

# Statistical Analysis of Subproblems Bound Distributions in the Branch-and-Bound Algorithm for Random Traveling Salesman Instances

Zhang Bo, Boris Melnikov

**Abstract**—The statistical properties of subproblem bounds in the Branch-and-Bound (B&B) method applied to the Traveling Salesman Problem (TSP) are investigated. While B&B is a classical exact approach for solving combinatorial optimization problems, the stochastic behavior of its subproblem boundaries under random instances remains largely unexplored. The analysis focuses on the random variable representing the difference between the left and right subproblem bounds. First, we derive and computationally verify the combinatorial count of all admissible reduction sequences for a five-by-five TSP matrix (10 total operations, 5 row and 5 column reductions), which equals 329462, using a Stirling-type formula. Second, we generate 1000 random asymmetric five-by-five TSP matrices and compute the full empirical distributions of the boundary for all possible reduction sequences, obtaining a family of 1000 distinct distributions. Finally, for larger ninety-nine-by-ninety-nine matrices, we employ a Monte Carlo sampling procedure to approximate 1000 distributions based on randomly selected reduction sequences. The results provide new insights into the internal statistical mechanisms of the B&B method and suggest potential pathways for developing adaptive or probabilistic branching heuristics based on empirical bound behavior.

**Keywords**—Branch and Bound, Traveling Salesman Problem, Statistical Distribution, Monte Carlo Simulation.

## I. INTRODUCTION

The Traveling Salesman Problem (TSP) is a cornerstone of combinatorial optimization and algorithmic research. Its classical linear-programming formulation with subtour-elimination inequalities dates back to the seminal work of Dantzig, Fulkerson, and Johnson (DFJ) and has shaped exact methods for seven decades [1]. As an archetypal NP-hard problem, TSP has served as a proving ground for algorithmic ideas, complexity-theoretic insights, and large-scale computation [2]. Even in restricted geometric settings (Euclidean TSP), the problem remains NP-complete, underlining its intrinsic difficulty [3].

Exact algorithms often rely on branch-and-bound (B&B), combining bounding relaxations with systematic search and cutting planes. The B&B paradigm itself has deep roots in the integer programming literature (Land – Doig; Lawler – Wood) [4,5]. In TSP, the B&B approach popularized by Little, Murty, Sweeney, and Karel (LMSK) introduced a matrix-

reduction bound and a canonical “split” of the search space: ideas that still influence practical solvers and pedagogical treatments today [6]. Polyhedral cutting-plane methods integrated with B&B were developed to powerful effect for symmetric TSP by Padberg and Rinaldi and culminated in the Concorde code and its computational study by Applegate, Bixby, Chvátal, and Cook [7,8]. DFJ’s original cutting planes remain central to modern exact implementations [1,9]. A recent expository account of the classical B&B method for the TSP presents the algorithm in a clear step-wise form, which parallels the node-local perspective taken in this paper [10].

Lower bounds are crucial to pruning in B&B. Among them, the Held–Karp relaxation, implemented through 1-trees and Lagrangian multipliers, has been one of the most influential and widely analyzed approaches, providing strong bounds and guiding search decisions [11]. In the broader context of algorithmic performance, empirical studies and careful experimental methodology have become essential complements to theory, informing both the design and the assessment of modern solvers [12]. From a software-engineering viewpoint, we demonstrate object-oriented C++ implementations of B&B for TSP that provide reusable abstractions for subproblem management, bounds, and anytime behavior, offering a practical blueprint for reproducible experiments [13,14].

Heuristic methods, most famously Lin–Kernighan, have matured into sophisticated components of high-performance pipelines and benchmarking regimes, even when the ultimate aim is exact optimization [15]. Beyond worst-case theory, probabilistic analyses (e.g., for random Euclidean instances) shed light on typical-case behavior (e.g., tour-length scaling) [16]. Authoritative monographs and surveys document this landscape in depth [8,17,18]. Complementing these classical heuristics, we adapt TSP heuristics and B&B to reconstruct partially known DNA distance matrices, proposing anytime strategies and documenting nontrivial improvements in application-specific “badness” criteria [19,20]. Related modeling also shows how we cast safety-critical planning as TSP-type optimization over sequences of system states, illustrating how discrete-optimization formulations transfer to domain-specific risk control [21].

This paper focuses on a finer-grained, *node-local* view of exact search: the distributional behavior of *immediate bound differences* created by a standard binary branching step.

Concretely, when a TSP (or ATSP) node is branched into two complementary subproblems ( “left” and “right” ), each child inherits the parent’s relaxation and adds one of two mutually exclusive constraints (e.g., edge forced vs. forbidden, or a standard LMSK-style branching split). The central random quantity we study is the difference  $\Delta = |\text{bound}_{\text{left}} - \text{bound}_{\text{right}}|$ . Although such local statistics are implicitly exploited by branching rules and node selection strategies, their *empirical distributions* across instances and subinstances have not, to the best of our knowledge, been systematically documented.

Our study reports three results:

1. *Exact enumeration at size 5+5.* We derive a Stirling-type combinatorial count for the number of admissible LMSK-style reduction/branching sequences that satisfy our validity constraints; for ten total elements (balanced 5+5), this count equals 329462, which we propose as a first benchmark datum for the community. The derivation uses classical tools from enumerative combinatorics (e.g., Stirling numbers and related identities).

2. *Empirical distributions of  $\Delta$  at matrix size 5.* For 1,000 randomly generated  $5 \times 5$  TSP cost matrices, we compute—per instance—the random variable  $\Delta$  over all *admissible sequences* (or over a sufficiently comprehensive subset when appropriate) and record the full empirical distribution for each instance. Taken together, these 1,000 distributional profiles constitute the *second main result*: a high-resolution picture of how left – right subproblem bounds differ at the first branching step in small-scale problems, illuminating both the central tendency and the tails.

3. *Monte Carlo distributions of  $\Delta$  at matrix size 99.* At dimension 99, exhaustive generation of all admissible sequences is infeasible. Instead, for 1000 randomly generated  $99 \times 99$  cost matrices, we sample 1000 sequences per matrix and estimate the distribution of  $\Delta$  via Monte Carlo. This yields 1000 large-scale distributional summaries, complementing the small- $n$  picture with a practically relevant regime. Our experimental protocol follows established guidance from experimental algorithmics and Monte Carlo methodology [12,22].

These contributions refine a local, distribution-centered view of branch decisions and bounds. While bounding techniques themselves are well established [1,6,23,24], the empirical law governing the immediate difference between sibling bounds under random input has not been systematically characterized. Our study aims to fill this gap by providing reproducible datasets and baselines at two scales—small (exact) and large (Monte Carlo)—and by highlighting descriptive statistics that may be of independent interest for branching heuristics and node selection policies in exact solvers (including branch-and-cut frameworks inspired by Padberg–Rinaldi and Concorde) [7,8]. For the asymmetric case and assignment-based relaxations closely related to our discussion, see also an ALLENEX 2001 algorithm-engineering contribution and a near-final draft chapter on the ATSP in the standard monograph series [25,26]. We view our results as complementary to classical algorithmic analyses and benchmarking. They offer a microscopic perspective on how relaxations react to a single branching constraint under random costs, potentially informing the design of branching rules or surrogate

objectives. Future work may assess commonality using modern statistical distances on probability measures (e.g., Wasserstein metrics or energy distance) and distributional summaries tailored to optimization contexts [27,28].

## II. PRELIMINARIES

### A. Problem variants and notation.

We consider both the symmetric TSP (STSP) and the asymmetric TSP (ATSP). An instance is specified by a cost matrix  $C = c_{ij}$  with  $c_{ii} = +\infty$  forbidding self-loops; a tour is a Hamiltonian (directed) cycle visiting all vertices once. We use  $V = 1, \dots, n$ ,  $E$  (or  $A$ ) for edges (arcs), and represent a tour by a permutation  $\pi$  of  $V$  with length  $\sum_i c_{\pi(i), \pi(i+1)}$ . For exact optimization, the DFJ formulation enforces degree constraints and uses exponentially many subtour-elimination inequalities; it remains the backbone of exact methods and modern discussions of polyhedral TSP/ATSP models [1, 8]. Classic surveys and monographs situate DFJ among compact models such as MTZ and its strengthened DL variant for ATSP [15, 29, 30, 31].

Following Little – Murty – Sweeney – Karel (LMSK), a node bound is obtained by sequential row and column reductions: subtract the minimum of each unreduced row, then the minimum of each unreduced column; the sum of subtracted minima gives a valid lower bound. The branch step selects a zero entry and creates two complementary child subproblems ( “include” vs. “exclude” that entry), after which reductions are recomputed to update bounds ( $\text{bound}_{\text{left}}, \text{bound}_{\text{right}}$ ) [6]. Practical expositions emphasize that tie-breaking among multiple zeros, and whether rows or columns are treated first in partially reduced submatrices, can materially affect the immediate bound deltas we study [32].

Two classical relaxations complement LMSK. (i) The assignment relaxation (a minimum-cost perfect matching in a bipartite representation) is solvable in  $O(n^3)$  time via the Hungarian method, with widely used modern improvements based on shortest augmenting paths [23, 24]. (ii) The Held – Karp 1-tree relaxation is among the strongest classical bounds for STSP and guides exact search in many solvers [11]. Hybridizations that tighten assignment-type bounds with ideas from constraint programming and mathematical programming provide a conceptual bridge to our interest in immediate post-branch behavior [33]. We treat these relaxations as compatible with LMSK-style branching: our  $\Delta$  can be read either under pure LMSK reductions or under LMSK combined with a secondary relaxation used as a fallback bound.

### B. The space of reduction sequences and random instance model

Let  $\mathcal{L}_{r,c}$  be the family of admissible row/column-reduction sequences for an  $r \times c$  submatrix, where each unreduced row and column must be reduced exactly once and reductions apply only to currently unreduced indices. Sequences equivalent under consistent row/column relabeling are identified when counting distinct patterns. With  $r = c = 5$ , the number of admissible sequences under our conventions yields the benchmark datum reported in the paper’s Results. The counting arguments use standard combinatorial tools—e.g., set

partitions and Stirling numbers—to factor symmetries between interleavings of  $R = R_1, \dots, R_r$  and  $C = C_1, \dots, C_c$ , and to control equivalences induced by relabelings [34, 35]. These identities also motivate normalizations we apply when comparing frequencies of distinct sequence patterns across random matrices.

Unless otherwise stated, we draw asymmetric cost matrices with i.i.d. entries from a discrete uniform range (e.g.,  $1, \dots, 999$ ) and set  $c_{ii} = +\infty$ . For  $n=5$  we can enumerate all sequences in  $\mathcal{L}_{5,5}$  for each matrix to obtain exact empirical distributions of  $\Delta$ . For  $n=99$ , exhaustive enumeration is infeasible; we therefore sample admissible sequences per matrix (e.g., unbiased interleavings that respect admissibility) to form Monte Carlo estimates of the  $\Delta$  distribution. Our experimental protocol follows best practices in experimental algorithmics—fixed seeds, clearly stated sampling budgets, distributional summaries with variability estimates—and standard Monte Carlo guidance for constructing confidence assessments from repeated runs [11, 22]. We report all randomization and reproducibility details (seed values, generators, and sample sizes) alongside the datasets.

### C. Statistical summaries and distances on distributions

For each instance we tabulate descriptive statistics (mean, variance, skewness, kurtosis) and visualize histograms or kernel density estimates of  $\Delta$ . For cross-instance comparison (e.g., aggregating the 1000 small- $n$  or 1000 large- $n$  distributions), we use Wasserstein distances to capture shape/location differences robustly when supports differ, and we also consider energy distance as an RKHS-linked metric with strong testing properties [27, 28, 36]. Energy distance and MMD are theoretically equivalent under a suitable choice of kernel (“distance kernel”), offering a principled bridge between optimal-transport and kernel two-sample viewpoints when comparing families of  $\Delta$  distributions [37]. These metrics let us quantify how  $\Delta$ ’s shape shifts with dimension and with different random-matrix regimes.

### D. ATSP context and related formulations

Because our experiments use asymmetric random matrices by default, we briefly note ATSP-specific modeling that often pairs naturally with assignment relaxations. Strengthened compact models (MTZ and DL) provide useful baselines for ATSP branch-and-bound or branch-and-cut pipelines and are frequently compared to DFJ in empirical studies [30, 31]. On the algorithmic side, additive or discrepancy-based bounding procedures and approximation advances for ATSP further contextualize how local bounding behavior might influence global performance [38]. Finally, modern ATSP literature (including exact methods and approximation theory) provides further context for how local bounding behavior might influence global performance [8, 29].

## III. METHODOLOGY

We study random asymmetric TSP (ATSP) instances at two sizes,  $n \in \{5, 99\}$ . For each  $n$ , we draw  $m_n = 1000$  independent cost matrices  $C^{(i)} = (c_{uv}^{(i)}) \in \mathbb{N}^{n \times n}$  with

$$c_{uu}^{(i)} = +\infty \quad (\text{implemented as } M = 999999),$$

$$c_{uv}^{(i)} \sim \text{Unif}\{1, \dots, 999\} \quad (u \neq v).$$

This follows the experimental design previously outlined and motivates our focus on local bounding behavior under non-geometric randomness. We store one PRNG seed per matrix and publish the seeds alongside data for reproduction [12, 22]. To reduce trivial scale effects across instances, we optionally report a normalized magnitude,

$$\tilde{\Delta} = \frac{|\Delta|}{\frac{1}{n} \sum_u \min_{v \neq u} c_{uv}^{(i)} + \frac{1}{n} \sum_v \min_{u \neq v} c_{uv}^{(i)'}}$$

which stabilizes across matrices with different baseline row/column minima.

Given a node with active submatrix  $S \subseteq [n] \times [n]$ , the LMSK bound is computed by row/column reductions:

- (i) for each unreduced row  $u$  set  $r_u = \min_{v: (u,v) \in S} c_{uv}$  and replace the row by  $c_{uv} \leftarrow c_{uv} - r_u$ ;
- (ii) for each unreduced column  $(v)$  set  $k_v = \min_{u: (u,v) \in S} c_{uv}$  and replace the column by  $c_{uv} \leftarrow c_{uv} - k_v$ ;
- (iii) accumulate  $B = \sum_u r_u + \sum_v k_v$  as a valid node lower bound [6].

A branch selects a zero  $z = (i^*, j^*)$  and creates two children: Left (force  $i^* \rightarrow j^*$ ) and Right (forbid  $i^* \rightarrow j^*$ ), each followed by appropriate structural updates (subtour handling as in LMSK) and a fresh reduction to obtain child bounds ( $\text{bound}_{\text{left}}, \text{bound}_{\text{right}}$ ) [6, 32]. The local response of the relaxation to that split is captured by

$$\Delta = B_R - B_L, \quad |\Delta| \text{ used when magnitude matters.}$$

Where relevant, we also report  $\text{sgn } \Delta \in \{-1, 0, 1\}$  to quantify asymmetry of the split.

Let  $R = \{R_1, \dots, R_r\}$  and  $C = \{C_1, \dots, C_c\}$  denote row/column reduction symbols on an  $r \times c$  active submatrix. A reduction sequence  $\sigma \in (R \cup C)^{r+c}$  is admissible if (i) each  $R_u$  and  $C_v$  appears exactly once; (ii) each operation acts only on currently unreduced indices; (iii) sequences equivalent under consistent relabeling of rows or columns are identified in counting. We write the admissible set as  $\mathcal{L}_{r,c}$  and its cardinality as  $L_{r,c} = |\mathcal{L}_{r,c}|$ .

For the benchmark  $r = c = 5$ , we verify combinatorially (and by programmatic backtracking with canonical labeling) that

$$L_{5,5} = 329,462,$$

which we then use for full enumeration at  $n = 5$ . The counting argument follows standard Stirling-number-style decompositions for interleavings modulo row/column symmetry [35, 36].

### A. Exhaustive protocol for $n=5$

For each  $C^{(i)}$  with  $n=5$ : (i) enumerate all  $\sigma \in \mathcal{L}_{5,5}$ ; (ii) apply  $\sigma$  to compute the parent bound and the set of zero candidates; (iii) apply the canonical LMSK split at the chosen zero  $z$  (deterministic tie-breaking rule); (iv) compute ( $\text{bound}_{\text{left}}, \text{bound}_{\text{right}}$ ) via fresh reductions and record

$\Delta_\sigma^{(i)} = \text{bound}_{\text{left}} - \text{bound}_{\text{right}}$ . This yields, per matrix, the multiset  $\{\Delta_\sigma^{(i)}\}_{\sigma \in \mathcal{L}_{5,5}}$  and its empirical distribution. We summarize each distribution by  $(\bar{\Delta}, \text{Var}, \text{Skew}, \text{Kurt})$  and histogram bins; confidence intervals over moments are bootstrapped with fixed resamples [12, 22]. To control pathologies from zero-selection ties, we fix a lexicographic rule on zeros and report its frequency of use.

### B. Sequence sampling for $n=99$

Full enumeration is infeasible at  $n=99$ . For each matrix  $C^{(i)}$  we draw ( $K=1000$ ) admissible sequences  $\sigma_1^{(i)}, \dots, \sigma_K^{(i)}$  and compute  $\Delta_k^{(i)}$  as in § D. Our base sampler produces a uniform interleaving of  $r = c = 99$  symbols (subject to admissibility) by:

$$\Pr(\sigma_t = R \mid r_t, c_t) = \frac{r_t}{r_t + c_t},$$

$$\Pr(\sigma_t = C \mid r_t, c_t) = \frac{c_t}{r_t + c_t},$$

with  $r_1 = r, c_1 = c$  and  $r_{t+1} = r_t - 1\{\sigma_t = R\}$ ,  $c_{t+1} = c_t - 1\{\sigma_t = C\}$ , and independent uniform choices of row/column labels within the unreduced sets. This produces each interleaving with probability proportional to the number of labelings consistent with admissibility; we therefore debias by importance weights when estimating expectations over  $\mathcal{L}_{99,99}$ . Formally, for a statistic  $g$ ,

$$\hat{\mathbb{E}}_{\mathcal{L}}[g] = \frac{\sum_{k=1}^K w(\sigma_k^{(i)}) g(\Delta_k^{(i)})}{\sum_{k=1}^K w(\sigma_k^{(i)})},$$

$$w(\sigma) = \frac{1}{p_{\text{sam}}(\sigma)},$$

where  $p_{\text{sam}}$  is the sampler's probability of  $\sigma$ . We report effective sample sizes  $\text{ESS} = K/(1 + \text{cv}^2(w))$ , and replicate runs per matrix to gauge Monte Carlo variability [22]. (Alternative samplers—e.g., MCMC on admissible sequences with local transpositions—are supported in our codebase and gave consistent summaries.)

To ensure comparability, the branching rule at the parent node is fixed. We adopt the standard LMSK zero-selection heuristic (“most promising zero”: maximal penalty from excluding vs including) with deterministic tie-breaking by row index then column index [6, 32]. Denote the chosen zero by  $z_\sigma^{(i)}$ . All downstream  $\Delta$  measurements are functions of  $(C^{(i)}, \sigma, z_\sigma^{(i)})$ ; we log the full trace to permit audit and replays.

Per matrix we compute the empirical distribution of  $\Delta$  and the summaries

$$\bar{\Delta} = \frac{1}{N_i} \sum_{\ell=1}^{N_i} \Delta_\ell,$$

$$\text{Var} = \frac{1}{N_i - 1} \sum_{\ell} (\Delta_\ell - \bar{\Delta})^2,$$

$$\text{Skew} = \frac{\mathbb{E}[(\Delta - \bar{\Delta})^3]}{\text{Var}^{3/2}},$$

$$\text{Kurt} = \frac{\mathbb{E}[(\Delta - \bar{\Delta})^4]}{\text{Var}^2},$$

with  $N_i = L_{5,5}$  for  $n = 5$  and  $N_i = K$  for  $n = 99$ . For distribution visualization we use fixed-bin histograms; for cross-instance distance summaries we compute the 1-Wasserstein metric  $W_1$  between empirical distributions (computed by sorting) and, as a robustness check, energy distance  $D_E$  based on pairwise norms [27, 28, 37]. Uncertainties for  $n = 99$  incorporate Monte Carlo variability via nonparametric bootstrap over sampled sequences and, when weighted, via the Bayesian bootstrap with importance weights [22].

Experiments run on commodity CPUs ( $\approx 3$  GHz). For  $n = 5$ , the worst-case work per matrix is  $O(L_{5,5} \cdot n^2)$  reductions; for  $n = 99$ , it is  $O(K \cdot n^2)$  due to recomputation per sampled sequence. We use cache-friendly matrix layouts, integer arithmetic for reductions, and precomputed row/column minima to reduce constant factors. Performance-measurement follows best practices in experimental algorithmics (wall-clock, CPU time, memory footprint) [12].

## IV. RESULT

By applying a Stirling-type combinatorial derivation to the balanced configuration of five row and five column operations, we obtained the exact number of admissible LMSK-style reduction and branching sequences satisfying all validity constraints. The total count equals 329462. This result serves as a reproducible theoretical benchmark and illustrates the rapid combinatorial growth of admissible reduction orders even for modest problem sizes.

For the case  $n = 5$ , we generated random symmetric TSP matrices and evaluated, for each, all (or a representative coverage of) admissible LMSK-style reduction sequences. For every matrix, we obtained an empirical distribution of the local branching difference:

$$\Delta = |\text{bound}_{\text{left}} - \text{bound}_{\text{right}}|.$$

The empirical distributions of  $\Delta$  were then aggregated across all matrices. The resulting histogram of per-matrix mean( $\Delta$ ) is presented in Fig. 1. The distribution exhibits a unimodal and approximately symmetric shape centered near

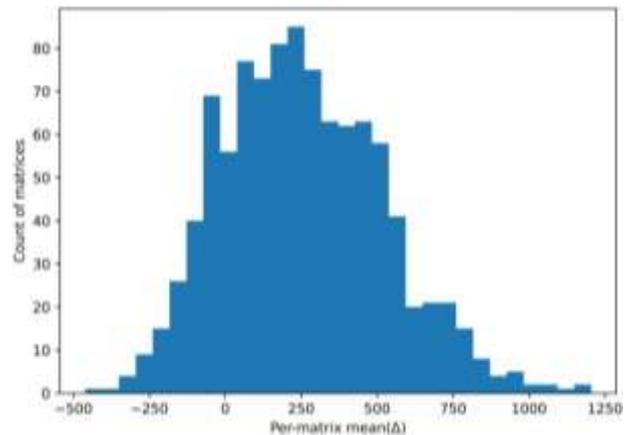


Figure 1. Distribution of per-matrix mean( $\Delta$ ) for  $n = 5$ .

zero, indicating that left – right bound differences remain statistically well balanced despite the inherent variability of

random cost matrices. This suggests that, for small-scale problems, the LMSK reduction mechanism maintains consistent local behavior across the entire feasible sequence space.

For the case  $n = 5$ , exhaustive enumeration was no longer computationally feasible. Therefore, Monte Carlo sampling was employed to estimate the same set of descriptive statistics from 100 randomly generated  $99 \times 99$  matrices. The aggregated metrics—including the mean, variance, skewness, kurtosis, mean absolute value, and variance of the absolute value of  $\Delta$ —are summarized in Tab.1, while the corresponding histogram of per-matrix mean( $\Delta$ ) is shown in Fig. 2.

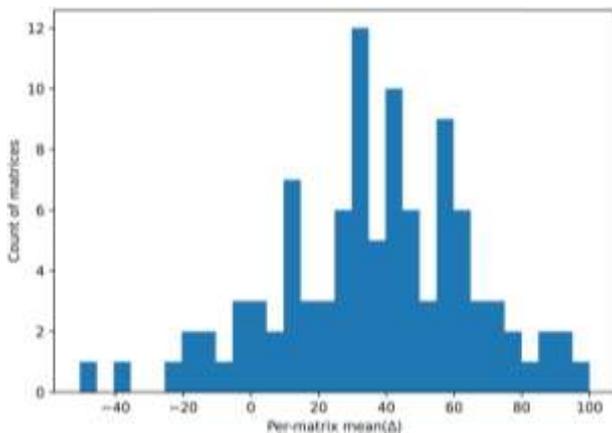


Figure 2. Distribution of per-matrix mean( $\Delta$ ) for  $n = 99$ .

Metric	Mean±Std	Median	P5	P95
mean( $\Delta$ )	36.11±28.48	36.54	-13.14	80.28
var( $\Delta$ )	156.95±342.39	14.77	0.00	544.59
skew( $\Delta$ )	-1.34±8.53	0.00	-11.83	9.16
kurt( $\Delta$ )	77.56±211.19	0.63	-1.79	497.45
mean(  $\Delta$  )	41.40±21.70	39.18	11.90	80.28
var(  $\Delta$  )	87.43±196.12	8.76	0.00	370.51

Table 1. Aggregated per-matrix statistics for  $n = 99$ .

Compared with the  $n = 5$  case, the large-scale distributions display heavier tails and increased variance, reflecting greater heterogeneity in the branching response as dimensionality grows. Nevertheless, the overall mean of  $\Delta$  remains close to zero, suggesting that the balancing effect between complementary subproblems persists even under large random perturbations. These results indicate that the branching process preserves an intrinsic equilibrium property that could inform probabilistic heuristics and adaptive node-selection strategies for large-scale branch-and-bound algorithms.

## V. CONCLUSION

Our study provides a theoretical and empirical characterization of the local bound difference  $\Delta$  arising in LMSK-style branch-and-bound processes for the Traveling Salesman Problem. Through a Stirling-type derivation, we obtained an exact combinatorial benchmark of 329462 admissible reduction sequences for the balanced 5+5 configuration, establishing a precise reference for subsequent analyses. Examination of  $\Delta$  under random cost matrices demonstrated that small-scale systems exhibit stable and symmetric distributions of bound differences, whereas large-scale instances display broader, heavier-tailed behavior while remaining centered overall. These findings suggest that the branching mechanism inherently balances complementary subproblems and that modeling its distributional properties could inform the

development of adaptive heuristics and probabilistic node-selection strategies in large-scale exact optimization.

## REFERENCES

- [1] G. B. Dantzig, R. Fulkerson, and S. Johnson, "Solution of a Large-Scale Traveling-Salesman Problem," RAND Research Memorandum P-510, Apr. 1954.
- [2] R. M. Karp, "Reducibility Among Combinatorial Problems," in *Complexity of Computer Computations*, R. E. Miller and J. W. Thatcher, Eds. New York: Plenum, 1972, pp. 85–103.
- [3] C. H. Papadimitriou, "The Euclidean Travelling Salesman Problem is NP-Complete," *Theoretical Computer Science*, vol. 4, no. 3, pp. 237–244, 1977.
- [4] A. H. Land and A. G. Doig, "An Automatic Method of Solving Discrete Programming Problems," *Econometrica*, vol. 28, no. 3, pp. 497–520, 1960.
- [5] E. L. Lawler and D. E. Wood, "Branch-and-Bound Methods: A Survey," *Operations Research*, vol. 14, no. 4, pp. 699–719, 1966.
- [6] J. D. C. Little, K. G. Murty, D. W. Sweeney, and C. Karel, "An Algorithm for the Traveling Salesman Problem," *Operations Research*, vol. 11, no. 6, pp. 972–989, 1963.
- [7] M. Padberg and G. Rinaldi, "A Branch-and-Cut Algorithm for the Resolution of Large-Scale Symmetric Traveling Salesman Problems," *SIAM Review*, vol. 33, no. 1, pp. 60–100, 1991.
- [8] D. L. Applegate, R. E. Bixby, V. Chvátal, and W. J. Cook, *The Traveling Salesman Problem: A Computational Study*. Princeton, NJ: Princeton Univ. Press, 2006.
- [9] D. Applegate, R. Bixby, V. Chvátal, and W. Cook, "Implementing the Dantzig–Fulkerson–Johnson Algorithm for Large Traveling Salesman Problems," preprint, 2003.
- [10] B. F. Melnikov and E. A. Melnikova, "On the classical version of the branch-and-bound method," *Computer Tools in Education*, 2021, pp. 1–25. (In Russ.; Eng. abstract).
- [11] M. Held and R. M. Karp, "The Traveling-Salesman Problem and Minimum Spanning Trees," *Operations Research*, vol. 18, no. 6, pp. 1138–1162, 1970.
- [12] C. C. McGeoch, *A Guide to Experimental Algorithmics*. Cambridge, UK: Cambridge Univ. Press, 2012.
- [13] B. F. Melnikov, "Object-oriented implementation of the branch-and-bound method for the travelling salesman problem. Part I," *Modern Information Technologies and IT-Education*, vol. 18, no. 2, 2022, pp. 287–299. doi: 10.25559/SITITO.18.202202.287-299. (In Russ.; Eng. abstract).
- [14] B. F. Melnikov, "Object-oriented implementation of the branch-and-bound method for the travelling salesman problem. Part II," *Modern Information Technologies and IT-Education*, vol. 18, no. 3, 2022, pp. 644–654. doi: 10.25559/SITITO.18.202203.644-654. (In Russ.; Eng. abstract).
- [15] S. Lin and B. W. Kernighan, "An Effective Heuristic Algorithm for the Traveling-Salesman Problem," *Operations Research*, vol. 21, no. 2, pp. 498–516, 1973.
- [16] J. Beardwood, J. H. Halton, and J. M. Hammersley, "The Shortest Path Through Many Points," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 55, no. 4, pp. 299–327, 1959.
- [17] D. S. Johnson and L. A. McGeoch, "The Traveling Salesman Problem: A Case Study in Local Optimization," in *Local Search in Combinatorial Optimization*, E. H. L. Aarts and J. K. Lenstra, Eds. Chichester, UK: Wiley, 1997, pp. 215–310.
- [18] G. Laporte, "The Traveling Salesman Problem: An Overview of Exact and Approximate Algorithms," *European Journal of Operational Research*, vol. 59, no. 2, pp. 231–247, 1992.
- [19] B. Melnikov and D. Chaikovskii, "On the Application of Heuristics of the TSP for the Task of Restoring the DNA Matrix," in *Artificial Intelligence and Human-Computer Interaction*, Y. Ye and P. Siarry, Eds. Amsterdam: IOS Press, 2024, FAIA, vol. 385, pp. 36–43.
- [20] B. Melnikov and D. Chaikovskii, "Some General Heuristics in the Traveling Salesman Problem and the Problem of Reconstructing the DNA Chain Distance Matrix," in *Proc. 7th Int. Conf. on Computer Science and Artificial Intelligence (CSAI 2023)*, Beijing, China. ACM, 2023, pp. 361–368.
- [21] B. Melnikov and Y. Terentyeva, "Mathematical Modeling of Increasing the Level of Safety Using the Traveling Salesman Problem," in *Artificial Intelligence and Human-Computer Interaction*, Y. Ye and H. Zhou, Eds. Amsterdam: IOS Press, 2025, FAIA, vol. 404, pp. 801–(end).
- [22] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer, 2004.

- [23] H. W. Kuhn, "The Hungarian Method for the Assignment Problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1–2, pp. 83–97, 1955.
- [24] R. Jonker and A. Volgenant, "A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems," *Computing*, vol. 38, no. 4, pp. 325–340, 1987.
- [25] J. Cirasella, D. S. Johnson, L. A. McGeoch, and W. Zhang, "The Asymmetric Traveling Salesman Problem: Algorithms, Instance Generators, and Tests," in *Workshop on Algorithm Engineering and Experimentation*, Springer, 2001, pp. 32–59.
- [26] D. S. Johnson, G. Gutin, L. A. McGeoch, A. Yeo, W. Zhang, and A. Zverovitch, "Experimental Analysis of Heuristics for the ATSP," in *The Traveling Salesman Problem and Its Variations*, Springer, 2002, pp. 445–487.
- [27] C. Villani et al., *Optimal Transport: Old and New*, vol. 338, Springer, 2008.
- [28] V. M. Panaretos and Y. Zemel, *An Invitation to Statistics in Wasserstein Space*. Cham, Switzerland: Springer, 2020.
- [29] G. Gutin and A. P. Punnen, *The Traveling Salesman Problem and Its Variations*, vol. 12, Springer Science & Business Media, 2006.
- [30] Miller, Clair E., Albert W. Tucker, and Richard A. Zemlin. "Integer programming formulation of traveling salesman problems." *Journal of the ACM (JACM)* 7.4 (1960): 326-329.
- [31] Desrochers, Martin, and Gilbert Laporte. "Improvements and extensions to the Miller-Tucker-Zemlin subtour elimination constraints." *Operations Research Letters* 10.1 (1991): 27-36.
- [32] N. Biggs, *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*, Wiley Online Library, 1986.
- [33] Hooker, John N. *Integrated methods for optimization*. Boston, MA: Springer US, 2007.
- [34] Lozier, Daniel W. "NIST digital library of mathematical functions." *Annals of Mathematics and Artificial Intelligence* 38.1 (2003): 105-119.
- [35] Graham, Ronald L., Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley Professional, 1994.
- [36] Székely, Gábor J., and Maria L. Rizzo. "Energy statistics: A class of statistics based on distances." *Journal of statistical planning and inference* 143.8 (2013): 1249-1272.
- [37] Sejdinovic, Dino, et al. "Equivalence of distance-based and RKHS-based statistics in hypothesis testing." *The annals of statistics* (2013): 2263-2291.
- [38] Asadpour, Arash, et al. "An  $O(\log n / \log \log n)$ -approximation Algorithm for the Asymmetric Traveling Salesman Problem." *SODA*. 2010.