Paul Breiding, Frank Sottile, James Woodcock.

Paul Breiding, Frank Sottile, James Woodcock.

Editor of the Journal of the Society for the
Foundations of Computational Mathematics

May 10, 2021

Response to the referees

Dear editor, dear referees,

Thank you for your reading and commenting on our manuscript.

Below, we list our detailed replies to the issues pointed out each referee. Our answers are in blue, the text from the referees in is black. We also attach a revised version of the manuscript with all changes highlighted in green.

Sincerely,

Paul Breiding, Frank Sottile, James Woodcock.

Detailed response to Referee 1

In Section 3, it is not very clear to me why Corollary 1 follows from Theorem 1 given the definition of an optimal homotopy before the corollary. It seems that Theorem 1 provides the exact number of solutions for the start system, hence property (1) is satisfied, but properties (2) and (3) are related to the homotopy method and not to Theorem 1. Could the authors add a few words to clarify this point?

We made some clarifications in the text. The EDD system has mixed volume number of solutions—general f, u have all solutions without multiplicity, and the polyhedral homotopy is optimal in this case.

The authors may want to check if their second main result, Theorem 2, could be directly obtained as a corollary of Proposition 2.6 in [9] for complete intersections.

No. Proposition 2.6 in [9] involves polynomials whose Newton polytope is a simplex, not a Box, as in Theorem 2.

Section 4.1: The phrase "The support of $\partial_i - \lambda(u_i - x_i)$ is obtained by adding ..." does not take into account that the support A of the polynomial f is a subset of N^n while the supports of the polynomials corresponding to the other Lagrange multiplier equations are subsets of N^{n+1}. The same comment holds for Section 6 where the $B(a)$ is a polytope in R^n while $P_i(a)$ is a polytope in R^{n+1}, you need to embed $B(a)$ to R^{n+1}

We added some text to make it clear that R^n is always embedded into R^{n+1} as $\{0\} \times R^n$.

I'm not sure that the following statements need further justification or are straightforward consequences of the definitions in which case my comment can be ignored:
a) the statement "Since $f_w \in C[x_I]$ is a general polynomial of support A_w" and
b) the argument in the proof of Lemma 3 that the specialization of a polynomial g wrt variable x_j is general w.r.t. the corresponding projection of its polytope.

These are standard. For (a): In the space of polynomials with support A_w, there is a hypersurface consisting of polynomials that are not general. The set of f of support A_w with f_w in this hypersurface is a hypersurface in the space of polyomials with support A, and therefore general.
For (b), there are finitely many substitutions of x_j such that the specialized polynomial is not general. Almost all polynomials do not have roots in that finite set.

Detailed response to Referee 2

We include comments only when we deviated from the suggestions of the referee, or when the change was substantial.

(1) Better to use capital letters for Euclidean Distance Degree.

Done.

(2) Line 5 of Introduction: "if X is nonempty". Is that true even if X is non reduced?

Varieties are reduced.

(3) Insert "and $e_0(a) := 1$" before the statement of Theorem 2.

We do not use e_0

(4) Line 2, page 4, replace "its" with "the": "...whereas Theorem 2 concerns [the] Newton polytope of f".

We reworded this differently, but thanks for pointing it out.

(7) Line 16, page 5, replace "all zeros" with "every zero": "...for [every zero] of $F(x)$ there is a continuation path which converges to it".

We did this and also revised the discussion of optimal homotopy.

(8) Line 9, page 6, replace "possibly" with "possible". Then I suggest changing the next sentence as "...we can take the [set of] intervals $\{r_1, \ldots, r_k\}$ [of] R^n ,...". Similarly, I suggest changing the sentence in line 12 as "... intervals $\{d_1, \ldots, d_k\}$ [of] R^n ."

Done. We also clarified the discussion in the previous paragraph.

(10) Line 3, page 8: in the sentence "Let $\partial_i A$ be the support of ∂_i", ∂_i is an operator, so it would be probably better to replace it with the polynomial $\partial_i f$. Similarly in line 6 "$\partial_i - \lambda(u_i - x_i)$.

Done. It was a typo.

(11) Equation (9), page 9: the notation $\partial_i A_w$ is a bit misleading. I would suggest using the notation $(\partial_i A)_w$ in equation (9), and then say that you adopt the shorthand $\partial_i A_w$. Moreover, at this point I would stress that there is always a strict inclusion $\partial_i(A_w) \subset (\partial_i A)_w$.....

We added this to the proof of Lemma 1, which was completely rewritten, and simplified.

(12) First line in the proof of Lemma 1, page 10: does the definition of A° depend on the index i.....

This was removed from the proof, so this question is moot.

(13) Proof of Lemma 1: the proof is correct, although I found quite difficult to follow its various steps. Hence I suggest this alternative one, which is divided in four steps:

The proof was rewritten along different lines

(16) Line 4 in the proof of Theorem 3: isn't the property $(\partial_i f w = 0) \Rightarrow (A_w \subset \{a \in N^n \mid a_i = 0\})$ true for any f and any vector w, as A_w is the support of f_w?

Yes. We removed the general.

(17) First line after equation (13): I would remove the redundant sentence "is a polynomial in only the variables x_I", as the previous expression $f_w \in C[x_I]$" already says it.

Reworded this sentence.

(18) Line 6, page 12: I suggest changing the letter "a" in the sentence "If $a \in A \setminus A_w$, then $w \cdot a > 0$ as $h^* = 0$", because two lines below the expression "$a + e_j \in A \setminus A_w$" considers another vector a and might create confusion.

We are using a as a variable for integer vecors, and these occurrences are all bound variables. We also changes the b to a.

(19) Seven lines before equation (17), page 12: remove the redundant sentence "Let $i \in I$ be an index with $w_i < 0$".

We reworded the setence.

(20) Two lines before Section 6, page 13, replace "thus" with "this": "But [this] contradicts...".

Got it. Thanks.

(21) Line 6 of Section 6, page 13, remove the comma after "R^n".

Done.

(22) First line of Section 6.2, page 14: the index m runs between 1 and n.

Thanks. Fixed.

(23) Line 6 in Section 6.3, page 15, missing parenthesis in the end of "$a_j(1+E(\pi_j(a))$". Moreover, here I would cross-refer the definition of $E(a)$ by making the definition before Theorem 4 an equation.

Done.

(24) First line of Section 6.4, page 15, typo "suport".

Fixed

(25) Fifth line of Section 6.4, page 15, typo "Simiarly".

Fixed

(26) Throughout the paper you use two different notations for indices running in an interval, either $1 \leq i \leq n$ or $i = 1, \ldots, 4$, but not both.

fixed

Detailed response to Referee 3

For the practitioner, the paper should clarify on how to determine if given coefficients are generic. For example, do the results explain why the ED degree of a circle $x^2+y^2 = 1$ is two instead of the predicted four by the mixed volume?

TBW

Constant term assumption: Throughout the article the authors made the assumption f has a generic constant term. The authors should say if this is necessary because they want X to be smooth.

TBW

The formula for Euclidean distance degrees is very nice, but the authors didn't mention how this is connected to applications.

TBW

Helmer's and Sturmfels' "Nearest Points on Toric Varieties" appears relevant.

TBW

The authors say "Corollary 5 is one of the few known instances of a structured problem for which we have an optimal homotopy available.", but there are numerous instances of polyhedral homotopies being optimal. Two textbook example are on Nash Equilibria and the eigenvalue problem.

We are not aware of instances other than a few including the ones you mentioned. Nevertheless, we have weakened the statement a little bit by saying that "Corollary 5 is an instance of ...". We hope that this is acceptable for the referee.

Aluffi's results: Since the authors are assuming $f = 0$ is a smooth hypersurface, it needs to be explained why Aluffi's cited results are not applicable.

TBW

I do not know what the authors mean by "in the context of learning the topological configuration of X" and the "area of learning."

We have clarified the first appearance and removed the second one.