A Framework for Comparing SATB Harmonization Approaches

Elia Pacioni
elia.pacioni@hevs.ch
HES-SO Valais-Wallis
Sierre, Switzerland
Universidad de Extremadura
Mérida, Spain

Francisco Fernández de Vega
fcofdez@unex.es
Universidad de Extremadura
Mérida, Spain

Abstract

Four-part harmonization (SATB) is a complex problem, characterized by a large search space and numerous harmonic constraints. Despite the many approaches proposed, direct comparison between results remains difficult due to the lack of a standardized evaluation framework. In this work, we propose a practical, reproducible approach based on the Sharpmony platform that integrates a rule-based correction system and an evolutionary environment for harmonizing melodies. The proposed solution enables the evaluation of different approaches using uniform criteria, thanks to its compatibility with standard score formats. As a case study, we compare a classical evolutionary approach with a modern evolutionary system that employs local search, directed mutation, and leverages data from Sharpmony students. The results show significant improvements, with error-free harmonizations and average times of about 14 minutes.

CCS Concepts

• **Computing methodologies** → **Artificial intelligence**; • **Applied computing** → **Sound and music computing**.

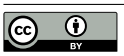
Keywords

4-part Harmonization, Genetic Algorithm, Directed Mutation, Local Search, Computational Creativity

ACM Reference Format:
Elia Pacioni and Francisco Fernández de Vega. 2026. A Framework for Comparing SATB Harmonization Approaches. In *Genetic and Evolutionary Computation Conference (GECCO Companion '26), July 13–17, 2026, San Jose, Costa Rica.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3795101.3814695

1 Introduction

Four-part harmonization (soprano, alto, tenor, bass; SATB) represents a longstanding problem in computational creativity and evolutionary computation. Despite decades of research, it remains open due to the vastness of the search space and the need to satisfy a complex set of musical rules governing harmonic progression, voice leading, and stylistic correctness.



This work is licensed under a Creative Commons Attribution 4.0 International License.
GECCO Companion '26, San Jose, Costa Rica
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2488-6/2026/07
https://doi.org/10.1145/3795101.3814695

Early work in this area demonstrated that harmonization can be formulated as an optimization problem [1]. In particular, genetic algorithms (GAs) were applied to generate harmonizations from a given melody, representing one of the first computational approaches to this problem [5]. However, these early systems were generally limited to simplified contexts and required considerable computational effort even for short musical excerpts. Subsequently, a variety of approaches within artificial intelligence have been explored, including symbolic rule-based methods, probabilistic models, and machine learning techniques [3, 4].

More recent research has focused on improving evolutionary approaches by integrating domain knowledge into the search process. In particular, techniques such as directed mutation [8], local search [6], and fitness precalculation have been proposed to reduce convergence time and improve solution quality. In parallel, several studies have highlighted the role of human-derived information as a means to constrain the search space and guide exploration toward more promising regions [9, 7, 6]. These approaches have led to significant improvements in both solution quality and computational performance. Despite these advances, a fundamental limitation remains: results reported in the literature are difficult to compare due to the lack of a standardized evaluation framework. Different studies rely on varying rule sets, datasets, and experimental conditions, making direct comparison between approaches unclear.

In this work, we take a first step toward addressing this issue by introducing a practical and reproducible evaluation framework for SATB harmonization based on the Sharpmony platform. Our approach enables the correction of musical scores generated by different methods within a unified, rule-based system developed in collaboration with music conservatory teachers. This allows for direct, uniform comparison across approaches. Furthermore, we show that evolutionary algorithms (EAs) can be effectively applied to this problem by framing it as a search task in the solution space. As a case study, we revisit early GA approaches and compare them with a modern evolutionary system. Finally, we outline how this approach can be extended to construct a broader benchmark for harmonization research, enabling systematic comparisons between evolutionary and non-evolutionary methods. We also present the Sharpmony evolutionary platform as an experimental environment that supports batch execution, parameter control, and detailed result analysis.

The paper is structured as follows. Section 2 provides an overview of the existing literature on the four-part harmonization problem. Section 3 presents the Sharpmony framework and how it is used in the research. Section 4 describes the methodological approaches adopted and the experimental setup. Section 5 reports and analyzes

the findings, while Section 6 offers the final conclusions of the study.

2 State of the Art

SATB harmonization originates in the compositional practice of the Baroque period, particularly in the works of Johann Sebastian Bach, which contributed to the formalization of the main harmonic and voice-leading rules still used today as theoretical and pedagogical references.

In the computational domain, the challenge of SATB harmonization has been addressed using different approaches. These approaches can be broadly classified into rule-based symbolic methods, probabilistic models, machine learning techniques, and evolutionary approaches.

Symbolic approaches explicitly model harmonic rules, ensuring a high level of stylistic correctness, but they tend to be rigid and lack scalability as the number of constraints increases. Probabilistic models, such as those based on Markov chains, are effective in predicting musical sequences from known contexts, but they do not systematically enforce harmonic rules. Similarly, machine learning techniques, in particular neural networks, have shown strong performance on imitation tasks but are less suited to handling complex explicit constraints.

Evolutionary approaches have been applied to the SATB harmonization problem, as they enable formulating harmonization as an optimization task. Early work in this direction demonstrated the approach's feasibility but also highlighted its scalability limitations. In particular, the approach proposed by McIntyre [5] harmonizes a given melody in a fixed key, but requires large populations and many generations even for short melodies, illustrating the computational cost of the problem.

Subsequent work has aimed to better structure the compositional process. For example, Kaliakatsos et al. [2] propose separating the generation of the chord progression from the assignment of notes to voices, reducing the problem's complexity but introducing additional decisions regarding chord selection. Similarly, EvoComposer adopts a multi-objective formulation and a hybrid evaluation function that combines music-theory rules with weights derived from statistical analysis of Bach chorales, thereby improving musical realism while remaining dependent on the chosen model.

More recently, research has focused on improving evolutionary approaches by integrating domain knowledge and strategies to guide search-space exploration. These include techniques such as directed mutation, local search, and fitness precomputation, which aim to reduce the number of non-promising solutions explored during the evolutionary process. These studies show that the main bottleneck lies not only in the problem representation but also in the interaction among search space size, the number of rules, and evaluation cost.

A further development involves integrating data derived from real teaching and learning processes. The analysis of harmonizations produced by students [9] shows that the portion of the search space explored in practice is only a small fraction of the theoretical one, suggesting that such data can guide the algorithm toward more relevant regions of the solution space.

Despite these advances, a fundamental limitation remains: different approaches are evaluated under heterogeneous experimental conditions, including different rule sets, representations, and input data. This makes direct comparison difficult and prevents an objective assessment of improvements. This work addresses this issue by taking a first step toward a uniform and model-independent evaluation framework.

3 Sharpmony Framework

Sharpmony is a web-based and Android application designed to support the teaching and learning of 4-part harmonization. The system is currently in production and actively used by students, independent users, and music conservatory teachers, with 4,840 registered users, 24 conservatories involved, and more than 21,000 exercises processed by the platform. At its core, Sharpmony implements a rule-based system derived from traditional harmony teaching. The platform incorporates more than 50 harmonic rules [1] and exceptions, covering aspects such as chord progression, voice leading, spacing, and stylistic constraints. These rules are formally defined and serve as the basis for both correcting student exercises and evolving scores.

Sharpmony also collects harmonization exercises created by users, which are then analyzed to extract chord pairs and identify any associated errors. This information is used to build a comprehensive database of chord pairs, capturing for each pair the key, tonality, degree information, as well as the number and types of harmonic errors detected. The initial database was built from the analysis of about 13,000 student exercises and contained 67,240 chord pairs, representing less than 1% of the possible chord-pair space in a single key. This initial subset corresponds to the exercises analyzed for database construction at that stage, while the current platform statistics reflect the larger number of exercises processed by Sharpmony over time. Since then, the database has continued to grow through the use of the evolutionary platform, which adds newly explored chord pairs during search. At present, it contains 168,897 chord pairs, of which 48,726 are error-free, and 142,051 contain fewer than four errors. The database is used to guide the search process. The algorithm searches among chord pairs already stored in the database and prioritizes those that are compatible with the current melodic note and produce fewer local errors. If no suitable pair is found, the algorithm can generate and evaluate a new chord pair from the full search space and then store the resulting information for future use. In this way, the database acts both as a knowledge base extracted from real practice and as an evolving repository enriched during the search process.

In addition to the correction system, a dedicated evolutionary platform has been developed within the Sharpmony ecosystem. This platform, currently available for internal use, allows the execution of evolutionary processes for harmonization tasks. As shown in Figure 1, it provides a configurable interface that allows parameters such as the number of generations, population size, mutation probability, and crossover probability to be defined.

The system includes a job manager that handles multiple executions and supports parallel processing. During execution, users

[1] See the haromonic manual: https://sharpmony.unex.es/index.php?r=tutorials%2Fharmonic-manual

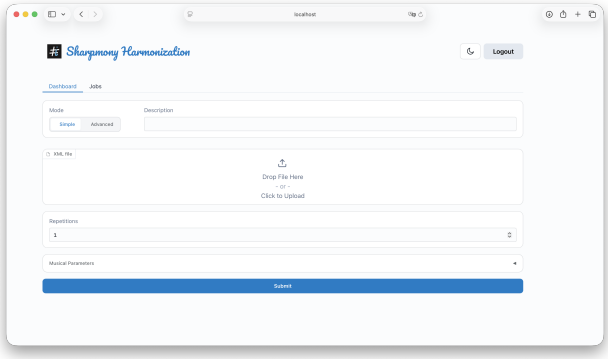


Figure 1: Sharpmony dashboard used to run the EA, displaying configurable parameters and enabling the selection of harmonic rules to be included in the fitness evaluation.

can monitor the progress of each run in real time, including the evolution of fitness values across generations. The platform also provides tools for analyzing results, including the computation of mean and standard deviation across multiple runs and the comparison of different configurations. This combination makes Sharpmony a suitable framework for systematically comparing different SATB harmonization approaches under uniform conditions.

4 Methodology and Experimental Setup

The experiments are based on the melody (see Figure 2) proposed and harmonized by McIntyre [5]. Using the same musical input allows the results to be evaluated under comparable conditions.

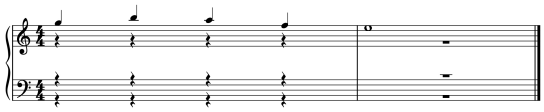


Figure 2: Melody selected by McIntyre and used for the comparative analysis.

All experiments are executed using the Sharpmony evolutionary platform (Section 3). To evaluate the stability of the proposed approach, 10 independent runs are performed using the same parameter settings. The results are aggregated using mean and standard deviation.

The methodology is based on two consecutive EAs. The first algorithm evolves the chord progression associated with the input melody. The second algorithm generates the complete SATB harmonization by assigning notes to the remaining voices. A description of the chromosome is available in [7]. In the experiments, the first EA runs for 100 generations. The second one is configured with 100 generations, a population of 8 individuals, a mutation probability of 30%, and a crossover probability of 40%. The parameter configuration follows the setup validated in our previous studies [8, 7], where it was tuned through preliminary experimentation.

The search process is guided in three ways. First, directed mutation is used so that mutation is applied to chord positions associated with errors. Second, mutation relies on a database of chord pairs derived from Sharpmony student exercises. Third, local search is integrated throughout the evolutionary process. While in previous approaches it is applied only within directed mutation, in the present configuration it is used from the beginning of the evolutionary process.

Each run produces a complete harmonization, which is evaluated using the Sharpmony rule-based correction system. The evaluation is based on the number of harmonic errors detected in the final score, considering the full set of more than 50 rules and exceptions. To slightly reduce the rigidity of the search, and since the original melody contains a range error, the algorithm is allowed to modify the octave of the melody notes, without significantly changing the melody itself.

5 Results

Figure 3 shows the harmonization produced by McIntyre's algorithm. Two types of errors are detected in the score: voice range violations (highlighted in brown) and wrong note duplications (highlighted in fuchsia). These errors indicate that the generated harmonization does not satisfy all the harmonic constraints considered in this work.

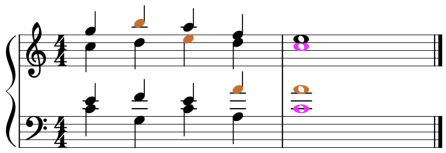


Figure 3: Output produced by McIntyre's algorithm; colored notes indicate errors identified through correction based on Sharpmony's harmonic rules.

In contrast, all runs of the proposed approach converge to solutions with no detected errors. As shown in Figure 4, convergence is reached after approximately 12 generations, even though the algorithm was configured to run for 100 generations. Building on these results, Figure 5 shows an example of harmonization produced by the proposed method using the same melody McIntyre adopted. The resulting score is error-free, i.e., no violations are detected by Sharpmony with respect to the 50+ harmonic rules formalized from traditional harmony teaching.

Table 1 summarizes the quantitative results across the 10 runs. The average execution time is approximately 13 minutes, with some variability across runs. The values reported for Found and New, respectively, are the number of chord transitions retrieved from the student-based database and those generated by exploring the full search space. These results show that a significant portion of the search is guided by the database, while still allowing the discovery of new configurations.

Compared to previous work (e.g., [7, 6]) , the execution time is significantly reduced. In earlier approaches, convergence required substantially longer runtimes, whereas in the present configuration, solutions are obtained in a few minutes. This improvement can be attributed to the integration of local search throughout the evolutionary process, which enables faster refinement of candidate

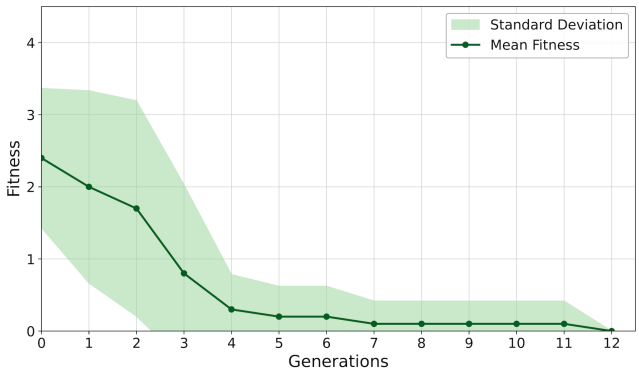


Figure 4: Average fitness across multiple runs of the proposed EA, with standard deviation.

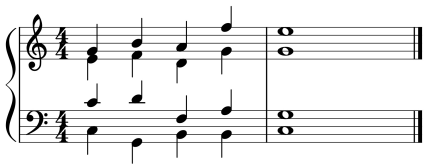


Figure 5: Example of a result produced by the evolutionary process.

Metric	Mean	Std. Dev.
Time	13:47	7:24
Found	859.4	365.18
New	70.4	48.51

Table 1: Summary of execution times, chords found in the database (Found), and those from the general search space (New).

solutions and reduces the number of generations required to reach convergence, and to the redesign of the kernel, where revising the computation of degree and inversion has led to a significant speedup.

To assess generalizability, the approach has also been tested on a larger musical example, an 8-bar score (Figure 6). In this case, the algorithm also produces a complete harmonization with no detected errors, suggesting that the method generalizes beyond short input melodies.

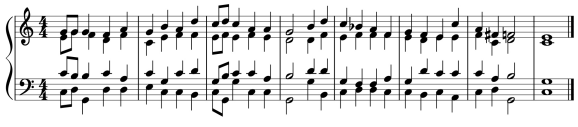


Figure 6: Example of a result produced by the evolutionary process with a bigger score.

6 Conclusion

This work introduces a practical, reproducible framework for evaluating SATB harmonization approaches using the Sharpmony platform. The experimental results show that the proposed evolutionary approach consistently produces valid harmonizations under the considered rule set, outperforming classical evolutionary methods under the same conditions. The results also indicate that EAs can effectively address both simple and more complex harmonization tasks, as demonstrated by experiments on melodies of varying lengths. The approach also demonstrates improved computational efficiency, producing valid solutions within reasonable execution times and requiring less computational effort than previous methods. Beyond the specific results, this work highlights the importance of consistent evaluation in the SATB harmonization domain. The lack of uniform evaluation criteria has historically limited the comparability of different approaches. The proposed framework provides a first step toward addressing this issue by enabling reproducible and comparable assessments. Future work will extend the experimental comparison to a broader set of approaches from the literature, including both evolutionary and non-evolutionary methods, to establish a shared benchmark for systematic evaluation.

Acknowledgments

This work was partially supported by the Spanish Ministry of Economy and Competitiveness (PID2020-115570GB-C21, PID2023-147409NB-C22), funded by MCIN/AEI/10.13039/501100011033, and the Junta de Extremadura (GR24142).

References

[1] Andrew Horner and David E Goldberg. 1991. *Genetic algorithms and computer-assisted music composition.* Vol. 51. Ann Arbor, MI: Michigan Publishing, University of Michigan Library.

[2] Maximos Kaliakatsos-Papakostas and Emilios Cambouropoulos. 2014. Probabilistic harmonization with fixed intermediate chord constraints. In *ICMC*.

[3] Chien-Hung Liu and Chuan-Kang Ting. 2016. Computational intelligence in music composition: a survey. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 1, 1, 2–15.

[4] Omar Lopez-Rincon, Oleg Starostenko, and Gerardo Ayala-San Martín. 2018. Algoritmic music composition based on artificial intelligence: a survey. In *2018 international conference on electronics, communications and computers (CONIELECOMP)*. IEEE, 187–193.

[5] Ryan A McIntyre. 1994. Bach in a box: the evolution of four part baroque harmony using the genetic algorithm. In *Proceedings of the first ieee conference on evolutionary computation. ieee world congress on computational intelligence*. IEEE, 852–857.

[6] Elia Pacioni and Francisco Fernández De Vega. 2025. Combining local search and directed mutation in evolutionary approaches to 4-part harmony. In *Artificial Intelligence in Music, Sound, Art and Design*. Penousal Machado, Colin Johnson, and Iria Santos, (Eds.) Springer Nature Switzerland, Cham, 141–155. ISBN: 978-3-031-90167-6.

[7] Elia Pacioni and Francisco Fernández De Vega. 2025. Advancing 4-part evolutionary harmony through analysis of human–machine approaches to teaching–learning. *Complexity*, 2025, 1, 3086287. doi:https://doi.org/10.1155/cplx/3086287.

[8] Elia Pacioni and Francisco Fernández De Vega. 2024. On the impact of directed mutation applied to evolutionary 4-part harmony models. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*. Springer, 311–325.

[9] Elia Pacioni and Francisco Fernández De Vega. 2024. Paving the way towards evolutionary machine teaching: an application to 4-part harmony. In *International Conference on Artificial Evolution (Evolution Artificielle)*. Springer, 61–76.