Scale Invariance and Decomposability of the Binary Plot Allocation Problem: A Rigorous Mathematical Analysis

Quantum Optimization Project

November 2025

Abstract

We establish rigorous mathematical foundations for the scale invariance properties of the binary plot allocation problem used in agricultural land optimization. We prove that under area-normalized objectives and purely combinatorial constraints, the optimal allocation pattern is independent of the total land area. Furthermore, we derive conditions under which the problem decomposes into identical subproblems, enabling significant computational savings through solution replication. These results have important implications for quantum-classical hybrid optimization, where problem size is constrained by quantum hardware limitations.

Contents

1 Introduction and Motivation

In agricultural land allocation, we seek to optimally assign crops to land parcels while satisfying various constraints on diversity, capacity, and resource usage. When the land is discretized into equal-sized plots, the problem becomes a binary optimization problem amenable to quantum annealing and other combinatorial optimization techniques.

A fundamental question arises: *If we solve the allocation problem for a representative subset of land (e.g., 100 hectares with 25 plots), can we scale the solution to larger areas (e.g., 1000 hectares) without re-solving?*

This document provides rigorous proofs that, under specific conditions, the answer is affirmative. This result has profound implications for quantum optimization, where hardware constraints limit problem size.

2 Problem Formulation

2.1 Basic Notation

Let us establish the fundamental notation:

Definition 2.1 (Problem Parameters). *The binary plot allocation problem $\mathcal{P}$ is characterized by:*

- $\mathcal{F} = \{1, 2, \ldots, n\}$*: Set of n plots (equal-sized land parcels)*
- $\mathcal{C} = \{1, 2, \ldots, m\}$*: Set of m crops*
- $A \in \mathbb{R}_{>0}$*: Total land area (hectares)*
- $a = A/n$*: Area of each individual plot*
- $B_c \in \mathbb{R}_{\geq 0}$*: Benefit density of crop c (value per unit area)*
- $\mathcal{G} = \{G_1, G_2, \ldots, G_K\}$*: Partition of $\mathcal{C}$ into K food groups*

Definition 2.2 (Decision Variables). *The decision variables are:*

$$Y_{p,c} \in \{0, 1\} \quad \forall p \in \mathcal{F}, \forall c \in \mathcal{C} \tag{1}$$

where $Y_{p,c} = 1$ if and only if crop c is assigned to plot p.

2.2 Objective Function

Definition 2.3 (Area-Normalized Objective). *The objective function maximizes the total agricultural benefit, normalized by total land area:*

$$Z(Y) = \frac{1}{A} \sum_{p \in \mathcal{F}} \sum_{c \in \mathcal{C}} a \cdot B_c \cdot Y_{p,c} \tag{2}$$

Lemma 2.4 (Objective Simplification). *The objective function (2) simplifies to:*

$$Z(Y) = \frac{1}{n} \sum_{p \in \mathcal{F}} \sum_{c \in \mathcal{C}} B_c \cdot Y_{p,c} \tag{3}$$

which is independent of the total area A.

Proof. Substituting $a = A/n$ into (2):

$$Z(Y) = \frac{1}{A} \sum_{p \in \mathcal{F}} \sum_{c \in \mathcal{C}} \frac{A}{n} \cdot B_c \cdot Y_{p,c} \tag{4}$$

$$= \frac{1}{A} \cdot \frac{A}{n} \sum_{p \in \mathcal{F}} \sum_{c \in \mathcal{C}} B_c \cdot Y_{p,c} \tag{5}$$

$$= \frac{1}{n} \sum_{p \in \mathcal{F}} \sum_{c \in \mathcal{C}} B_c \cdot Y_{p,c} \tag{6}$$

The total area A cancels completely. $\square$

2.3 Constraint Set

Definition 2.5 (Combinatorial Constraint Set). *The constraint set $\mathcal{K}$ consists of purely combinatorial (count-based) constraints:*

(C1) Plot Assignment Constraint: *Each plot is assigned to at most one crop:*

$$\sum_{c \in \mathcal{C}} Y_{p,c} \leq 1 \quad \forall p \in \mathcal{F} \tag{7}$$

(C2) Food Group Minimum Constraints: *For each food group G_k, at least $N_k^{\min}$ plots must be assigned:*

$$\sum_{p \in \mathcal{F}} \sum_{c \in G_k} Y_{p,c} \geq N_k^{\min} \quad \forall k \in \{1, \ldots, K\} \tag{8}$$

(C3) Food Group Maximum Constraints: *For each food group G_k, at most $N_k^{\max}$ plots can be assigned:*

$$\sum_{p \in \mathcal{F}} \sum_{c \in G_k} Y_{p,c} \leq N_k^{\max} \quad \forall k \in \{1, \ldots, K\} \tag{9}$$

(C4) Maximum Plots Per Crop: *Each crop c can be assigned to at most K_c plots:*

$$\sum_{p \in \mathcal{F}} Y_{p,c} \leq K_c \quad \forall c \in \mathcal{C} \tag{10}$$

Remark 2.6. *Crucially, none of the constraints (C1)–(C4) depend on the total area A or the individual plot area a. They are purely combinatorial, depending only on counts of assignments.*

2.4 Complete Problem Formulation

Definition 2.7 (Binary Plot Allocation Problem). *The complete optimization problem is:*

$$\mathcal{P}(n, A, \mathcal{K}): \quad Z^* = \max_{Y \in \{0,1\}^{n \times m}} Z(Y) \quad \textit{subject to } Y \in \mathcal{K} \tag{11}$$

where $\mathcal{K}$ denotes the feasible region defined by constraints (C1)–(C4).

3   Scale Invariance Theorem

We now establish the fundamental scale invariance property.

Theorem 3.1 (Pure Scale Invariance)**.** *Consider two instances of the binary plot allocation problem:*

$$\mathcal{P}_1 = \mathcal{P}(n, A_1, \mathcal{K}) \tag{12}$$
$$\mathcal{P}_2 = \mathcal{P}(n, A_2, \mathcal{K}) \tag{13}$$

with identical parameters $(n, m, \{B_c\}, \{N_k^{\min}\}, \{N_k^{\max}\}, \{K_c\})$ but different total areas $A_1 \neq A_2$.
  Then:

  (i) The feasible regions are identical: $\mathcal{K}_1 = \mathcal{K}_2$

  (ii) The optimal objectives are equal: $Z_1^* = Z_2^*$

  (iii) The optimal solutions coincide: $Y_1^* = Y_2^*$

Proof. We prove each claim:

(i) Feasibility equivalence: The feasible region $\mathcal{K}$ is defined by constraints (C1)–(C4). Examining each:

  - Constraint (C1): $\sum_c Y_{p,c} \leq 1$ — no dependence on A

  - Constraint (C2): $\sum_{p,c \in G_k} Y_{p,c} \geq N_k^{\min}$ — no dependence on A

  - Constraint (C3): $\sum_{p,c \in G_k} Y_{p,c} \leq N_k^{\max}$ — no dependence on A

  - Constraint (C4): $\sum_p Y_{p,c} \leq K_c$ — no dependence on A

Since none of the constraint coefficients or bounds depend on A, the feasible regions are identical.

(ii) Objective equivalence: By Lemma 2.4, the objective function is:

$$Z(Y) = \frac{1}{n} \sum_{p,c} B_c \cdot Y_{p,c} \tag{14}$$

This expression contains no reference to A. Therefore, $Z_1(Y) = Z_2(Y)$ for all feasible Y.

(iii) Solution equivalence: Since $\mathcal{K}_1 = \mathcal{K}_2$ and $Z_1 \equiv Z_2$, we have:

$$Y_1^* = \arg\max_{Y \in \mathcal{K}_1} Z_1(Y) = \arg\max_{Y \in \mathcal{K}_2} Z_2(Y) = Y_2^* \tag{15}$$

(In case of multiple optima, the set of optimal solutions is identical.)   □

Corollary 3.2 (Area Independence)**.** *The optimal assignment pattern Y^* and optimal benefit density Z^* are invariant under arbitrary positive rescaling of the total land area.*

4   Decomposability Theorem

We now address the more powerful result: when can we solve a smaller problem and replicate the solution?

4.1 Problem Scaling

Definition 4.1 (Scaled Problem). *Given a base problem $\mathcal{P}_1 = \mathcal{P}(n, A, \mathcal{K}_1)$, the k-fold scaled problem is:*

$$\mathcal{P}_k = \mathcal{P}(kn, kA, \mathcal{K}_k) \tag{16}$$

where the constraint parameters are scaled as follows:

- *Plot assignment (C1): unchanged (applies per-plot)*

- *Food group bounds: $N_{k,g}^{\min} = k \cdot N_{1,g}^{\min}$, $N_{k,g}^{\max} = k \cdot N_{1,g}^{\max}$*

- *Max plots per crop: $K_{k,c} = k \cdot K_{1,c}$*

Definition 4.2 (Separable Constraints). *The constraint set $\mathcal{K}_k$ is* separable *if we can partition the kn plots into k groups $\mathcal{F}_1, \ldots, \mathcal{F}_k$ (each of size n) such that:*

(a) Each constraint involves variables from only one group

(b) The constraint structure is identical across groups

4.2 Tiled Constraints

Definition 4.3 (Tiled Constraint Set). *The* tiled *constraint set $\mathcal{K}_k^{tile}$ is defined as follows. Partition the plots $\{1, \ldots, kn\}$ into k groups:*

$$\mathcal{F}_j = \{(j-1)n + 1, \ldots, jn\} \quad \text{for } j = 1, \ldots, k \tag{17}$$

The constraints for each group j are:

(C1) Plot Assignment:

$$\sum_{c \in \mathcal{C}} Y_{p,c} \leq 1 \quad \forall p \in \mathcal{F}_j \tag{18}$$

(C2)-(C3) Food Group Constraints (per tile):

$$N_g^{\min} \leq \sum_{p \in \mathcal{F}_j} \sum_{c \in G_g} Y_{p,c} \leq N_g^{\max} \quad \forall g, \forall j \tag{19}$$

(C4) Max Plots Per Crop (per tile):

$$\sum_{p \in \mathcal{F}_j} Y_{p,c} \leq K_c \quad \forall c \in \mathcal{C}, \forall j \tag{20}$$

Lemma 4.4 (Tiled Constraints are Separable). *The tiled constraint set $\mathcal{K}_k^{tile}$ is separable in the sense of Definition 4.2.*

Proof. By construction, each constraint in $\mathcal{K}_k^{tile}$ involves only variables $Y_{p,c}$ where $p \in \mathcal{F}_j$ for a single tile j. The constraints for different tiles are completely decoupled. Furthermore, the constraint structure (coefficients and bounds) is identical across all tiles. $\square$

4.3 Main Decomposability Result

Theorem 4.5 (Solution Decomposability). *Consider the base problem $\mathcal{P}_1 = \mathcal{P}(n, A, \mathcal{K}_1)$ with optimal solution $Y^* \in \{0,1\}^{n \times m}$ and optimal objective Z^*.*
  Consider the k-fold tiled problem $\mathcal{P}_k^{tile}$ with tiled constraints $\mathcal{K}_k^{tile}$.
  Then:

  (i) The replicated solution $\tilde{Y} = [Y^*, Y^*, \ldots, Y^*]$ (k copies) is feasible for $\mathcal{P}_k^{tile}$

  (ii) The replicated solution is optimal: $\tilde{Y} \in \arg\max_{Y \in \mathcal{K}_k^{tile}} Z_k(Y)$

 (iii) The optimal objectives are equal: $Z_k^* = Z_1^*$

Proof. We establish each claim rigorously.

(i) Feasibility of replicated solution: Define the replicated solution $\tilde{Y} \in \{0,1\}^{kn \times m}$ by:

$$\tilde{Y}_{p,c} = Y^*_{((p-1) \mod n)+1, c} \quad \forall p \in \{1, \ldots, kn\}, \forall c \in \mathcal{C} \tag{21}$$

For each tile j, the variables $\{\tilde{Y}_{p,c} : p \in \mathcal{F}_j\}$ are a copy of Y^*.
Constraint (C1): For any plot $p \in \mathcal{F}_j$:

$$\sum_{c \in \mathcal{C}} \tilde{Y}_{p,c} = \sum_{c \in \mathcal{C}} Y^*_{p',c} \leq 1 \tag{22}$$

where $p' = ((p-1) \mod n) + 1$ is the corresponding plot in the base problem. The inequality holds because Y^* satisfies (C1).
Constraints (C2)-(C3): For each tile j and food group g:

$$\sum_{p \in \mathcal{F}_j} \sum_{c \in G_g} \tilde{Y}_{p,c} = \sum_{p=1}^{n} \sum_{c \in G_g} Y^*_{p,c} \tag{23}$$

Since Y^* satisfies $N_g^{\min} \leq \sum_{p, c \in G_g} Y^*_{p,c} \leq N_g^{\max}$, the tiled constraints are satisfied.
Constraint (C4): For each tile j and crop c:

$$\sum_{p \in \mathcal{F}_j} \tilde{Y}_{p,c} = \sum_{p=1}^{n} Y^*_{p,c} \leq K_c \tag{24}$$

Again inherited from feasibility of Y^*.

(ii) Optimality of replicated solution: The objective for the tiled problem is:

$$Z_k(\tilde{Y}) = \frac{1}{kn} \sum_{p=1}^{kn} \sum_{c \in \mathcal{C}} B_c \cdot \tilde{Y}_{p,c} \tag{25}$$

Since the tiled constraints are separable (Lemma 4.4), we can decompose:

$$Z_k(\tilde{Y}) = \frac{1}{kn} \sum_{j=1}^{k} \sum_{p \in \mathcal{F}_j} \sum_{c \in \mathcal{C}} B_c \cdot \tilde{Y}_{p,c} \tag{26}$$

$$= \frac{1}{kn} \sum_{j=1}^{k} \left(\sum_{p=1}^{n} \sum_{c \in \mathcal{C}} B_c \cdot Y_{p,c}^* \right) \tag{27}$$

$$= \frac{1}{kn} \cdot k \cdot \left(\sum_{p=1}^{n} \sum_{c \in \mathcal{C}} B_c \cdot Y_{p,c}^* \right) \tag{28}$$

$$= \frac{1}{n} \sum_{p=1}^{n} \sum_{c \in \mathcal{C}} B_c \cdot Y_{p,c}^* \tag{29}$$

$$= Z_1^* \tag{30}$$

Now suppose for contradiction that $\tilde{Y}$ is not optimal, i.e., there exists $\hat{Y} \in \mathcal{K}_k^{\text{tile}}$ with $Z_k(\hat{Y}) > Z_k(\tilde{Y}) = Z_1^*$.

By separability, we can write:

$$Z_k(\hat{Y}) = \frac{1}{k} \sum_{j=1}^{k} Z^{(j)}(\hat{Y}) \tag{31}$$

where $Z^{(j)}(\hat{Y}) = \frac{1}{n} \sum_{p \in \mathcal{F}_j} \sum_c B_c \hat{Y}_{p,c}$ is the contribution from tile j.

Since $Z_k(\hat{Y}) > Z_1^*$ and $Z_k(\hat{Y})$ is the average of $Z^{(1)}, \ldots, Z^{(k)}$, there must exist some tile j^* with:

$$Z^{(j^*)}(\hat{Y}) > Z_1^* \tag{32}$$

But the restriction of $\hat{Y}$ to tile j^* is feasible for the base problem $\mathcal{P}_1$ (since the constraints are identical). This contradicts the optimality of Y^* for $\mathcal{P}_1$.

Therefore, $\tilde{Y}$ is optimal.

(iii) Objective equality: From part (ii), we have $Z_k^* = Z_k(\tilde{Y}) = Z_1^*$. $\qquad\square$

Corollary 4.6 (Computational Reduction)**.** *To solve the tiled problem $\mathcal{P}_k^{tile}$ with kn plots, it suffices to:*

1. *Solve the base problem $\mathcal{P}_1$ with n plots*

2. *Replicate the solution k times*

The computational complexity is $O(solve(n))$ rather than $O(solve(kn))$, providing potentially exponential savings for combinatorial optimization.

5 Conditions for Decomposability

5.1 When Decomposition Holds

Proposition 5.1 (Sufficient Conditions for Decomposability)**.** *The decomposability result (Theorem 4.5) holds when:*

(S1) The objective is linear and area-normalized

(S2) All constraints are count-based (not area-based)

(S3) *Constraints are applied locally (per tile) rather than globally*

(S4) *The constraint bounds are identical across tiles*

Proof. Conditions (S1) and (S2) ensure the objective and constraints are independent of total area (Section 3). Conditions (S3) and (S4) ensure separability (Definition 4.2), which is required for the decomposition argument in Theorem 4.5. □

5.2 When Decomposition Fails

Proposition 5.2 (Failure Conditions). *The decomposability result* fails *under any of the following conditions:*

(F1) Global constraints: If there exist constraints that couple multiple tiles, e.g.,

$$\sum_{p=1}^{kn} Y_{p,c} \leq K_c^{\text{global}} \tag{33}$$

then the problem does not decompose. Replicating a feasible base solution may violate the global bound.

Example 5.3. *Base problem: $n = 25$ plots, $K_c = 5$ (max 5 plots per crop). Scaled problem with global constraint: $kn = 250$ plots, $K_c^{global} = 20$. Replicated solution uses $k \cdot 5 = 50 > 20$ plots, violating the constraint.*

(F2) Non-linear objectives: If the objective has diminishing returns, e.g.,

$$Z(Y) = \sum_{c \in \mathcal{C}} f_c \left(\sum_{p=1}^{kn} a \cdot Y_{p,c} \right) \tag{34}$$

where f_c is concave, then replication may not be optimal.

Example 5.4. *Market saturation: $f_c(x) = \log(1 + x)$. The marginal value of additional area for crop c decreases. A diverse allocation across tiles may outperform replication.*

(F3) Global diversity constraints: If there exist constraints requiring variety across the entire domain:

$$|\{c \in \mathcal{C} : \exists p, Y_{p,c} = 1\}| \geq D^{\text{global}} \tag{35}$$

Replication uses the same crops in each tile, potentially failing to achieve global diversity.

(F4) Shared resource constraints: If tiles compete for a shared resource (water, labor, equipment):

$$\sum_{j=1}^{k} \sum_{p \in \mathcal{F}_j} \sum_{c \in \mathcal{C}} r_c \cdot Y_{p,c} \leq R^{\text{total}} \tag{36}$$

where R^{total} does not scale with k, decomposition fails.

6 Implications for Quantum Optimization

6.1 QPU Size Limitations

Current quantum processing units (QPUs) are limited in the number of qubits and connectivity. For the binary plot allocation problem:

- Number of binary variables: $n \cdot m$ (plots $\times$ crops)

- Number of logical qubits needed: $O(n \cdot m)$

- Physical qubits after embedding: $O(n \cdot m \cdot \chi)$ where χ is chain length

Proposition 6.1 (QPU Scaling via Decomposition). *If the decomposability conditions (Proposition 5.1) hold, then a problem with kn plots can be solved using a QPU sized for n plots, with:*

1. *Same solution quality (optimal)*

2. *$k\times$ reduction in qubit requirements*

3. *No additional QPU calls (solve once, replicate k times)*

6.2 Practical Algorithm

Algorithm 1 Scalable Plot Allocation via Decomposition

Require: Target area A_{target}, number of tiles k, crops $\mathcal{C}$, constraints $\mathcal{K}$
Ensure: Optimal allocation Y^* for A_{target}
 1: $A_{\text{base}} \leftarrow A_{\text{target}}/k$
 2: $n_{\text{base}} \leftarrow$ desired plots per tile $\triangleright$ Choose n_{base} to fit QPU
 3: Verify decomposability conditions (Prop. 5.1)
 4: **if** conditions not satisfied **then**
 5:      **return** solve full problem (no decomposition)
 6: **end if**
 7: $Y^*_{\text{base}} \leftarrow \text{SolveOnQPU}(n_{\text{base}}, A_{\text{base}}, \mathcal{K})$ $\triangleright$ Solve base problem
 8: $Y^* \leftarrow \text{Replicate}(Y^*_{\text{base}}, k)$ $\triangleright$ Tile the solution
 9: **return** Y^*

---

7 Relationship to Continuous Formulation

7.1 Discretization Gap

Definition 7.1 (Continuous Relaxation). *The continuous relaxation replaces binary variables with:*

$$A_{p,c} \in [0, a] \quad \text{(area allocated to crop c on plot p)} \tag{37}$$

with objective:

$$Z^{cont}(A) = \frac{1}{A} \sum_{p,c} B_c \cdot A_{p,c} \tag{38}$$

Definition 7.2 (Discretization Gap). *The discretization gap is:*

$$\epsilon = \frac{Z^{cont,*} - Z^{bin,*}}{Z^{cont,*}} \tag{39}$$

where $Z^{cont,*}$ and $Z^{bin,*}$ are the optimal continuous and binary objectives.

Theorem 7.3 (Gap Preservation Under Scaling). *If the binary problem has discretization gap ϵ for n plots, then the k-fold tiled problem also has gap ϵ.*

Proof. By Theorem 4.5, $Z_k^{\text{bin},*} = Z_1^{\text{bin},*}$.

By similar arguments, the continuous problem also decomposes (linear objective, separable constraints), giving $Z_k^{\text{cont},*} = Z_1^{\text{cont},*}$.

Therefore:

$$\epsilon_k = \frac{Z_k^{\text{cont},*} - Z_k^{\text{bin},*}}{Z_k^{\text{cont},*}} = \frac{Z_1^{\text{cont},*} - Z_1^{\text{bin},*}}{Z_1^{\text{cont},*}} = \epsilon_1 \tag{40}$$

□

Remark 7.4. *The discretization gap depends on the* granularity *(plots per unit area), not the total area. Finer discretization (more plots per hectare) reduces the gap but increases computational cost.*

8 Numerical Verification

The theoretical results can be verified empirically:

1. Solve the base problem $\mathcal{P}_1(n = 25, A = 100)$

2. Solve the scaled problem $\mathcal{P}_k(kn = 250, A = 1000)$ with tiled constraints

3. Verify: $Z_k^* = Z_1^*$ and $Y_k^* =$ replication of Y_1^*

See the accompanying benchmark script `scaling_formulation_benchmark.py` for empirical validation.

9 Conclusion

We have established rigorous mathematical foundations for the scale invariance of the binary plot allocation problem:

1. **Pure Scale Invariance** (Theorem 3.1): The optimal allocation is independent of total land area when the number of plots and combinatorial constraints are fixed.

2. **Solution Decomposability** (Theorem 4.5): Under tiled (local) constraints, the optimal solution to a large problem is the replication of the optimal solution to a smaller subproblem.

3. **Conditions** (Propositions 5.1, 5.2): Decomposition requires local constraints and fails under global constraints, non-linear objectives, or shared resources.

4. **Gap Preservation** (Theorem 7.3): The discretization gap from the continuous relaxation is preserved under scaling.

These results enable significant computational savings in quantum optimization by allowing problems to be solved at a scale that fits current QPU limitations, then scaled to arbitrary sizes through replication.

A Notation Summary

Symbol	Meaning
n	Number of plots
m	Number of crops
A	Total land area
$a = A/n$	Area per plot
$\mathcal{F}$	Set of plots
$\mathcal{C}$	Set of crops
G_k	Food group k
B_c	Benefit density of crop c
$Y_{p,c}$	Binary: crop c assigned to plot p
$N_k^{\min}, N_k^{\max}$	Bounds on food group k
K_c	Max plots for crop c
$Z(Y)$	Objective function
Z^*	Optimal objective value
Y^*	Optimal solution
$\mathcal{K}$	Feasible region (constraint set)
k	Scaling factor (number of tiles)
ϵ	Discretization gap

Table 1: Notation used in this document