#### DISCUSSION

# Comments on: Recent progress on the combinatorial diameter of polytopes and simplicial complexes

Jesús A. De Loera

Published online: 1 October 2013

© Sociedad de Estadística e Investigación Operativa 2013

#### 1 Introduction

Linear programming is the workhorse of modern optimization, but despite its importance, we still face plenty of open mathematical questions about it. Dantzig's simplex method is certainly one of the most popular algorithms for solving linear programs. Its intuitive geometry is quite simple. The simplex method searches the graph of the polyhedron, from a vertex of the one-skeleton to a better neighboring one according to some *pivot rule*, which selects an improving neighbor. Geometrically, the simplex method traces a path on the graph of the polytope. The *diameter* of the graph of a polytope is the length of the longest shortest path among all possible pairs of vertices.

The historic results of Francisco Santos disproving the Hirsch bound on the diameter (Santos 2012), mentioned in the main article (see Sect. 4 of Santos 2013), have dramatically advanced the understanding of the geometry of the simplex method and motivated new active research. Santos' present article in TOP is a wonderful snapshot of what we know today about the diameter problem and the power of the abstract point of view on dealing with optimization problems (which recently has seen active work, see De Loera et al. (2013) and the many references therein). Today, we still do not know exact bounds for the growth of the diameter; Santos lower-bound is far from the best upper bounds discussed in Sect. 3 of Santos 2013. Here, I would like to make just three comments about his excellent article.

The author was partially supported by NSF grant DMS-0914107.

This comment refers to the invited paper available at doi:10.1007/s11750-013-0295-7.

J.A. De Loera (\simega)

Department of Mathematics, University of California, Davis, USA

e-mail: deloera@math.ucdavis.edu



## Comment 1: There is also very rich geometry in other algorithms for linear programming, with plenty of open questions waiting for us, also

Indeed, there are plenty of alternatives to the simplex method. There is the Ellipsoid method of Hačijan (1979), the first polynomial time algorithm for linear programming. There is the simple and beautiful Fourier–Motzkin elimination process, which can be used to prove the duality theory of linear programs. There is the Criss-Cross method, pivoting methods that are allowed to go out of the feasible region, walking not on the graph of it the polytope, but rather on the graph of the hyperplane arrangement defining it (Fukuda and Terlaky 1997). There is the family of Relaxation methods based on linear projections (Agmon 1954; Motzkin and Schoenberg 1954; Betke and Gritzmann 1992; Betke 2004; Goffin 1982; Telgen 1982) and many many other methods of solution and analysis (see, e.g., Bertsimas and Vempala 2010; Borgwardt 2009; Chubanov 2012; Dunagan and Vempala 2008; Spielman and Teng 2004) all based on different geometric principles. We could really fill an entire book on the geometry of all linear programming algorithms.

To prove my point, let me take a quick look at one lovely geometric challenge that arises from the main competitors of the simplex method, the interior point methods. To state the results, let us consider the pair of linear programming problems in *primal* and *dual* formulation:

Maximize 
$$\mathbf{c}^T \mathbf{x}$$
 subject to  $A\mathbf{x} = \mathbf{b}$  and  $\mathbf{x} \ge 0$ ; (1)

Minimize 
$$\mathbf{b}^T \mathbf{y}$$
 subject to  $A^T \mathbf{y} - \mathbf{s} = \mathbf{c}$  and  $\mathbf{s} \ge 0$ . (2)

Here, A is an  $m \times n$  matrix. The primal-dual interior point methods are among the most computationally successful algorithms for linear optimization. While the simplex methods follow an edge path on the boundary, the interior point methods follow the *central path*. The famous primal-dual central path is given by the following system of quadratic and linear polynomial equations:

$$A\mathbf{x} = \mathbf{b}, \quad A^T \mathbf{y} - \mathbf{s} = \mathbf{c}, \quad \text{and} \quad x_i s_i = \lambda \quad \text{for } i = 1, 2, \dots, n.$$
 (3)

The system has several properties: For all  $\lambda > 0$ , the system of polynomial equations has a unique real solution  $(\mathbf{x}^*(\lambda), \mathbf{y}^*(\lambda), \mathbf{s}^*(\lambda))$  with the properties  $\mathbf{x}^*(\lambda) > 0$  and  $\mathbf{s}^*(\lambda) > 0$ . The point  $\mathbf{x}^*(\lambda)$  is the optimal solution of the *logarithmic barrier function* for (1), which is defined as

$$f_{\lambda}(\mathbf{x}) := \mathbf{c}^T \mathbf{x} + \lambda \sum_{i=1}^n \log x_i.$$

Any limit point  $(\mathbf{x}^*(0), \mathbf{y}^*(0), \mathbf{s}^*(0))$  of these solutions for  $\lambda \to 0$  is the unique solution of the complementary slackness constraints, and thus yields an optimum point.

Optimization workers only look at the central path as connecting the optimal solution of the linear programs in question with its *analytic center* within one single cell, with  $x_i, s_i \ge 0$  (a cell of the arrangement is the polyhedron defined by a choice of signs in the constraints); but the central path is a portion of an *algebraic curve* that



J.A. De Loera

extends beyond a single feasibility region (given by sign constraints on variables  $x_i$ ,  $s_i$ ). Instead of studying the problem with only constraints  $x_i$ ,  $s_i \ge 0$ , one can ask for all feasible programs arising from any set of sign conditions. There are at most  $\binom{m-1}{n}$  such feasible sign vectors; (A, b) is said to be in general position if this number is attained. Then the central curve passes through all the vertices of a hyperplane arrangement.

In practical computations, interior point methods follow a piecewise-linear approximation to the central path. One way to estimate the number of Newton steps needed to reach the optimal solution is to bound the *total curvature* of the central path. The intuition is that curves with small curvature are easier to approximate with fewer line segments. This idea has been investigated by various authors (see, e.g., Monteiro and Tsuchiya 2008; Sonnevend et al. 1992/1991; Vavasis and Ye 1996; Zhao and Stoer 1993), and has yielded interesting results. For example, Vavasis and Ye (1996) found that the central path contains no more than  $n^2$  turning points. This finding led to an interior-point algorithm whose running time depends only on the constraint matrix A. Thus, in a way, curvature can be regarded as the continuous analogue of the diameter in the simplex method.

Dedieu et al. (2005) first investigated the differential geometric properties of the central curve of interior point methods. Their main theorem is as follows: Let (A,b,c) be as above with (A,b) in general position. Then the *average* total curvature of the primal, the dual, and the primal-dual central paths of the strictly feasible polytopes defined by (A,b) is at most  $2\pi (n-1)$  (primal), at most  $2\pi n$  (dual), and at most  $2\pi n$  (primal-dual), respectively. In particular, it is independent of the number m of constraints. Although traditionally the central path is only followed approximately by interior point methods with some kind of Newton steps in De Loera et al. (2012) the authors obtained explicit exact algebraic formulas for the primal central curve and gave improved bounds for the total curvature in terms of the degree of the Gauss maps of the curve. It is surprising that the formulas can be read from a matroid associated to the matrix A and the cost vector c.

Of course, for practical applications the more relevant quantity is not the average total curvature but rather the curvature in a single feasible region. This has been investigated by A. Deza, T. Terlaky, and Y. Zinchenko in a series of papers. Dedieu et al. conjectured that the curvature (in a single cell) could only grow linearly in the dimension, but Deza et al. (2009) constructed central paths that are forced to visit small neighborhoods near of all vertices of a cube, "à la Klee–Minty." In Deza et al. (2008), they proved that even for d=2 the total curvature can grow linearly in the number of facet constraints. They have conjectured the following curvature analogue of diameter:

**Conjecture 1** The curvature of a polytope, defined as the largest possible total curvature of the associated central path with respect to the various cost vectors, is no more than  $2\pi m$ , where m is the number of facets of the polytope.

This is a pretty conjecture that combines both the differential geometry of the central curve and the combinatorics of the input polyhedra. Many more of these kinds of challenges await us for each of the different linear programming algorithms.



### Comment 2: We still need to work harder to understand the geometry of pivoting

The simplex method is governed by a chosen pivoting rule, i.e., a method of choosing adjacent vertices with better objective function value. Starting with the historical 1972 construction of Klee and Minty (1972), showing that Dantzig's original pivoting rule could require exponentially many steps, researchers have discarded many of the popular pivot rules as good candidates for polynomial behavior. By 2010, almost all known natural deterministic pivoting rules were known to require an exponential number of steps to solve some linear programs (see Amenta and Ziegler 1999; Ziegler 2004), but three conspicuous pivot rules resisted the attacks of researchers until then. The most famous "untamed" pivot rules were Zadeh's rule (also known as the *least entered* rule) 2009, the randomized pivot rules of Random-Edge originally proposed by G. Dantzig, and *Random-Facet* proposed by Kalai (1992), and in a different form by Matoušek et al. (1996).

At any nonoptimal vertex, the Zadeh pivot rule chooses the decreasing edge that leaves the facet that has been left least often in the previous moves. In case of ties, a tie-breaking rule is used to determine the decreasing edge to be taken. Any other pivot rule can be used as a tie-breaking rule. The rule was proposed by Norman Zadeh in a 1980 technical report from the department of operations research of Stanford University. It has now appeared reprinted in Zadeh (2009). Now, the random edge pivot rule chooses, from among all improving pivoting steps (or edges) from the current basic feasible solution (or vertex), one uniformly at random. The description of Random-Facet pivoting is a bit more complicated as there are several versions: roughly, at any nonoptimal vertex v choose one facet F containing v uniformly at random and solve the problem restricted to F by applying the algorithm recursively to one of its facets. The recursion decreases the dimension of the polytope at each iteration, thus it will eventually restrict to a one-dimensional face, which is solved by following that edge. One repeats the process until reaching an optimum.

No nonpolynomial lower bounds were known until recently for these three pivot rules. Prior evidence of exponential behavior was given in Matoušek (1994) and Matoušek and Szabó (2006) that both the random edge and random facet pivot rules do not have a polynomial bound when used in a certain class of oriented graphs, which include the graph of a polyhedron oriented by an objective function. Morris (2002) showed bad behavior existed for random edge in the related setting of linear complementarity problems (see also Avis and Moriyama 2009; Gärtner and Kaibel 2007 for more on abstract graphs). But there was also evidence of good behavior in special cases (e.g., Balogh and Pemantle 2007; Kaibel et al. 2004/05) and the random facet rule can be shown to perform an expected subexponential number of steps in the worst case (Kalai 1992; Matoušek et al. 1996). This outperforms the deterministic pivot rules so far.

The results that these pivot rules are not always polynomial for specific simple LPs is an exciting breakthrough in the theory of the simplex method. The breakthrough came in two nice papers. The team of Friedmann et al. (2011) provided the first lower bound of the form  $2^{\Omega(n^{\alpha})}$ , for some  $\alpha > 0$ , for both the Random-Edge and the Random-Facet pivot rule in the one-pass variant. Based on Fearnley (2010), Friedmann et al. (2011) and Friedmann (2009), Friedmann constructed



J.A. De Loera

an exponential lower bound for the number of steps in Zadeh's rule (Friedmann 2011). Their new constructions use the close relation between simplex-type algorithms for solving linear programs and the policy iteration algorithms for the stochastic 1-player games called *Markov decision processes* or MDPs Bertsekas 2005; Puterman 1994). It should be remarked that the diameter of the resulting polytopes is actually smaller than the Hirsch bound. Curiously, Ye (2011) showed that the simplex method using Dantzig's pivot rule (where one chooses the entering variable with the largest reduced cost coefficient). is strongly polynomial for the linear programs derived from Markov Decision Processes with Fixed Discount (which is not the setting for the other papers, but is an important case of MDPs).

Ye's result has inspired others to revisit the classical analysis of pivot rules. In the case of pivot rule of Dantzig, and based on Ye's analysis, Kitahara and Mizuno (2011, 2012) have shown bounds for the number of basic feasible solutions (BFSs) visited by the simplex method with Dantzig's pivot rule. The bounds uses the relative sizes of the nonzero coordinates of the vertices: Given a linear program of the form  $\max\{c^Tx:Ax=b,x\geq 0\}$  where A is a real  $d\times n$  matrix, the number of different BFSs generated by this version of the simplex method is bounded by  $n\lceil d\frac{\gamma}{\delta}\log(d\frac{\gamma}{\delta})\rceil$ , where  $\delta$  and  $\gamma$  are the minimum and the maximum values of all the positive elements of primal BFSs and  $\lceil a \rceil$  denotes the smallest integer greater than a. Interestingly, they also presented a variant of Klee–Minty's LP, for which the number of iterations for this variant is equal to the ratio  $\frac{\gamma}{\delta}$ .

Beyond the theoretical analysis of pivot rules, there are a number of things we can learn from experiments. Ziegler (2004) reported on studies analyzing some of the well-known NETLIB collection of benchmark problems using the shadow boundary method, a pivot rule guided by the two-dimensional projection of the polytope. This pivot rule deserves more investigation and plays an interesting role in the smoothed analysis of linear programs (Spielman and Teng 2004). There is so much we still to understand.

### Comment 3: There are many interesting special polyhedra, which already present a worthy challenge to determine their diameter

It is clear that some additional structure can always help and I think one can exploit this even more that we have done so far. Indeed, as mentioned in Sect. 4.5 of Santos' article, some families of bounded coefficients on the vertices have been shown to satisfy (low-degree) polynomial diameter bounds (Brightwell et al. 2006; De Loera et al. 2009), but we it is frustrating that we do not know whether the Hirsch conjecture holds for them. A rather interesting family is the special case in which A is a totally unimodular matrix, which recently had a great improvement and it contains all network polytopes, in particular the classical transportation polytopes. As already stressed in Theorem 4.28 of the article, the bounds by Bonifas et al. (2011) provided a bound that is polynomial in the dimension and the largest absolute value of a sub-determinant of the defining integer matrix A. They exploited the metric information of the problem. For unimodular matrices, this gives the beautiful bound  $O(d^4 \log d)$  and  $O(d^{3.5} \log d)$  for general or unbounded polytopes. This greatly improves over the previous best bound for totally unimodular matrices by Dyer and Frieze (1994).



For the specific case of transportation polytopes, one is truly very close of the Hirsch bound. The current record is the theorem of Hurkens:

**Theorem 1** (Hurkens 2013) *The diameter of every*  $p \times q$  *transportation polytope is at most* 4(p+q-2).

The big question is can we close the gap further to reach the Hirsch bound in the case of transportation problems? In other words, is the diameter of  $p \times q$  transportation always less than or equal to p + q - 1.

The Hirsch conjecture holds for all known many cases of special transportation polytopes that restrict the margins. For example, the conjecture is true for Birkhoff's polytope and for some special right-hand sides (see, e.g., Borgwardt 2013). Our work continues today to try to settle this important instance.

**Acknowledgements** I want to thank Francisco Santos for many wonderful mathematical collaborations and a friendship of 20 years. I also want to thank the editors of this volume for the invitation to contribute a commentary to this special issue. I am also grateful to the Technische Universität München for the hospitality received during the time of writing this article.

#### References

Agmon S (1954) The relaxation method for linear inequalities. Can J Math 6:382-392

Amenta N, Ziegler GM (1999) Deformed products and maximal shadows of polytopes. In: Advances in discrete and computational geometry, South Hadley, MA, 1996. Contemp math, vol 223. Am Math Soc, Providence, pp 57–90

Avis D, Moriyama S (2009) On combinatorial properties of linear program digraphs. In: Polyhedral computation. CRM proc lecture notes, vol 48. Am Math Soc, Providence, pp 1–13

Balogh J, Pemantle R (2007) The Klee–Minty random edge chain moves with linear speed. Random Struct Algorithms 30(4):464–483. doi:10.1002/rsa.20127

Bárász M, Vempala S (2010) A new approach to strongly polynomial linear programming. http://www.cc.gatech.edu/~vempala/papers/affine.pdf

Bertsekas DP (2005) Dynamic programming and optimal control, 3rd edn. Athena Scientific, Belmont

Bertsimas D, Vempala S (2004) Solving convex programs by random walks. J ACM 51(4):540–556 (electronic). doi:10.1145/1008731.1008733

Betke U (2004) Relaxation, new combinatorial and polynomial algorithms for the linear feasibility problem. Discrete Comput Geom 32(3):317–338. doi:10.1007/s00454-004-2878-4

Betke U, Gritzmann P (1992) Projection algorithms for linear programming. Eur J Oper Res 60(3):287–295. http://www.sciencedirect.com/science/article/pii/037722179290080S. doi:10.1016/0377-2217 (92)90080-S

Bonifas N, Di Summa M, Eisenbrand F, Hähnle N, Niemeier M (2011) On sub-determinants and the diameter of polyhedra. Available at arXiv:1108.4272

Borgwardt KH (2009) Probabilistic analysis of simplex algorithms. In: Floudas CA, Pardalos PM (eds) Encyclopedia of optimization. Springer, Berlin, pp 3073–3084

Borgwardt S (2013, in press) On the diameter of partition polytopes and vertex-disjoint cycle cover. Math Program

Brightwell G, van den Heuvel J, Stougie L (2006) A linear bound on the diameter of the transportation polytope. Combinatorica 26(2):133–139. doi:10.1007/s00493-006-0010-5

Chubanov S (2012) A strongly polynomial algorithm for linear systems having a binary solution. Math Program 134(2):533–570

De Loera JA, Hemmecke R, Köppe M (2013) Algebraic and geometric ideas in the theory of discrete optimization. MOS-SIAM series on optimization, vol 14. SIAM, Philadelphia

De Loera JA, Kim ED, Onn S, Santos F (2009) Graphs of transportation polytopes. J Comb Theory, Ser A 116(8):1306–1325. doi:10.1016/j.jcta.2009.03.010



J.A. De Loera

De Loera JA, Sturmfels B, Vinzant C (2012) The central curve in linear programming (with Jesùs De Loera and Bernd Sturmfels). Found Comput Math 12:509–540

- Dedieu JP, Malajovich G, Shub M (2005) On the curvature of the central path of linear programming theory. Found Comput Math 5(2):145–171. doi:10.1007/s10208-003-0116-8
- Deza A, Terlaky T, Zinchenko Y (2008) Polytopes and arrangements: diameter and curvature. Oper Res Lett 36(2):215–222. doi:10.1016/j.orl.2007.06.007
- Deza A, Terlaky T, Zinchenko Y (2009) Central path curvature and iteration-complexity for redundant Klee–Minty cubes. In: Advances in applied mathematics and global optimization, Adv mech math, vol 17. Springer, New York, pp 223–256
- Dunagan J, Vempala S (2008) A simple polynomial-time rescaling algorithm for solving linear programs. Math Program, Ser A 114(1):101–114. doi:10.1007/s10107-007-0095-7
- Dyer M, Frieze A (1994) Random walks, totally unimodular matrices, and a randomised dual simplex algorithm. Math Program, Ser A 64(1):1–16. doi:10.1007/BF01582563
- Fearnley J (2010) Exponential lower bounds for policy iteration. In: Proceedings of the 37th ICALP, pp 551–562
- Friedmann O (2009) An exponential lower bound for the parity game strategy improvement algorithm as we know it. In: Proceedings of the 24th LICS, pp 145–156
- Friedmann O (2011) A subexponential lower bound for Zadeh's pivoting rule for solving linear programs and games. In: Proceedings of the 15th conference on Integer Programming and Combinatorial Optimization, IPCO'11, New York, NY, USA
- Friedmann O, Hansen T, Zwick U (2011) Subexponential lower bounds for randomized pivoting rules for the simplex algorithm. In: Proceedings of the 43rd ACM symposium on Theory of Computing, STOC'11, San Jose, CA, USA
- Fukuda K, Terlaky T (1997) Criss-cross methods: a fresh view on pivot algorithms. Math Program, Ser B 79(1–3):369–395. doi:10.1016/S0025-5610(97)00062-2. Lectures on mathematical programming (ismp97), Lausanne (1997)
- Gärtner B, Kaibel V (2007) Two new bounds for the random-edge simplex-algorithm. SIAM J Discrete Math 21(1):178–190 (electronic). doi:10.1137/05062370X
- Goffin JL (1982) On the nonpolynomiality of the relaxation method for systems of linear inequalities. Math Program 22(1):93–103. doi:10.1007/BF01581028
- Hačijan LG (1979) A polynomial algorithm in linear programming. Dokl Akad Nauk SSSR 244(5):1093–
- Hurkens C (2013) Personal communication
- Kaibel V, Mechtel R, Sharir M, Ziegler GM (2004/05) The simplex algorithm in dimension three. SIAM J Comput 34(2):475–497 (electronic). doi:10.1137/S0097539703434978
- Kalai G (1992) A subexponential randomized simplex algorithm. In: Proceedings of the 24th ACM symposium on theory of computing, STOC. Victoria, Columbia
- Kitahara T, Mizuno S (2011b) Klee–Minty's lp and upper bounds for Dantzig's simplex method. Oper Res Lett 39(2):88–91
- Kitahara T, Mizuno S (2012) A bound for the number of different basic solutions generated by the simplex method. Math Program 134(2):533–570
- Klee V, Minty GJ (1972) How good is the simplex algorithm? In: Inequalities, III, Proc third sympos, Univ. California, Los Angeles, Calif., 1969 Academic Press, New York, pp 159–175. Dedicated to the memory of Theodore S Motzkin
- Matoušek J (1994) Lower bounds for a subexponential optimization algorithm. Random Struct Algorithms 5(4):591-607. doi:10.1002/rsa.3240050408
- Matoušek J, Sharir M, Welzl E (1996) A subexponential bound for linear programming. Algorithmica 16(4–5):498–516. doi:10.1007/s004539900062
- Matoušek J, Szabó T (2006) RANDOM EDGE can be exponential on abstract cubes. Adv Math 204(1):262–277. doi:10.1016/j.aim.2005.05.021
- Monteiro RDC, Tsuchiya T (2008) A strong bound on the integral of the central path curvature and its relationship with the iteration-complexity of primal-dual path-following LP algorithms. Math Program, Ser A 115(1):105–149. doi:10.1007/s10107-007-0141-5
- Morris WD Jr. (2002) Randomized pivot algorithms for *P*-matrix linear complementarity problems. Math Program, Ser A 92(2):285–296. doi:10.1007/s101070100268
- Motzkin TS, Schoenberg IJ (1954) The relaxation method for linear inequalities. Can J Math 6:393–404 Puterman ML (1994) Markov decision processes: discrete stochastic dynamic programming. Wiley series in probability and mathematical statistics: applied probability and statistics. Wiley, New York. A Wiley-Interscience Publication



- Santos F (2012) A counterexample to the Hirsch conjecture. Ann Math 176(1):383–412. doi:10.4007/annals.2012.176.1.7
- Santos F (2013) Recent progress on the combinatorial diameter of polytopes and simplicial complexes. TOP Real Soc Estadí stica e Investigación Operativa, this issue
- Sonnevend G, Stoer J, Zhao G (1992/1991) On the complexity of following the central path of linear programs by linear extrapolation. II. Math Program, Ser B 52(3):527–553. doi:10.1007/BF01582904 Interior point methods for linear programming: theory and practice (Scheveningen, 1990)
- Spielman DA, Teng SH (2004) Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time. J ACM 51(3):385–463 (electronic). doi:10.1145/990308.990310
- Telgen J (1982) On relaxation methods for systems of linear inequalities. Eur J Oper Res 9(2):184–189. doi:10.1016/0377-2217(82)90071-6
- Vavasis SA, Ye Y (1996) A primal-dual interior point method whose running time depends only on the constraint matrix. Math Program, Ser A 74(1):79–120. doi:10.1016/0025-5610(95)00043-7
- Ye Y (2011, to appear) The simplex and policy-iteration methods are strongly polynomial for the Markov decision problem with a fixed discount rate. Math Oper Res. http://www.stanford.edu/~yyye/SimplexMDP4.pdf
- Zadeh N (2009) What is the worst case behavior of the simplex algorithm? In: Polyhedral computation. CRM proc lecture notes, vol 48. Am Math Soc, Providence, pp 131–143
- Zhao G, Stoer J (1993) Estimating the complexity of a class of path-following methods for solving linear programs by curvature integrals. Appl Math Optim 27(1):85–103. doi:10.1007/BF01182599
- Ziegler GM (2004) Typical and extremal linear programs. In: The sharpest cut, MPS/SIAM ser optim. SIAM, Philadelphia, pp 217–230

