at level ℓ , each with at most sq outgoing arcs. Summing over $0 \leq \ell < k$ gives

$$sq^2 \sum_{\ell=0}^{k-1} \binom{\ell+s-1}{s-1} = sq^2 \binom{k+s-1}{s}$$

arcs. When $k = 0$, the hull is $\{0\}$ if $g = 0$, and is empty otherwise; equations suffice in either case. $\square$

Proposition 25 (Dependence on numerical total supply). *If d has $s \geq 1$ supply vertices, $p \geq 1$ demand vertices, and total supply k , then*

$$(25) \quad \text{xc}(Q_{d,g}) \leq spq(|V| + 1)|E| + sq^2 \binom{k+s-1}{s}.$$

Proof. Apply Proposition 23 and Lemma 24 to the original margins, without extracting multiples of the group exponent. $\square$

For fixed s , this bound is polynomial in numerical total supply, but need not be polynomial in its binary encoding length. It complements the integral network of Theorem 11 and the demand-independent bound of Theorem 4. Reversing arcs gives a corresponding bound for a fixed number of sinks.

Corollary 26 (Two sources and unit demands). *If d has two supply vertices and p unit-demand vertices, then*

$$(26) \quad \text{xc}(Q_{d,g}) \leq 2pq(|V| + 1)|E| + q^2 p(p + 1).$$

Proof. Use $s = 2$ and $k = p$ in (25). $\square$

6.3. Removing multiples of the group exponent. Now suppose all column sums are positive, and write $N = |I|$ and

$$N_j = |\{(i, j, \alpha) \in I\}|.$$

If some $N_j = 0$, then $T_g(a, b)$ is empty. Hence assume $N_j \geq 1$ for every j . Let ρ be the exponent of Γ and put

$$(27) \quad c_j = \max\{0, \lfloor b_j/\rho \rfloor - N_j + 1\}, \quad \beta_j = b_j - \rho c_j, \quad K = \sum_{j=1}^p \beta_j.$$

The choice of c_j must work for every feasible array, before its entries are known. Dividing each of the N_j entries by ρ loses less than one whole unit per entry. Thus c_j is a guaranteed number of groups of ρ that can be removed from column j . The remainder need not be smaller than ρ entrywise; it is the column sum that is uniformly bounded. Specifically,

$$(28) \quad 0 \leq \beta_j \leq \rho N_j - 1, \quad K \leq \rho N - p < \rho spq.$$

Define the set of residual row sums by

$$\mathcal{A} = \{u \in \mathbb{Z}_{\geq 0}^s : u \leq a, \sum_i u_i = K, u_i \equiv a_i \pmod{\rho} \ (i \in [s])\}.$$

For $u \in \mathcal{A}$, let P_u be the ordinary transportation polytope on I with row sums $(a - u)/\rho$ and column sums c , and let $R_{u,g} = T_g(u, \beta)$, permitting zero column sums in this notation.

Lemma 27. *The following identity holds:*

$$(29) \quad T_g(a, b) = \text{conv} \bigcup_{u \in \mathcal{A}} (\rho P_u + R_{u,g}).$$

Moreover,

$$(30) \quad |\mathcal{A}| \leq \binom{\lfloor K/\rho \rfloor + s - 1}{s - 1}.$$

Proof. Take an integer array x defining $T_g(a, b)$. In column j , division of each entry by ρ gives

$$\sum_{i, \alpha: (i, j, \alpha) \in I} \left\lfloor \frac{x_{ij\alpha}}{\rho} \right\rfloor \geq \left\lfloor \frac{b_j}{\rho} \right\rfloor - N_j + 1.$$

Indeed, the sum of the N_j remainders is less than ρN_j . The left side is also nonnegative, so it is at least c_j . Choose nonnegative integers $z_{ij\alpha} \leq \lfloor x_{ij\alpha}/\rho \rfloor$ summing to c_j in each column. Set $w = x - \rho z$. Then $w \geq 0$ is integral and has column sums β . Its row sums u belong to $\mathcal{A}$, and $z \in P_u$. Since $\rho\alpha = 0$ for every $\alpha \in \Gamma$, the label of w is g . Thus $x \in \rho P_u + R_{u,g}$.

Conversely, P_u is integral: after negating its row-sum equations, its constraint matrix is the incidence matrix of a bipartite multigraph. For every integer $z \in P_u$ and every integer w defining $R_{u,g}$, the array $\rho z + w$ has margins (a, b) and label g . Taking convex combinations proves $\rho P_u + R_{u,g} \subseteq T_g(a, b)$, and hence (29).

For the count, write $a_i \equiv \sigma_i \pmod{\rho}$ with $0 \leq \sigma_i < \rho$. Every $u \in \mathcal{A}$ has the form $u_i = \sigma_i + \rho v_i$, where $v_i \geq 0$ is integral and

$$\sum_i v_i = (K - \sum_i \sigma_i)/\rho \leq \lfloor K/\rho \rfloor.$$

When $\mathcal{A}$ is nonempty, the sum on the left has a fixed nonnegative integer value. Counting its weak compositions into s parts, and dropping the upper bounds $u \leq a$, proves (30). $\square$

The column sums c_j can be large, but they occur only as right-hand sides of ordinary transportation equations. All group-constrained work is confined to the residual supply K .

Example 28 (A bounded residual with growing demands). Take two sources, three sinks, and $\Gamma = \mathbb{Z}/3\mathbb{Z}$, so $\rho = 3$. Suppose each source–sink pair has just one available label, given by

$$(\alpha_{ij}) = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}.$$

Such labels arise from a bipartite directed graph with one arc from each source to each sink. We suppress the label index and write an array as a 2×3 matrix. Consider

$$x = \begin{pmatrix} 10 & 8 & 7 \\ 7 & 6 & 4 \end{pmatrix}, \quad a = (25, 17), \quad b = (17, 14, 11).$$

Its label is $8 + 2 \cdot 7 + 7 + 4 = 33 = 0$ in Γ . Here $N_j = 2$ for every column, so (27) gives

$$c = (4, 3, 2), \quad \beta = (5, 5, 5), \quad K = 15.$$

The residual column bound $\beta_j \leq \rho N_j - 1$ is attained. One admissible extraction is

$$z = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}, \quad w = x - 3z = \begin{pmatrix} 4 & 2 & 4 \\ 1 & 3 & 1 \end{pmatrix}.$$

The extracted array z has column sums c and row sums $(5, 4)$. The residual w has column sums β , row sums $u = (10, 5)$, and the same label 0 as x . Thus $z \in P_{(10,5)}$ and $w \in R_{(10,5),0}$.

The residual row sums are the only margins that must be enumerated. They satisfy $u_1 + u_2 = 15$, $u_1 \equiv 1 \pmod{3}$, and $u_2 \equiv 2 \pmod{3}$, giving just

$$\mathcal{A} = \{(1, 14), (4, 11), (7, 8), (10, 5), (13, 2)\}.$$

For each choice, an ordinary transportation polytope handles the extracted part, and a network handles the 15 residual units.

To see why large demands cause no further growth, replace x by $x + 3H$ for any nonnegative integer 2×3 array H . The new extracted column sums are c plus the column sums of H . Taking $z + H$ as the extraction leaves exactly the same w , β , and K . The source residues are unchanged, and the larger source sums leave the same five possibilities in $\mathcal{A}$. The demand magnitudes enter the ordinary transportation equations, while the residual construction stays the same size.

6.4. Combining the extracted and residual parts. For fixed u , apply the counting network of Lemma 24 to $R_{u,g} = T_g(u, \beta)$. It uses at most $sq^2 \binom{K+s-1}{s}$ arcs. The original demand magnitudes no longer occur in its state space.

Proposition 29. *With K as in (27),*

$$(31) \quad \text{xc}(T_g(a, b)) \leq \binom{\lfloor K/\rho \rfloor + s - 1}{s - 1} \left(N + sq^2 \binom{K + s - 1}{s} + 1 \right).$$

Proof. For each $u \in \mathcal{A}$, the polytope P_u has N nonnegativity inequalities, and Lemma 24 supplies the residual formulation. Taking their Cartesian product and projecting by $(z, w) \mapsto \rho z + w$ gives a formulation for $\rho P_u + R_{u,g}$ with the sum of these sizes. Apply the convex-hull-of-union formulation [3] to (29), using one additional nonnegative mixing variable for each u .

At zero mixing weight, the homogeneous column-sum equations force all transportation variables to zero, and acyclicity forces all residual-network flows to zero. At positive mixing weight, division by that weight recovers the corresponding product formulation. Thus infeasible terms can carry only zero weight, and the projection is exactly the union hull in (29). Counting terms by (30) proves (31). $\square$

Proof of Theorem 4. Use the available triples from Proposition 23. They satisfy $N \leq spq$ and can be found by reachability in the one-unit network of Theorem 11. If a column has no available triple,

the target hull is empty. Otherwise combine Propositions 23 and 29 to obtain

$$(32) \quad \text{xc}(Q_{d,g}) \leq spq(|V| + 1)|E| + \binom{\lfloor K/\rho \rfloor + s - 1}{s - 1} \left(spq + sq^2 \binom{K + s - 1}{s} + 1 \right),$$

where $K < \rho spq$. For fixed s , the second term is $O_s(p^{2s-1}q^{2s+1}\rho^s)$.

For $s = 2$, the transportation contribution is at most

$$(\lfloor K/\rho \rfloor + 1)(2pq + q^2K(K + 1) + 1), \quad K < 2pq\rho,$$

which is $O(p^3q^5\rho^2)$. All network sizes and enumerations obey the same polynomial bounds for fixed s . Large margins enter only the transportation right-hand sides and the elementary divisions in (27), whose binary encoding lengths are polynomial in the input length. Reversing arcs gives the assertion for a fixed number of sinks. $\square$

Corollary 30. *Under the hypotheses of Theorem 15, let L be its binary-doubling count and let K be the residual total from (27), for the available triples in Proposition 23. If every sink has an available triple, then*

$$(33) \quad \text{xc}(Q_{d,g}) \leq pq(|V| + 1)|E| + \min\{q^2L, pq + q^2K\}, \quad K < \rho pq.$$

If a sink has no available triple, the hull is empty. The same bounds hold with supply and demand reversed.

Proof. With one source, the set $\mathcal{A}$ in (29) consists of the single value $u = K$: it satisfies $K \leq a_1$ and $K \equiv a_1 \pmod{\rho}$. There is no union to formulate. The extracted transportation polytope uses $N \leq pq$ inequalities and the residual network uses q^2K . Proposition 23 therefore gives the second term in the minimum; Theorem 15 gives the first. $\square$

Neither construction uniformly improves the other. Binary doubling gives the smaller displayed bound when $L < K + p/q$, whereas residual extraction removes all demand-magnitude dependence from the inequality count.

7. STRICTLY MODULAR COGRAPHIC INTEGER HULLS

We use the graph and lattice ingredients of the slack-space reduction in [27]. Expressing them through a selected row basis yields the balance Db and allows the same reduction to handle arbitrary integer right-hand sides. The two ingredients enforce different requirements. The graph represents the real linear relations among slacks: $y = b - Ax$ must satisfy $Dy = Db$. The finite quotient group then records the additional condition that the recovered x is integral. We construct the graph first and the group second.

7.1. From a row basis to a directed graph. Throughout this section, A has rank $n \geq 1$, is strictly Δ -modular, and has cographic row matroid $M(A^\top)$. The right-hand side b is an arbitrary integer vector unless stated otherwise. Choose n independent rows and reorder A so that

$$A = \begin{bmatrix} A_B \\ A_N \end{bmatrix}, \quad U = AA_B^{-1} = \begin{bmatrix} I_n \\ T \end{bmatrix}, \quad T = A_N A_B^{-1}, \quad R = [-T \ I_{m-n}].$$

Here $|\det A_B| = \Delta$. Every square minor of T is 0 or ± 1 . Indeed, complete any minor with complementary identity rows to obtain a maximal minor of U . Its determinant, up to sign, is the chosen minor of T , and multiplication by $\det A_B$ gives a maximal minor of A . Consequently T and R are totally unimodular. Also,

$$(34) \quad \ker R = \text{col } A.$$

Since R is an orthogonal representation of full complementary rank, its column matroid is $M(A^\top)^*$, which is graphic.

Lemma 31. *There is a directed multigraph with m arcs and $m - n + 1$ vertices whose incidence matrix D satisfies $\ker D = \ker R$.*

Proof. First suppose $m > n$. Take a graph with edge labels matching the columns of R and with the same cycle matroid. Remove isolated vertices and identify one vertex from each component. This preserves the cycle matroid and makes the graph connected. Since its cycle-matroid rank is $m - n$, it has $m - n + 1$ vertices. Choose any orientation and delete one row of its incidence matrix to obtain a full-row-rank matrix $E = [E_1 \ E_2]$, where the last $m - n$ columns form a basis. Normalizing gives

$$E_2^{-1}E = [H \ I_{m-n}].$$

This matrix is totally unimodular: a minor of H extends with complementary identity columns to a maximal minor, which is 0 or ± 1 because E is TU and $\det E_2 = \pm 1$. The matrices H and $-T$ have the same nonzero pattern: the support of each nonbasis column is determined by its fundamental circuit relative to the common basis. The uniqueness of totally unimodular signings [5, 26] supplies diagonal sign matrices S and C_B such that $H = S(-T)C_B$. It follows that

$$[H \ I] = SR \operatorname{diag}(C_B, S).$$

Reversing the indicated arc orientations absorbs the column signs, and row operations preserve the kernel. Restoring the deleted incidence row also preserves the kernel. This gives the required D .

If $m = n$, R has no rows and has full kernel $\mathbb{R}^m$. Use one vertex with m loops. $\square$

7.2. The lattice congruence. Let

$$(35) \quad \Gamma = \mathbb{Z}^n / A_B \mathbb{Z}^n, \quad g = [b_B], \quad \gamma_i = \begin{cases} [e_i], & 1 \leq i \leq n, \\ 0, & n < i \leq m. \end{cases}$$

The brackets denote a coset, and $|\Gamma| = \Delta$.

Lemma 32. *For arbitrary $b \in \mathbb{Z}^m$, the map $x \mapsto b - Ax$ is a bijection from $P(A, b) \cap \mathbb{Z}^n$ onto*

$$\left\{ y \in \mathbb{Z}_{\geq 0}^m : Dy = Db, \sum_{i=1}^m y_i \gamma_i = g \right\}.$$

Its inverse on this set is $y \mapsto A_B^{-1}(b_B - y_B)$.

Proof. If $x \in P(A, b) \cap \mathbb{Z}^n$, then $y = b - Ax$ is nonnegative and integral. Since $DA = 0$, we have $Dy = Db$. Its first n entries satisfy $[y_B] = [b_B]$ in (35).

Conversely, let y satisfy the displayed conditions. The congruence gives $b_B - y_B \in A_B \mathbb{Z}^n$, so $x = A_B^{-1}(b_B - y_B)$ is integral. The balance equation implies $b - y \in \ker D = \operatorname{col} A$. Agreement of Ax with $b - y$ on the basis rows therefore implies agreement on every row. Thus $Ax = b - y \leq b$. Full column rank gives uniqueness. $\square$

Proof of (1) in Theorem 1. Lemma 32 identifies the integer points by an affine bijection whose inverse extends to their convex hulls. Here $b \in \operatorname{col} A$, so $Db = 0$. Apply Theorem 2 to the graph from Lemma 31, with group (35). Its numbers of arcs and vertices are m and $m - n + 1$, respectively, giving $\Delta m(m - n + 1)$ inequalities.

For constructibility, a graphic representation of a totally unimodular graphic matrix can be found in polynomial time by standard recognition and representation algorithms [4]. Smith normal form computes the quotient in (35). Enumerating its Δ elements and constructing the auxiliary network requires polynomial time in Δ and the input encoding length. All rational matrices in the affine projection have polynomial encoding length. $\square$

If equations are counted as pairs of inequalities, the auxiliary incidence equations add at most $2\Delta(1 + (m - n + 1)(m - n))$ inequalities. This leaves the asymptotic bound $O(\Delta m^2)$ unchanged. Equations defining the affine projection contribute only polynomially many additional constraints.

Corollary 8 sharpens (1) to $\Delta m|R|$ for any directed feedback vertex set R of the graph representing D . In particular, its minimum cardinality can replace $m - n + 1$ in the size bound; an explicit construction uses a supplied set R .

7.3. Removing repeated inequalities. Deleting repeated inequalities gives a row bound in terms of the dimension.

Lemma 33. *A simple cographic matroid of rank $n \geq 1$ has at most $3n$ elements.*

Proof. Represent it as the bond matroid of an undirected multigraph, and identify vertices across components as in Lemma 31. Write e and v for the resulting numbers of edges and vertices. The graph is connected and $n = e - v + 1$. Simplicity of the bond matroid excludes bonds of size one and two, hence excludes bridges and two-edge cuts.

If $v = 1$, all edges are loops and $e = n$. If $v \geq 2$, each vertex has at least three incident nonloop edges: otherwise its cut, or a bond contained in that cut, would have size at most two. Thus $2e \geq 3v$, even when loops are present. Substituting $v = e - n + 1$ yields $e \leq 3n - 3 \leq 3n$. $\square$

Lemma 34. *Under the hypotheses of Theorem 1, deleting zero rows and repeated identical rows leaves at most $6n$ rows and preserves all hypotheses and the polyhedron.*

Proof. If two nonzero rows a_i, a_j are proportional, say $a_j = \lambda a_i$, extend a_i to a row basis and replace it by a_j . Strict Δ -modularity gives $|\lambda| = 1$. Thus any parallel class in $M(A^\top)$ contains at most two distinct rows, negatives of each other.

Write $b = A\bar{x}$. Identical rows have identical right-hand sides, and zero rows have right-hand side zero. Deleting repeated copies and zero rows therefore preserves $P(A, b)$ and the rank. It preserves strict Δ -modularity, the cographic property under deletion, and the condition that the restricted right-hand side lies in the column space.

The simplification of $M(A^\top)$ is a simple cographic matroid of rank n , so it has at most $3n$ elements by Lemma 33. Keeping at most two rows per parallel class gives at most $6n$ rows. $\square$

Completion of the proof of Theorem 1. Apply (1) after Lemma 34, with $m' \leq 6n$. Since $n \geq 1$, one obtains the slightly sharper estimate

$$\text{xc}(P_I(A, b)) \leq \Delta m'(m' - n + 1) \leq \Delta(30n^2 + 6n) \leq 36\Delta n^2.$$

The deletions can be carried out in polynomial time. $\square$

For every fixed Δ , a linear lower-bound baseline is obtained from $A = \text{diag}(\Delta, 1, \dots, 1)$ and $b = 0$. Its row matroid is free and hence cographic, and $P_I(A, 0)$ is the negative orthant, of extension complexity n . For $\Delta = 1$, the linear upper bound noted after Theorem 1 is therefore of the correct order. For fixed $\Delta \geq 2$, the bounds established here leave a gap between this linear baseline and the quadratic upper bound.

7.4. Right-hand sides outside the column space.

Corollary 35. *For any incidence representation D as in Lemma 31, put $d = Db$ and $k = \frac{1}{2}\|d\|_1$. Then*

$$\text{xc}(P_I(A, b)) \leq \Delta(m - n + 1 + k2^{k-1})m.$$

If d has one supply vertex and p demand vertices of demands $\ell_1, \dots, \ell_p$, then

$$\text{xc}(P_I(A, b)) \leq p\Delta(m - n + 2)m + \Delta^2 L,$$

where L is as in Theorem 15. The same bound holds with supply and demand reversed. More generally, define t, h, L_t from $d = Db$ as in Theorem 3. Then

$$\text{xc}(P_I(A, b)) \leq 2^t \Delta(m - n + 1 + h2^{h-1})m + 4^t \Delta^2 L_t.$$

If d has $s \geq 1$ supply vertices and $p \geq 1$ demand vertices, then

$$\text{xc}(P_I(A, b)) \leq sp\Delta(m - n + 2)m + O_s(p^{2s-1}\Delta^{3s+1}).$$

Proof. Lemma 32 gives an affine isomorphism from $P_I(A, b)$ to $Q_{Db, [b_B]}$ for a group of order Δ , on a graph with m arcs and $m - n + 1$ vertices. Apply Theorems 11, 15, 3, and 4, using $\rho \leq \Delta$ in the last bound. $\square$

The balance parameters are computed from the chosen incidence representation of the given system, before the row reduction used for the translated-cone bound. The conclusions hold for every such representation, so their numerical bounds can be minimized over the valid choices of D . Finding a representation that minimizes a bound is a separate algorithmic question. Corollaries 12 and 30 and Proposition 25 transfer through the same slack map, with $q = \Delta$, $|V| = m - n + 1$, and $|E| = m$.

The single-source condition permits bounded polyhedra of positive dimension. For an integer $h \geq 2$, take

$$A = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ h \end{bmatrix}, \quad D = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}.$$

The rows of A represent the cographic rank-one matroid on two parallel elements; A is strictly bimodular and $\ker D = \text{col } A$. The polyhedron $P(A, b) = [-h/2, 1/2]$ is a bounded nondegenerate interval, while $Db = (-h - 1, h + 1)^\top$ has one source and one sink. Thus the right-hand-side extension applies to polyhedra beyond translated cones.

7.5. Lower bounds for the group dependence. The following consequence of the matching lower bound shows that polynomial dependence on the group order is necessary, up to the exponent, even for cographic translated cones. It also rules out replacing group order by group exponent as the sole group parameter in a polynomial bound.

Proposition 36. *For every sufficiently large integer k , there is a one-vertex group-labeled graph with $2k^2$ loops and a group of order 2^{2k} and exponent 2 whose circulation hull has extension complexity $2^{\Theta(k)}$. There is also a square strictly 2^{2k} -modular integer matrix A of order $2k^2$, with cographic row matroid and quotient group $\mathbb{Z}^{2k^2}/A\mathbb{Z}^{2k^2}$ of exponent 2, and an integer $b \in \text{col}(A)$ such that*

$$\text{xc}(P_I(A, b)) = 2^{\Theta(k)}.$$

Proof. Take the doubled complete bipartite graph from Corollary 14, with k vertices in each part. For each pair (i, j) and color $c \in \{0, 1\}$, replace its edge by a loop labeled

$$\gamma_{ijc} = (\mathbf{e}_i, \mathbf{e}_j, c) \in (\mathbb{Z}/2\mathbb{Z})^k \times (\mathbb{Z}/2\mathbb{Z})^k \times \mathbb{Z}/2\mathbb{Z}.$$

Let Γ be the subgroup generated by these labels. The blue labels span the vectors $(u, v, 0)$ with $\sum_i u_i = \sum_j v_j$, a subspace of dimension $2k - 1$; the difference between a red and a blue label supplies the last coordinate. Thus $|\Gamma| = 2^{2k}$ and its exponent is 2. Set $g = (\mathbf{1}, \mathbf{1}, 1) \in \Gamma$ and write Q for the resulting circulation hull.

For every defining integer point of Q , the total multiplicity at each of the original left and right vertices is odd, and the total red multiplicity is odd. In particular $\sum_{i,j,c} y_{ijc} \geq k$. Equality forces every original vertex to have multiplicity exactly one. Hence the face

$$Q \cap \left\{ y : \sum_{i,j,c} y_{ijc} = k \right\}$$

is exactly the odd-red perfect matching polytope from [18, Theorem 1.2]. Its exponential lower bound also holds for Q . Theorem 2 gives the upper bound $2^{2k} \cdot 2k^2 = 2^{O(k)}$.

Put $M = 2k^2$ and let $\Phi : \mathbb{Z}^M \rightarrow \Gamma$ be the label homomorphism. Choose an integer matrix A whose columns form a lattice basis of $\ker \Phi$. Then A is nonsingular and $|\det A| = [\mathbb{Z}^M : \ker \Phi] = 2^{2k}$;

its only maximal minor is its determinant. Its row matroid is free, hence cographic. Choose an integer b with $\Phi(b) = g$, for instance the incidence vector of a perfect matching with exactly one red edge. The affine bijection $x \mapsto b - Ax$ identifies $P_I(A, b)$ with Q . Moreover $\mathbb{Z}^M / A\mathbb{Z}^M \cong \Gamma$, and $b \in \text{col}(A) = \mathbb{R}^M$. This proves the second assertion. $\square$

In this construction, $\Delta = 2^{2k}$ and $n = 2k^2 = \frac{1}{2}(\log_2 \Delta)^2$, while $\text{xc}(P_I(A, b)) = \Delta^{\Theta(1)}$. Hence no uniform bound $\Delta^{o(1)} n^C$ can hold for any fixed C : on this family it would be $2^{o(k)}$, contradicting Proposition 36. The polynomial dependence on Δ in Theorem 1 is therefore necessary, up to the exponent.

Since the quotient groups in this family have exponent 2, no bound $f(\rho)(|V| + |E|)^C$ can hold for all circulation hulls with an absolute constant C ; nor can $f(\rho)(m + n)^C$ hold for all the cographic translated-cone hulls when ρ denotes the exponent of the quotient group. This obstruction lets Δ grow and does not give a superlinear lower bound for fixed Δ .

8. CONCLUDING REMARKS

The circulation construction combines cycles based at unrelated vertices in a single integral network. Through the cographic slack-space representation, it gives an $O(\Delta n^2)$ bound whose exponent in the dimension is independent of Δ . For prescribed balances, two further decompositions control different sources of complexity: rounded balances lead to shared formulations with exponential dependence on terminal count and polynomial dependence on binary balance length, while extraction of multiples of the group exponent leaves bounded transportation demands. The latter yields polynomial-size formulations for a fixed number of sources, with an inequality count independent of the demand magnitudes.

The fixed-source result also admits a transportation interpretation. Consider integer shipments from a fixed number of depots to arbitrarily many customers, with prescribed supplies and demands. Requiring the aggregate shipments using designated options to total a multiple of a batch size is a finite-group constraint: label those options by 1 in the cyclic group of that order and all others by 0, with target 0. In the absence of route capacities, the resulting integer hull has an extended formulation whose inequality count is polynomial in the number of customers and the batch size, independently of shipment volumes.

Adding capacity constraints to an extension does not generally give the capacitated integer hull. On one vertex with one loop of label 1 in $\mathbb{Z}/2\mathbb{Z}$, target 1 gives

$$S_1 = \{1, 3, 5, \dots\}, \quad Q_1 = [1, \infty).$$

Intersecting Q_1 with $[0, 2]$ gives $[1, 2]$, whereas the capacitated integer hull is $\{1\}$. The same issue prevents an arbitrary affine slice of a circulation hull from being identified with the corresponding integer hull. The prescribed-balance constructions address this distinction directly.

The failure persists with unit capacities on two arcs. With two loops of label 1 in $\mathbb{Z}/2\mathbb{Z}$ and target 1, the uncapacitated hull is $\{y \geq 0 : y_1 + y_2 \geq 1\}$. Its intersection with $[0, 1]^2$ contains $(1, 1)$, whereas the capacitated integer hull is $\text{conv}\{(1, 0), (0, 1)\}$. Indeed, the uncapacitated hull contains $(1, 0)$ and $(0, 1)$ and has recession cone $\mathbb{R}_{\geq 0}^2$, which gives the stated description.

The cographic row-matroid hypothesis is substantive. For a graphic row matroid, the orthogonal representation is cographic and need not be representable by a node-arc incidence matrix. The conic bimodular examples discussed in [27, Section 6], based on [6], rule out the corresponding general extension-complexity claim. Arbitrary balances already encounter the matching obstruction of Corollary 14.

Remark 37 (Divisible balances). Let ρ be the exponent of Γ . If $d = \rho b$ with $b \in \mathbb{Z}^V$, then

$$Q_{d,0} = \{y \in \mathbb{R}_{\geq 0}^E : Dy = d\}.$$

The ordinary flow polyhedron $P_b = \{x \geq 0 : Dx = b\}$ is integral. Every ρx with $x \in P_b \cap \mathbb{Z}^E$ has label zero and balance d . Hence $\rho P_b = \text{conv}\{\rho x : x \in P_b \cap \mathbb{Z}^E\} \subseteq Q_{d,0}$. The reverse inclusion follows from the balance and nonnegativity equations. Empty polyhedra are included. This extends the zero-balance identity of Proposition 9.

Several quantitative questions remain. Does a bound $f(s) \text{poly}(p, q, |V|, |E|)$ hold with an absolute polynomial degree? Corollary 14 has $s = p = k$ and $q = 2$, so any such bound requires $f(s) = 2^{\Omega(s)}$; an exponential or larger parameter factor is not excluded. Can the quadratic dimension dependence in Theorem 1 be improved, or can a superlinear lower bound be established for fixed $\Delta \geq 2$? For nonzero targets, which label structures permit a smaller effective group? Proposition 36 shows that the exponent alone cannot control the group dependence. A further direction is to identify additional network constraints that preserve compact formulations; even unit capacities require a new exactness argument.

REFERENCES

- [1] Manuel Aprile, Samuel Fiorini, Gwenaél Joret, Stefan Kober, Michał Seweryn, Stefan Weltge, and Yelena Yuditsky. Integer programs with nearly totally unimodular matrices: The cographic case. *Mathematics of Operations Research*, 2026. Published online 9 June 2026.
- [2] Stephan Artmann, Robert Weismantel, and Rico Zenklusen. A strongly polynomial algorithm for bimodular integer linear programming. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1206–1219. ACM, 2017.
- [3] Egon Balas. Disjunctive programming. *Annals of Discrete Mathematics*, 5:3–51, 1979.
- [4] Robert E. Bixby and William H. Cunningham. Converting linear programs to network problems. *Mathematics of Operations Research*, 5(3):321–357, 1980.
- [5] Paul Camion. Caractérisation des matrices unimodulaires. *Cahiers du Centre d’Études de Recherche Opérationnelle*, 5:181–190, 1963.
- [6] Alfonso Cevallos, Stefan Weltge, and Rico Zenklusen. Lifting linear extension complexity bounds to the mixed-integer setting. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 788–807. SIAM, 2018.
- [7] Michele Conforti, Samuel Fiorini, Tony Huynh, Gwenaél Joret, and Stefan Weltge. The stable set problem in graphs with bounded genus and bounded odd cycle packing number. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms*, pages 2896–2915. SIAM, 2020.
- [8] Michele Conforti, Samuel Fiorini, Tony Huynh, and Stefan Weltge. Extended formulations for stable set polytopes of graphs without two disjoint odd cycles. *Mathematical Programming*, 192:547–566, 2022.
- [9] Jana Cslovjceksek, Martin Koutecký, Alexandra Lassota, Michał Pilipczuk, and Adam Polak. Parameterized algorithms for block-structured integer programs with large entries. *TheoretiCS*, 4:15:1–15:49, 2025.
- [10] Jesús A. De Loera, Raymond Hemmecke, Shmuel Onn, and Robert Weismantel. N-fold integer programming. *Discrete Optimization*, 5(2):231–241, 2008.
- [11] Friedrich Eisenbrand and Thomas Rothvoss. A parameterized linear formulation of the integer hull. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 4560–4571. SIAM, 2026.
- [12] Samuel Fiorini and Kanstantsin Pashkovich. Uncapacitated flow-based extended formulations. *Mathematical Programming*, 153:117–131, 2015.
- [13] Samuel Fiorini, Gwenaél Joret, Stefan Weltge, and Yelena Yuditsky. Integer programs with bounded subdeterminants and two nonzeros per row. *Journal of the ACM*, 72(1):3:1–3:50, 2025.
- [14] Ralph E. Gomory. Some polyhedra related to combinatorial problems. *Linear Algebra and its Applications*, 2(4):451–558, 1969.

- [15] Dmitry Griбанov, Ivan Shumilov, Dmitry Malyshev, and Panos Pardalos. On Δ -modular integer linear problems in the canonical form and equivalent problems. *Journal of Global Optimization*, 88(3):591–651, 2024.
- [16] Christoph Hertrich and Georg Loho. Neural networks and (virtual) extended formulations. arXiv:2411.03006v4, 2026. Version 4, 28 May 2026.
- [17] Klaus Jansen, Kai Kahler, Lis Piroton, and Malte Tutas. New algorithm for combinatorial n -folds and applications. In *20th International Symposium on Parameterized and Exact Computation*, volume 358 of *Leibniz International Proceedings in Informatics*, pages 15:1–15:15. Schloss Dagstuhl, 2025.
- [18] Xinrui Jia, Ola Svensson, and Weiqiang Yuan. The exact bipartite matching polytope has exponential extension complexity. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1635–1654. SIAM, 2023.
- [19] Matthias Köppe, Quentin Louveaux, Robert Weismantel, and Laurence A. Wolsey. Extended formulations for Gomory corner polyhedra. *Discrete Optimization*, 1(2):141–165, 2004.
- [20] Alexandra Lassota and Koen Ligthart. Solving 4-block integer linear programs faster using affine decompositions of the right-hand sides. In *Integer Programming and Combinatorial Optimization*, volume 16588 of *Lecture Notes in Computer Science*, pages 173–187. Springer, 2026.
- [21] R. Kipp Martin, Ronald L. Rardin, and Brian A. Campbell. Polyhedral characterization of discrete dynamic programming. *Operations Research*, 38(1):127–138, 1990.
- [22] Sarah Morell, Ina Seidel, and Stefan Weltge. Minimum-cost integer circulations in given homology classes. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms*, pages 2725–2739. SIAM, 2021.
- [23] Martin Nägele, Richard Santiago, and Rico Zenklusen. Congruency-constrained TU problems beyond the bimodular case. *Mathematics of Operations Research*, 49(3):1303–1348, 2024.
- [24] Martin Nägele, Christian Nöbel, Richard Santiago, and Rico Zenklusen. Advances on strictly Δ -modular IPs. *Mathematical Programming*, 210:731–760, 2025.
- [25] Martin Nägele, Christian Nöbel, and Rico Zenklusen. On the complexity of the odd-red bipartite perfect matching polytope. arXiv:2603.18232, 2026.
- [26] James Oxley. *Matroid Theory*. Oxford University Press, 2nd edition, 2011.
- [27] Joseph Paat, Zach Walsh, and Luze Xu. Extended formulations for the integer hull of strictly Δ -modular cographic polyhedral cones. *Operations Research Letters*, 60:107235, 2025.
- [28] Alexander Schrijver. *Theory of Linear and Integer Programming*. Wiley, 1986.

DEPARTMENT OF INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH, INDIAN INSTITUTE OF TECHNOLOGY BOMBAY, MUMBAI, INDIA 400076

Email address: `abhardwaj@iitb.ac.in`