# Improving Ranking Quality and Fairness in Swiss-System Chess Tournaments

PASCAL FÜHRLICH, Hasso Plattner Institute, University of Potsdam, Germany
ÁGNES CSEH, Institute of Economics, Centre for Economic and Regional Studies, Hungary
PASCAL LENZNER, Hasso Plattner Institute, University of Potsdam, Germany

The International Chess Federation (FIDE) imposes a voluminous and complex set of player pairing criteria in Swiss-system chess tournaments and endorses computer programs that are able to calculate the prescribed pairings. The purpose of these formalities is to ensure that players are paired fairly during the tournament and that the final ranking corresponds to the players' true strength order. We contest the official FIDE player pairing routine by presenting alternative pairing rules. These can be enforced by computing maximum weight matchings in a carefully designed graph. We demonstrate by extensive experiments that a tournament format using our mechanism (1) yields fairer pairings in the rounds of the tournament and (2) produces a final ranking that reflects the players' true strengths better than the state-of-the-art FIDE pairing system.
The full version [3] of this extended abstract can be found here: https://arxiv.org/abs/2112.10522.

CCS Concepts: • **Theory of computation** → **Algorithmic mechanism design**.

Additional Key Words and Phrases: Swiss system, tournament, ranking, fairness, weighted matching

**ACM Reference Format:**
Pascal Führlich, Ágnes Cseh, and Pascal Lenzner. 2022. Improving Ranking Quality and Fairness in Swiss-System Chess Tournaments. In *Proceedings of the 23rd ACM Conference on Economics and Computation (EC '22), July 11–15, 2022, Boulder, CO, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3490486.3538298

The Swiss-system tournament format is widely used in competitive games like most e-sports, badminton, and chess, the last of which this paper focuses on. In such tournaments, the number of rounds is predefined, but the pairing of players in each of these rounds depends on the results of previous rounds. Designing player pairings for Swiss-system tournaments is a challenging combinatorial optimization problem. In Swiss-system chess tournaments, there are two well-defined and rigid *absolute* and two milder *quality* pairing criteria prescribed by FIDE [2]:

(A1) No two players play against each other more than once.
(A2) In each round before the last one, the difference of matches played with white and matches played with black pieces is between $-2$ and 2 for every player.
(Q1) Opponents have equal or similar score.
(Q2) Each player has a balanced color distribution.

Criterion (A1) ensures variety, while criterion (A2) ensures fairness, since the player with white pieces starts the game, and thus has an advantage over her opponent. These absolute criteria must be obeyed at any cost, which often enforces the relaxation of the two quality criteria (Q1) and (Q2).


*EC '22, July 11–15, 2022, Boulder, CO, USA*

ACM ISBN 978-1-4503-9150-4/22/07.
https://doi.org/10.1145/3490486.3538298

We present a novel algorithm-based mechanism for calculating pairings in Swiss-system chess tournaments. As our pairings deviate from the ones prescribed by FIDE, they contrast the state-of-the-art declarative approach and contest Dutch BBP, the FIDE-endorsed implementation [1]. We compare the two systems in terms of ranking quality, number of floaters, and color balance quality.

(1) Using an extensible and easy-to-understand approach based on maximum weight matchings, we implemented the standard pairing systems (these pair players within the same score group) Dutch, Burstein, and Monrad and invented the new systems Random and Random2.
(2) The pairing systems in descending order by expected **ranking quality** are: Burstein > Random2 > Dutch = Dutch BBP > Random > Monrad. In particular, our implementation Dutch yields similar ranking quality as the one reached by the Dutch BBP pairing engine.
(3) We utilize our weighted matching model to define a novel measure called *normalized strength difference*, which we identify as the main reason for a good ranking quality. This also explains why our approaches Burstein and Random2 outperform Dutch BBP.
(4) The pairing systems in ascending order by expected **number of floaters** are: Burstein < Random2 = Dutch = Monrad < Dutch BBP < Random. Compared to Dutch BBP, our mechanism is fairer in terms of matching more players within their own score group.
(5) All our pairing systems ensure the same **color balance quality** as Dutch BBP, with Random even reaching a better color balance. Moreover, we show that our approach can easily be modified to enforce an even stronger color balance. This does not significantly affect the ranking quality—only the number of floaters increases slightly.

As the above points demonstrate, our implementations Burstein and Random2 either outperform or are on a par with Dutch BBP. Our implementation Dutch yields pairings that perform just as well or even better than the ones prescribed by the FIDE-endorsed Dutch BBP.

*Discussion and Outlook:* The experimental results for our matching-based approaches Burstein and Random2 demonstrate that it is possible to outperform the state-of-the-art FIDE pairing criteria in terms of both ranking quality and fairness. A clear advantage of our algorithmic approach is that it is easily extendable: as Random and Random2 already demonstrate, a new pairing system can be implemented simply by specifying how edge weights are calculated. Similarly, as we have also shown, the color balance can be adjusted by simply changing a parameter. Alternatively, we can reach an alternating black-white sequence for each player instead of just minimizing the color difference in each round. Also, the flexibility of the maximum weight matching model proved to be essential for uncovering the driving force behind the achieved high ranking quality: the normalized strength difference. Hence, our approach was not only valuable for computing better pairings but also in the analysis of the obtained ranking quality. Furthermore, the flexibility of our matching-based approach likely allows to incorporate additional quality criteria like measuring fairness via the average opponent ratings. Last but not least, other fields using rankings derived from pairwise comparisons might also benefit from our work. Approaches similar to ours might be applied for organizing speed-dating type events or for computing trainee rotation schedules.

## ACKNOWLEDGMENTS

Cseh was supported by OTKA grant K128611 and the János Bolyai Research Fellowship.

## REFERENCES

[1] Jeremy Bierema. 2017. BBP Pairings, a Swiss-system chess tournament engine. https://github.com/BieremaBoyzProgramming/bbpPairings. Accessed: 2022-05-17.
[2] FIDE. 2020. FIDE Handbook. https://handbook.fide.com/. Accessed: 2022-05-17.
[3] Pascal Führlich, Ágnes Cseh, and Pascal Lenzner. 2021. Improving Ranking Quality and Fairness in Swiss-System Chess Tournaments. https://doi.org/10.48550/ARXIV.2112.10522