A Brief Tour in Twin-Width

Stéphan Thomassé ✉

Univ Lyon, CNRS, ENS de Lyon, Université Claude Bernard Lyon 1, LIP UMR5668, France

Abstract

This is an introduction to the notion of twin-width, with emphasis on how it interacts with first-order model checking and enumerative combinatorics. Even though approximating twin-width remains a challenge in general graphs, it is now well understood for ordered graphs, where bounded twin-width coincides with many other complexity gaps. For instance classes of graphs with linear FO-model checking, small classes, or NIP classes are exactly bounded twin-width classes. Some other applications of twin-width are also presented.

2012 ACM Subject Classification Theory of computation → Finite Model Theory; Theory of computation → Parameterized complexity and exact algorithms

Keywords and phrases Twin-width, matrices, ordered graphs, enumerative combinatorics, model theory, algorithms, computational complexity, Ramsey theory

Digital Object Identifier 10.4230/LIPIcs.ICALP.2022.6

Category Invited Talk

1 Introduction

One of the most natural ways to understand discrete structures is to measure their complexity. A reasonable expectation is that the class of structures with bounded measure is not too difficult to understand and manipulate. Ideally, bounded measure classes should be simple with respect to several points of view such as "computationally hard problems can be solved efficiently" or "the number of structures of size n is a small function of n" and should also enjoy some stability like "the measure should stay bounded if one performs moderate deterministic changes to the structures".

The first difficulty to provide a general purpose complexity measures on graphs is that it quickly boils down to the basic question: What is a simple 01-matrix? Fortunately this problem has already been addressed long ago and the answer is very simple: a matrix M is complex if it contains all small matrices up to some (large) size. Consequently, the complexity measure $vc(M)$ could be the maximum k for which all 01-valued $k \times k$-matrices appear in M. This is equivalent to the well-known Vapnik-Cervonenkis dimension and indeed classes of matrices with bounded VC-dimension have moderate growth $O(2^{n^{2-\varepsilon}})$ and some problems are computationally easier (for instance the minimum hitting set problem can be approximated). Interestingly, these two properties characterize bounded VC-dimension for classes of matrices closed under submatrices. This is the bounded/unbounded VC-dimension gap, which is (arguably) the first question one should ask when investigating a class of structures.

The exact same idea can be used to measure the complexity of a permutation matrix M (exactly one 1 per row and per column): let $mt(M)$ be the maximum k for which all $k \times k$ permutation matrices appear in M. Marcus and Tardos [17], proving the Stanley-Wilf conjecture, showed that the growth of a class of permutation matrices with bounded measure mt is c^n. Using their method, Guillemot and Marx [14] showed that checking if a fixed $k \times k$ permutation matrix F is contained in an $n \times n$ permutation matrix P can be done in linear time $f(k)n$ (when both coded as permutations). Their breakthrough method was a completely new win/win scheme: they showed that unless one can detect F in P, then P can be iteratively contracted in linear time and the contraction scheme allows *then* to test if indeed F is a subpermutation of P. Note that permutation matrices are ordered

49th International Colloquium on Automata, Languages, and Programming (ICALP 2022).
Editors: Mikołaj Bojańczyk, Emanuela Merelli, and David P. Woodruff;
Article No. 6; pp. 6:1–6:5
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

matrices, where rows and columns are linearly ordered. We showed in Twin-width IV [9] that the natural generalization of the *mt* parameter to general ordered matrices is the following: a matrix M has *grid rank* k if this is the maximum value for which M has a $k \times k$ block division in which every block has rank at least k.

Guillemot and Marx concluded their paper by asking if their technique for permutations could be generalized for graphs. This was the goal of our paper Twin-width I [12] where we defined the twin-width of a graph G as the minimum degree of error in a contraction sequence of G (we iteratively contract pairs of vertices, two contracted groups of vertices forming an error edge if there is both an edge and a non edge between them). Precisely, the twin-width $tww(G)$ of a graph G on n vertices is the minimum k such that: there exists a sequence of partitions $P_n, \ldots, P_1$ of $V(G)$ where each P_{i-1} is obtained from P_i by merging two parts, and such that for every part X in P_j, the number of parts Y in P_j which are not homogeneous with X is at most k. Here two disjoint sets X, Y are *homogeneous* if the relation between $x \in X$ and $y \in Y$ does not depend of the choices of x, y (therefore homogeneity, and hence twin-width, is also defined for binary multirelations). So bounded twin-width corresponds to maximum degree in every *error graph* G_j which vertices are the parts of P_j and edges are the non homogeneous pairs. If we impose further that all components of graphs G_j have bounded size, we have shown in Twin-width VI [10] that the obtained parameter is equivalent to rank-width. Hence twin-width generalizes rank-width but also captures strict minor closed classes, or strict permutation graphs.

Mimicking Guillemot and Marx argument for permutations, it is not hard to show that if one has access to such a sequence $P_n, \ldots, P_1$ certifying twin-width k, then one can test if some fixed graph H of size t is an induced subgraph of G in linear time $f(k,t).n$. One of the main result of Twin-width I is that we can moreover test any FO-formula of depth t in time $f(k,t).n$. Also, generalizing Marcus-Tardos' result on permutations, we could prove in Twin-width II [6] that the number of (labelled) graphs of size n with twin-width bounded by some constant is at most $c^n.n!$ (we call this a *small class*). This result implies in particular that the class of (sub)cubic graphs (degree at most 3) does not have bounded twin-width, since it is not small. But so far we have no "deterministic" construction of a cubic graph with arbitrarily high twin-width. One can naturally wonder if these two implications (easyness of FO-model checking and small property) could be equivalent to bounded twin-width.

This is unfortunately not the case: since FO-model checking can be solved in linear time on bounded degree graphs, there are classes of graphs on which FO-model checking is tractable and for which twin-width is unbounded. From the counting point of view, we conjectured in [6] the equivalence between bounded twin-width and being a small class. Sadly again, we could prove in Twin-width VII [8] that there are (countable) Cayley graphs of finitely generated groups with unbounded twin-width, while any such Cayley graph defines a small class. Thus bounded twin-width for general graphs does not seem to be equivalent to some computational complexity class, nor it seems to be definable via counting. Nevertheless, bounded twin-width is a particularly stable notion since every first order interpretation of a bounded twin-width class has also bounded twin-width. For instance squares of planar graphs have bounded twin-width.

We still have a very limited understanding of twin-width, and especially for bounded degree graphs: not only we do not have an algorithm to approximate it, but we do not know what could be a certificate of high twin-width, and we are not even able to construct by hand a cubic graph of high twin-width. So why twin-width is so hard to handle, given that it enjoys so many nice properties? The answer is that twin-width indeed corresponds to a crucial complexity gap, but for ordered graphs (a binary birelation consisting of a graph and a linear order on its vertices) rather than for graphs.

Indeed, in Twin-width IV [9] we could show that for classes of ordered graphs, bounded twin-width, linear FO-model checking and being a small class are equivalent. Moreover these three characterizations are in turn equivalent to the fact that the adjacency matrices of the graphs, ordered by their linear order, have bounded grid-rank. Finally, approximating twin-width for ordered graphs can be done in polynomial time. In particular, the bounded/unbounded twin-width gap for ordered matrices is as fundamental as the one of VC-dimension as it has many equivalent formulations coming from other domains. For instance, for ordered graphs, the NIP property in model theory coincides with bounded twin-width. Since any graph G with twin-width k has a linear order L for which (G, L) has twin-width k, the difficulty of twin-width for general graphs seems to come from the fact that we have lost the information encoded in the linear order.

Hence an appealing strategy to show that a graph G has (reasonably) bounded twin-width is to be able to guess a suitable linear order T on its vertices since we have the machinery to efficiently approximate the twin-width of the ordered graph (G, T). For instance, we could prove that minor closed classes have bounded twin-width by using a Lex-DFS to provide the linear order. We can also take advantage of the stability of twin-width by FO-interpretation. Let us illustrate this on some example: Assume that we want to approximate the twin-width of a bipartite graph G with bipartition A, B in which B is linearly ordered by $<$. We would not have a clue if B would not have been ordered, and we can directly conclude if both A and B are ordered, so what about this "semi-ordered" case? The answer is quite easy: associate to each vertex $a \in A$ the characteristic 01 vector of its neighbors in B, ordered by $<$, and sort A by lexicographic order. This is a first-order interpretation, so the order we find on A cannot increase twin-width too much. Hence now A is ordered, and twin-width can be approximated. Note that if several vertices of A have the same neighborhood, we can pairwise contract them since this does not affect twin-width. This is probably the best advice to try to compute twin-width: find an order. Is there a general algorithm to find it?

To conclude, let us observe that as a way to characterize simple matrices, twin-width is a very general tool which can apply to many topics. For instance bounded twin-width is a group invariant and finitely generated groups can have either bounded or unbounded twin-width (but again we have no explicit presentation of any unbounded twin-width group). Another field where matrices are central is linear programming. We have showed in Twin-width III [7] that when a matrix has bounded twin-width, there is a constant duality gap between minimum hitting set and maximum packing (while bounded VC-dimension only bridges the fractional gap for hitting set). We provide in the references a list of recent publications with better bounds on twin-width of classes ([1, 2, 3, 15, 18]), on computing twin-width ([4, 5]), on using twin-width for algorithms ([11, 16]) and for data-structures [20], and more topics ([13, 19, 21, 22]) that we cannot unfortunately cover here. We believe that there are many other aspects of twin-width waiting to be discovered.

References

1 Jungho Ahn, Kevin Hendrey, Donggyu Kim, and Sang-il Oum. Bounds for the twin-width of graphs. *CoRR*, abs/2110.03957, 2021. arXiv:2110.03957.

2 Jakub Balabán and Petr Hlinený. Twin-width is linear in the poset width. In Petr A. Golovach and Meirav Zehavi, editors, *16th International Symposium on Parameterized and Exact Computation, IPEC 2021, September 8-10, 2021, Lisbon, Portugal*, volume 214 of *LIPIcs*, pages 6:1–6:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.IPEC.2021.6.

3 Jakub Balabán, Petr Hlinený, and Jan Jedelský. Twin-width and transductions of proper k-mixed-thin graphs. *CoRR*, abs/2202.12536, 2022. `arXiv:2202.12536`.

4 Pierre Bergé, Édouard Bonnet, and Hugues Déprés. Deciding twin-width at most 4 is NP-complete. *CoRR*, abs/2112.08953, 2021. `arXiv:2112.08953`.

5 Édouard Bonnet, Dibyayan Chakraborty, Eun Jung Kim, Noleen Köhler, Raul Lopes, and Stéphan Thomassé. Twin-width VIII: delineation and win-wins. *CoRR*, abs/2204.00722, 2022. `doi:10.48550/arXiv.2204.00722`.

6 Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width II: small classes. In Dániel Marx, editor, *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 1977–1996. SIAM, 2021. `doi:10.1137/1.9781611976465.118`.

7 Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width III: max independent set, min dominating set, and coloring. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 35:1–35:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. `doi:10.4230/LIPIcs.ICALP.2021.35`.

8 Édouard Bonnet, Colin Geniet, Romain Tessera, and Stéphan Thomassé. Twin-width VII: groups. *CoRR*, abs/2204.12330, 2022. `doi:10.48550/arXiv.2204.12330`.

9 Édouard Bonnet, Ugo Giocanti, Patrice Ossona de Mendez, Pierre Simon, Stéphan Thomassé, and Szymon Torunczyk. Twin-width IV: ordered graphs and matrices. *CoRR*, abs/2102.03117, 2021. `arXiv:2102.03117`.

10 Édouard Bonnet, Eun Jung Kim, Amadeus Reinald, and Stéphan Thomassé. Twin-width VI: the lens of contraction sequences. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 1036–1056. SIAM, 2022. `doi:10.1137/1.9781611977073.45`.

11 Édouard Bonnet, Eun Jung Kim, Amadeus Reinald, Stéphan Thomassé, and Rémi Watrigant. Twin-width and polynomial kernels. In Petr A. Golovach and Meirav Zehavi, editors, *16th International Symposium on Parameterized and Exact Computation, IPEC 2021, September 8-10, 2021, Lisbon, Portugal*, volume 214 of *LIPIcs*, pages 10:1–10:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. `doi:10.4230/LIPIcs.IPEC.2021.10`.

12 Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. *J. ACM*, 69(1):3:1–3:46, 2022. `doi:10.1145/3486655`.

13 Édouard Bonnet, O-joung Kwon, and David R. Wood. Reduced bandwidth: a qualitative strengthening of twin-width in minor-closed classes (and beyond). *CoRR*, abs/2202.11858, 2022. `arXiv:2202.11858`.

14 Sylvain Guillemot and Dániel Marx. Finding small patterns in permutations in linear time. In Chandra Chekuri, editor, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 82–101. SIAM, 2014. `doi:10.1137/1.9781611973402.7`.

15 Hugo Jacob and Marcin Pilipczuk. Bounding twin-width for bounded-treewidth graphs, planar graphs, and bipartite graphs. *CoRR*, abs/2201.09749, 2022. `arXiv:2201.09749`.

16 Stefan Kratsch, Florian Nelles, and Alexandre Simon. On triangle counting parameterized by twin-width. *CoRR*, abs/2202.06708, 2022. `arXiv:2202.06708`.

17 Adam Marcus and Gábor Tardos. Excluded permutation matrices and the stanley-wilf conjecture. *J. Comb. Theory, Ser. A*, 107(1):153–160, 2004. `doi:10.1016/j.jcta.2004.04.002`.

18 William Pettersson and John Sylvester. Bounds on the twin-width of product graphs. *CoRR*, abs/2202.11556, 2022. `arXiv:2202.11556`.

19 Michal Pilipczuk and Marek Sokolowski. Graphs of bounded twin-width are quasi-polynomially χ-bounded. *CoRR*, abs/2202.07608, 2022. `arXiv:2202.07608`.

20 Michal Pilipczuk, Marek Sokolowski, and Anna Zych-Pawlewicz. Compact representation for matrices of bounded twin-width. In Petra Berenbrink and Benjamin Monmege, editors, *39th International Symposium on Theoretical Aspects of Computer Science, STACS 2022, March 15-18, 2022, Marseille, France (Virtual Conference)*, volume 219 of *LIPIcs*, pages 52:1–52:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.STACS.2022.52.

21 Wojciech Przybyszewski. VC-density and abstract cell decomposition for edge relation in graphs of bounded twin-width. *CoRR*, abs/2202.04006, 2022. arXiv:2202.04006.

22 André Schidler and Stefan Szeider. A SAT approach to twin-width. *CoRR*, abs/2110.06146, 2021. arXiv:2110.06146.