Rend. Lincei Mat. Appl. 29 (2018), 465–485
DOI 10.4171/RLM/817

Algebraic Geometry — *Identifiability beyond Kruskal's bound for symmetric tensors of degree 4*, by Elena Angelini, Luca Chiantini and Nick Vannieuwenhoven, communicated on February 9, 2018.[1]

Abstract. — We show how methods of algebraic geometry can produce criteria for the identifiability of specific tensors that reach beyond the range of applicability of the celebrated Kruskal criterion. More specifically, we deal with the symmetric identifiability of symmetric tensors in $\mathrm{Sym}^4(\mathbb{C}^{n+1})$, i.e., quartic hypersurfaces in a projective space $\mathbb{P}^n$, that have a decomposition in $2n+1$ summands of rank 1. This is the first case where the reshaped Kruskal criterion no longer applies. We present an effective algorithm, based on efficient linear algebra computations, that checks if the given decomposition is minimal and unique. The criterion is based on the application of advanced geometric tools, like Castelnuovo's lemma for the existence of rational normal curves passing through a finite set of points, and the Cayley–Bacharach condition on the postulation of finite sets. In order to apply these tools to our situation, we prove a reformulation of these results, hereby extending classical results such as Castelnuovo's lemma and the analysis of Geramita, Kreuzer, and Robbiano, Cayley–Bacharach schemes and their canonical modules, *Trans. Amer. Math. Soc.* 339:443–452, 1993.

Key words: Identifiability of symmetric tensors, Waring decomposition, Kruskal's criterion, Cayley–Bacharach property, Castelnuovo's Lemma

Mathematics Subject Classification: 14J70, 14C20, 14N05, 15A69, 15A72

1. Introduction

The aim of this paper is the continuation of the study, started in [COV17b], of conditions which imply the identifiability of symmetric tensors, in a numerical range where the celebrated Kruskal criterion does not apply. Recall that a symmetric tensor T is *identifiable* if there exists a unique decomposition $T = T_1 + \cdots + T_r$ with a minimal number of symmetric rank-1 terms, up to scaling and reordering of the summands. This decomposition is called a *symmetric tensor rank decomposition* or *Waring decomposition*.

Beyond its theoretical interest, identifiability plays a central role in many applications of symmetric tensors. An important class of applications is found in algebraic statistics and machine learning. Indeed, the parameters of several latent variable models, including topic models, latent Dirichlet allocation, and hidden Markov models, can be recovered from the unique Waring decomposition of a symmetric tensor T that is associated with the model, as shown in [AGHKT14,

[1] The first and second authors are members of the Italian GNSAGA-INDAM.

AMR09, SKLV17]. Since one wishes to recover and interpret the parameters of the model uniquely, it is important to verify that the symmetric decomposition (computed by some numerical algorithm) is unique. In 1977, Kruskal [K77] determined what is now still the most popular criterion for testing the identifiability of a specific decomposition of a given, general tensor. Kruskal showed that a given decomposition is identifiable if its length is smaller than some numerical condition on the *Kruskal ranks* associated with the decomposition. Applying Kruskal's result to a reshaped tensor is considered to be a state-of-the-art effective criterion for identifiability by [COV17b].

Despite its popularity, there are three limitations to Kruskal's test in the symmetric setting. First, Kruskal's criterion verifies whether the tensor has no other tensor rank decompositions. It is possible, in principle, that a symmetric tensor has only one symmetric decomposition but several tensor rank decompositions of the same length. Indeed, Shitov [S17] recently provided a counterexample to Comon's conjecture, which entails that this may happen. Second, the length r of the decompositions of which Kruskal's criterion can prove identifiability is much smaller than the range wherein generic symmetric identifiability holds [COV17a]. Third, Derksen proved in [D13] that the numerical condition on the Kruskal ranks is sharp in the sense that if s is the largest rank allowed by the numerical condition in Kruskal's criterion, then there exist unidentifiable tensors of rank $s+1$ that still satisfy the numerical condition.

REMARK 1.1. The construction in [D13] also applies to the symmetric case. Indeed, taking all Kruskal ranks equal to the maximum value, the example constructed in [D13] becomes a symmetric tensor decomposition $T = T_1 + \cdots + T_{s+1}$ that admits another, distinct symmetric tensor decomposition $T = T'_1 + \cdots + T'_{s+1}$; hence, [D13] also proved that Kruskal's criterion is sharp in the symmetric setting.

The foregoing reasons motivated a further study of the specific identifiability of symmetric tensors, as in [BC12, COV17b, MMS17]. However, results beyond the (reshaped) Kruskal criterion are sparse. Moreover, some of these criteria require the use of general computer algebra algorithms which rapidly become ineffective for high-dimensional varieties. In Proposition 6.3 of [COV17b], a criterion for the identifiability of quartics in $\mathbb{P}^3$ (symmetric tensors of type $4 \times 4 \times 4 \times 4$), having a decomposition with 7 summands was given; this is exactly the first value beyond the range of the reshaped Kruskal's criterion. The goal of this paper is to extend the analysis to symmetric tensors in $S^4\mathbb{C}^{n+1}$, for any n. For such tensors the reshaped Kruskal criterion can prove the identifiability only for decompositions with $2n$ or less summands. We propose here a test that verifies the identifiability even for decompositions of length $2n+1$. Moreover, the test is *effective*, in the sense of [COV17b]: it will give a positive answer except on a set of measure 0 in the variety of tensors decomposed with $2n+1$ or less summands.

The idea of the test is based on the following observation. If $T \in S^4\mathbb{C}^{n+1}$ has a decomposition with $2n+1$ summands it may happen (as predicted in [D13]) that the decomposition is not unique, even if its associated Kruskal ranks have the maximum value $n+1$, which is equivalent to say that the decomposition is in linear general position, see Definition 2.2. However, in this case, we prove the existence of a positive-dimensional family of decompositions for T, which includes the given one. Specifically, we prove that when T has two different decompositions of length $2n+1$, then the pre-images of the two decompositions in the Veronese map determine a finite subset of points Z in a projective space $\mathbb{P}^n = \mathbb{P}(\mathbb{C}^{n+1})$ which lies in a rational normal curve. This rational normal curve in $\mathbb{P}^n$ induces the existence of the positive-dimensional family of decompositions for the tensor T, which can be detected using only linear algebra. These insights yield the new criterion (see Section 6.1).

We notice that the failure of identifiability for tensors of rank $2n+1$ in $S^4\mathbb{C}^{n+1}$ is caused by the existence of a low dimensional variety – the rational normal curve – that contains the pre-images of the points of the decomposition and forces the existence of other decompositions of the same tensor. The link between the failure of identifiability and the existence of subvarieties of positive dimension containing the decompositions is somehow familiar in the study of tensors. This phenomenon can occur for generic tensors, where the varieties that force the existence of many decompositions of T are the *contact varieties*; see [CC06]. In the case of symmetric tensors of very low rank, it was shown in [BC13] that unidentifiable cases occur only if the decompositions are contained in a positive-dimensional subvariety. In this paper, we prove that the same fact holds for specific tensors of rank $2n+1$ in $S^4\mathbb{C}^{n+1}$, and we expect that a similar characterization of unidentifiable specific tensors can be proved in other cases which lie just beyond Kruskal's range.

The basic tool in our analysis is provided by the study of the geometry of finite sets Z in the projective space $\mathbb{P}^n$. We will perform the analysis by means of classical methods in algebraic geometry, essentially related to the Hilbert function of Z. However, we cannot plainly use the large body of classical and modern results on finite sets in projective spaces. The reason is that when we have a decomposition of T, and we want to exclude the existence of a second decomposition, then we argue on the pre-images of the two decompositions, which determine two sets $A, B \subset \mathbb{P}^n$. In order to achieve our result, i.e., that when A is in linear general position then $Z = A \cup B$ lies in a rational normal curve, we can control only the geometry of A, as we know nothing about the hypothetical set B; in particular, we cannot place assumptions on the geometry of B. For this reason, we need to produce refinements of well-known geometric results, such as Castelnuovo's Lemma (see Lemma 5.3), in which we sharpen the hypothesis on the generality of the position of the points in Z in Lemma 5.4. Similarly, we introduce a *relative* version of the Cayley–Bacharach condition (see Definition 4.1), and extend a result of Geramita, Kreuzer and Robbiano (see Theorem 4.9 below).

We hope that our analysis can be of independent interest in the theory of finite sets in projective spaces. We also believe that it can support the idea that geometric results on the geometry of sets of points can produce interesting consequences for the theory of symmetric tensors. We also strongly believe that further analyses of the same type can provide new applications of algebraic geometry methods in the study of tensors, as well as stimulate the research on the Hilbert functions of finite sets, by indicating which refinements of known results could produce non-trivial applications to tensor analysis.

The rest of this article is structured as follows. In the next section some elementary results about the Hilbert function of finite sets are recalled. Kruskal's identifiability criterion of tensors is recalled in Section 3. We then investigate, in Section 4, the Hilbert function of sets with the Cayley–Bacharach property. In Section 5, the assumptions in the classic Castelnuovo Lemma are relaxed. Finally, we apply the results from aforementioned sections to the identifiability of fourth-order symmetric tensors whose rank is one higher than the range in which the (reshaped) Kruskal criterion applies.

2. Preliminaries

2.1. Notation

Let T be a homogeneous polynomial in $n+1$ variables of degree d over $\mathbb{C}$, i.e., $T \in S^d\mathbb{C}^{n+1}$. T is associated to an element of $\mathbb{P}(S^d\mathbb{C}^{n+1})$, which by abuse of notation we denote by T.

For any $m \in \mathbb{N}$, let $\mathbb{P}^m = \mathbb{P}^m_{\mathbb{C}}$ be the m-dimensional complex projective space and let $v_d : \mathbb{P}^n \to \mathbb{P}^N$ be the Veronese embedding of $\mathbb{P}^n$ of degree d, where $N = \binom{n+d}{d} - 1$. Let $A \subset \mathbb{P}^n$ be a finite set. We denote by $\ell(A)$ the cardinality of A and we define

$$v_d(A) = \{v_d(P_1), \dots, v_d(P_{\ell(A)})\} \subset \mathbb{P}^N,$$

where $P_i \in A$.

With the above notations we give the following definitions.

Definition 2.1. Let $A \subset \mathbb{P}^n$ be a finite set. A *computes* T if $T \in \langle v_d(A) \rangle$, the linear space spanned by the points of $v_d(A)$.

Recall that the Kruskal rank of a finite set $A \subset \mathbb{P}^n$ is defined as the maximum value k such that all subsets of k points from A are linearly independent. By definition, the maximum value for the Kruskal rank of A is thus $\min\{\ell(A), n+1\}$, which is also the generic value.

Definition 2.2. A finite set $A \subset \mathbb{P}^n$ is in *linear general position* (LGP) if the Kruskal rank of A is maximal, i.e., equal to $\min\{\ell(A), n+1\}$. This implies that any subset of A of cardinality at most $n+1$ is linearly independent.

Definition 2.3. Let $A \subset \mathbb{P}^n$ be a finite set which computes T. A is *minimal* if we cannot find a proper subset A' of A such that $T \in \langle v_d(A') \rangle$.

REMARK 2.4. If $A \subset \mathbb{P}^n$ is a finite set that computes T and satisfies the minimality property, then the points of $v_d(A)$ are linearly independent, i.e.,

$$\dim(\langle v_d(A) \rangle) = \ell(A) - 1.$$

Recall that a *rational normal curve* is a curve $\Gamma \subset \mathbb{P}^n$ corresponding to the Veronese embedding of $\mathbb{P}^1$ of degree n. Rational normal curves are the only irreducible curves of degree n in $\mathbb{P}^n$.

REMARK 2.5. If Z is a finite subset of a rational normal curve $\Gamma \subset \mathbb{P}^n$, then Z is always in LGP; see [H92] at the bottom of page 10.

REMARK 2.6. It is classically known that curves are never defective, i.e., their secant varieties always have the expected dimension. Thus, if $\Gamma \subset \mathbb{P}^n$ is a rational normal curve, then the dimension of the k-secant variety $\sigma_k(\Gamma)$ is the expected

$$\min\{n, 2k - 1\}.$$

2.2. *The Hilbert function of finite sets and its difference*

We collect in this section a series of definitions and propositions which are well known to people working in algebraic geometry, but maybe not so familiar to other people working in tensor analysis. The main definition is the *Hilbert function* of a finite set in a projective space, which is a basic tool for our results on the identifiability of symmetric tensors. Readers can find detailed studies on the Hilbert functions, and their relations with tensor analysis, in the book of A. Iarrobino and V. Kanev [IK99].

DEFINITION 2.7. Let $Y \subset \mathbb{C}^{n+1}$ be a finite, reduced set of cardinality ℓ and let $j \in \mathbb{N}$.

The *evaluation map of degree j on Y* is the linear map

$$ev_Y(j) : S^j\mathbb{C}^{n+1} \to \mathbb{C}^\ell$$

which sends $F \in S^j\mathbb{C}^{n+1}$ to the evaluation of F at the points of Y.

We will use the evaluation map to define the Hilbert function of a finite set $Z \subset \mathbb{P}^n$.

REMARK 2.8. Let $Z \subset \mathbb{P}^n$ be a finite set. Choose a set of homogeneous coordinates for the points of Z. We get a set of vectors $Y \subset \mathbb{C}^{n+1}$, for which the evaluation map $ev_Y(j)$ is defined for every j. If we change the choice of the homogeneous coordinates for the points of the fixed set Z, we get another set $Y' \subset \mathbb{C}^{n+1}$ and the evaluation map $ev_{Y'}(j)$ differs from $ev_Y(j)$ for the multiplication by a nonsingular diagonal matrix. Thus the rank of $ev_Y(j)$ and $ev_{Y'}(j)$ are the same for all j.

Thanks to the previous remark, we can give the following definition of the Hilbert function of a finite set of points.

DEFINITION 2.9. Let $Z \subset \mathbb{P}^n$ be a finite set. Choose any set of homogeneous coordinates Y for the points of Z. Define the *Hilbert function* of Z as the map

$$h_Z : \mathbb{Z} \to \mathbb{N}$$

such that $h_Z(j) = 0$ for $j < 0$ and $h_Z(j) = \operatorname{rank}(ev_Y(j))$ for $j \geq 0$.

REMARK 2.10. For any $j \geq 0$, the value $h_Z(j)$ provides the number of conditions that Z imposes to the elements of $S^j\mathbb{C}^{n+1}$, i.e., $h_Z(j) = \dim(\langle v_j(Z)\rangle) + 1$. In particular, if $h_Z(j) = \ell(Z)$, then we say that Z *imposes independent conditions to forms of degree* j.

DEFINITION 2.11. The *first difference of the Hilbert function* Dh_Z of Z is given by

$$Dh_Z(j) = h_Z(j) - h_Z(j-1),$$

where $j \in \mathbb{Z}$. The set of non-zero values of Dh_Z is called the *h-vector* of Z.

We recall some elementary and well-known properties of h_Z and Dh_Z that will be useful throughout the paper.

LEMMA 2.12. *We have*

(i) $Dh_Z(j) = 0$ *for* $j < 0$;
(ii) $h_Z(0) = Dh_Z(0) = 1$;
(iii) $Dh_Z(j) \geq 0$ *for all* j;
(iv) $h_Z(j) = \ell(Z)$ *for all* $j \gg 0$;
(v) $Dh_Z(j) = 0$ *for* $j \gg 0$ *and* $\sum_j Dh_Z(j) = \ell(Z)$;
(vi) $h_Z(i) = \sum_{0 \leq j \leq i} Dh_Z(j)$;
(vii) *if* $h_Z(j) = \ell(Z)$, *then* $Dh_Z(j+1) = 0$;
(viii) *if* $Z' \subset Z$, *then, for every* $j \in \mathbb{Z}$, *we have* $h_{Z'}(j) \leq h_Z(j)$ *and* $Dh_{Z'}(j) \leq Dh_Z(j)$.

The next property is a consequence of the Macaulay maximal growth principle; see, e.g., Section 3 of [BGM94] for a proof.

PROPOSITION 2.13. *If for some* $j > 0$, $Dh_Z(j) \leq j$, *then*

$$Dh_Z(j) \geq Dh_Z(j+1).$$

In particular, if for some $j > 0$, $Dh_Z(j) = 0$, *then* $Dh_Z(i) = 0$ *for all* $i \geq j$.

REMARK 2.14. Notice that if for some j we have $Dh_Z(j) = 0$, then $h_Z(j-1) = h_Z(j)$. By Proposition 2.13, for any $i \geq j$ also $Dh_Z(i) = 0$, i.e., $h_Z(j-1) = h_Z(i)$

for any $i \geq j$. Therefore, by parts (v) and (vi) of Lemma 2.12,

$$\ell(Z) = \sum_k Dh_Z(k) = \sum_{k=0}^{i} Dh_Z(k) = h_Z(i).$$

Thus, $h_Z(j-1)$ is equal to the cardinality of Z, i.e., the evaluation map in degree $j-1$ surjects. In this case, for every $P \in Z$ we can find a form of degree i that vanishes at $Z \setminus \{P\}$ and does not vanish at P. Therefore, when $h_Z(i) = \ell(Z)$, we will also say that *hypersurfaces of degree i separate the points of Z*.

REMARK 2.15. Assume $h_Z(i) = \ell(Z) - 1$. Then $h_Z(i+1) > h_Z(i)$, for otherwise, by Proposition 2.13, $Dh_Z(j) = 0$ for all $j > i$, thus $h_Z(j) = h_Z(i)$ for all $j > i$, contradicting property (iv) of Lemma 2.12. Thus, if $h_Z(i) = \ell(Z) - 1$ then necessarily $h_Z(i+1) = \ell(Z)$.

REMARK 2.16. There is an alternative way to look at the Hilbert function of a finite set $Z \subset \mathbb{P}^n$. Indeed, let J_Z be the ideal sheaf of Z and let $j \in \mathbb{N}$. Then, we have the short exact sequence

$$0 \to J_Z(j) \to \mathcal{O}_{\mathbb{P}^n}(j) \to \mathcal{O}_Z(j) \to 0. \tag{1}$$

Passing to cohomology, (1) provides the exact sequence

$$0 \longrightarrow H^0(J_Z(j)) \longrightarrow H^0(\mathcal{O}_{\mathbb{P}^n}(j)) \xrightarrow{ev_Y(j)} H^0(\mathcal{O}_Z(j)) \longrightarrow H^1(J_Z(j)) \longrightarrow 0,$$

regardless of the choice of Y with $[Y] = Z$. Therefore,

$$h_Z(j) = \binom{j+n}{j} - \dim(H^0(J_Z(j))), \tag{2}$$

and, by Lemma 2.12 (v) and (vi),

$$\dim(H^1(J_Z(j))) = \ell(Z) - h_Z(j) = \sum_{i>j} Dh_Z(i). \tag{3}$$

3. KRUSKAL'S CRITERION FOR SYMMETRIC TENSORS

In this section, we recall the specialization of the celebrated Kruskal criterion for the identifiability of a decomposition to symmetric tensors. In fact, in the context of this paper, we are interested only in the generic case where the Kruskal ranks are maximal. In this case, the highest rank r of which Kruskal's criterion can prove identifiability is maximal. The particular version of Kruskal's Lemma that is of relevance is recalled as the following corollary of the results in [K77].

COROLLARY 3.1 (Kruskal [K77]). *Let T be an $n_1 \times n_2 \times n_3$ tensor over $\mathbb{C}$ with $n_1 \geq n_2 \geq n_3 \geq 2$. Assume that $T = T_1 + \cdots + T_r$, where the T_i's are tensors of* *rank* 1. *Write* $T_i = v_{1i} \otimes v_{2i} \otimes v_{3i}$. *If the sets* $A_j = \{v_{j1}, \ldots, v_{jr}\}$ *are in LGP and*

$$r \leq \frac{1}{2}(\min(n_1, r) + \min(n_2, r) + \min(n_3, r)) - 1$$

then T *has complex rank* r *and it is identifiable, in the sense that the set* $\{T_1, \ldots, T_r\}$ *is unique, including multiplicities.*

A value of r not satisfying the above inequality is said to be *beyond Kruskal's range* of identifiability.

According to Corollary 20 of [COV17b], Kruskal's criterion can be specialized to the case of symmetric tensors, as follows. As before, we only present the following corollary in the generic case where the Kruskal ranks are maximal.

Corollary 3.2 (Reshaped Kruskal's criterion for symmetric tensors [COV17b]). *Let* $T \in S^d\mathbb{C}^{n+1}$ *with* $d \geq 3$ *and* $n \geq 1$. *Let* $A \subset \mathbb{P}^n$ *be a finite set of cardinality* $r = \ell(A)$ *computing* T, *and let* $d_1 + d_2 + d_3 = d$ *be a partition of* d *such that* $d_1 \geq d_2 \geq d_3$. *If* $v_{d_i}(A)$ *is in LGP for* $i = 1, 2, 3$ *and*

$$r \leq \frac{1}{2}\left(\min\left\{\binom{d_1+n}{d_1}, r\right\} + \min\left\{\binom{d_2+n}{d_2}, r\right\} + \min\left\{\binom{d_3+n}{d_3}, r\right\}\right) - 1,$$

then T *has complex rank* r *and it is identifiable.*

Remark 3.3. Direct computations show that for the case $d = 4$, which is the core of this paper, the maximum range of applicability is attained for $d_1 = 2$ and $d_2 = d_3 = 1$, so that $\ell(A) \leq 2n$. Values of $r = \ell(A) > 2n$ are beyond the reshaped Kruskal's range.

4. Identifiability and the Cayley–Bacharach property

For a finite set, we define the Cayley–Bacharach property $CB(i)$ as follows.

Definition 4.1. A finite set $Z \subset \mathbb{P}^n$ satisfies the *Cayley–Bacharach property in degree* i, abbreviated as $CB(i)$, if for all $P \in Z$, it holds that every form of degree i vanishing at $Z \backslash \{P\}$ also vanishes at P.

Remark 4.2. If Z satisfies $CB(i)$, then it satisfies $CB(i-1)$ too. Otherwise, one could find $P \in Z$ and a hypersurface $F \subset \mathbb{P}^n$ of degree $(i-1)$ such that $Z \backslash \{P\} \subset F$ and $P \notin F$. Therefore, if $H_P \subset \mathbb{P}^n$ is a hyperplane not containing P, then $F \cup H_P \in H^0(J_{Z\backslash\{P\}}(i)) \backslash H^0(J_Z(i))$, which contradicts the hypothesis.

Remark 4.3. If Z satisfies $CB(i)$, then for any $P \in Z$, we have

$$H^0(J_{Z\backslash\{P\}}(i)) = H^0(J_Z(i)).$$

It follows from equation (2) and Remark 4.2 that

$$h_Z(j) = h_{Z\setminus\{P\}}(j) \quad \text{and} \quad Dh_Z(j) = Dh_{Z\setminus\{P\}}(j) \quad \forall j \leq i. \tag{4}$$

The Cayley–Bacharach property could thus be interpreted as the converse of the separation property introduced in Remark 2.14.

REMARK 4.4. From Remark 2.14 it is clear that if Z satisfies $CB(i)$, then hypersurfaces of degree i cannot separate the points of Z. We must namely have:

$$h_Z(i) < \ell(Z) \quad \text{or, equivalently} \quad Dh_Z(i+1) > 0. \tag{5}$$

Indeed, if $h_Z(i) = \ell(Z)$, then from (4) and the exact sequence

$$0 \to H^0(J_{Z\setminus\{P\}}(i)) \to H^0(\mathcal{O}_{\mathbb{P}^n}(i)) \to H^0(\mathcal{O}_{Z\setminus\{P\}}(i)) \to H^1(J_{Z\setminus\{P\}}(i)) \to 0,$$

it follows that

$$h^1(J_{Z\setminus\{P\}}(i)) = \ell(Z\setminus\{P\}) - h_{Z\setminus\{P\}}(i) = \ell(Z) - 1 - h_Z(i) = -1,$$

which is not possible.

We notice that the converse statement is false. For instance, the set Z consisting of four points in $\mathbb{P}^2$, three of them aligned, does not satisfy $CB(1)$, while $h_Z(1) < 4$.

For brevity, we define the following value.

DEFINITION 4.5. The *socle degree* i_Z of Z is the maximum i such that (5) holds.

Note that i_Z is the maximum i such that $Dh_Z(i+1) > 0$, i.e., the last element of the h-vector of Z. Additionally, if Z does not satisfy $CB(i)$, then it does not satisfy $CB(j)$ for all $j \geq i$ either, by the contrapositive of Remark 4.2. Therefore, if Z satisfies $CB(i)$, then $i \leq i_Z$.

EXAMPLE 4.6. Some examples of the Cayley–Bacharach property are shown below.

(i) Let Z be a set of 6 general points in $\mathbb{P}^2$. Then $Dh_Z = (1, 2, 3)$, $i_Z = 1$ and Z satisfies $CB(1)$.
(ii) Let Z be a set of 6 general points lying on an irreducible conic of $\mathbb{P}^2$. Then $Dh_Z = (1, 2, 2, 1)$, $i_Z = 2$ and Z satisfies $CB(2)$, and, hence, $CB(1)$.
(iii) Let Z be a set of 6 general points in $\mathbb{P}^2$, with 5 of them on a line plus one point off the line. Then $Dh_Z = (1, 2, 1, 1, 1)$, $i_Z = 3$ and Z does not satisfy $CB(1)$.

The following holds.

LEMMA 4.7. *Let $Z = \{P_1, \dots, P_r\} \subset \mathbb{P}^n$ be a finite set satisfying $CB(i)$. If, for any $j \in \{1, \dots, r\}$, the set $Z_j = Z \setminus \{P_j\}$ does not satisfy $CB(i)$, then*

$$h_Z(i+1) = \ell(Z) = r, \quad \textit{and so} \quad Dh_Z(i+2) = 0. \tag{6}$$

PROOF. By hypothesis, for every $j \in \{1, \dots, r\}$ there exists a point $Q_j \in Z_j$ and a form $F_j \in H^0(\mathcal{O}_{\mathbb{P}^n}(i))$ which vanishes at $Z \setminus \{P_j, Q_j\}$ but not at Q_j. Notice that $F_j(P_j)$ cannot be 0, since Z satisfies $CB(i)$. Let $H_j \in H^0(\mathcal{O}_{\mathbb{P}^n}(1))$ be a linear form vanishing at Q_j but not at P_j and consider $G_j = F_j \cdot H_j$. For any $j \in \{1, \dots, r\}$, it holds that $G_j \in H^0(J_{Z_j}(i+1)) \setminus H^0(J_Z(i+1))$, which implies that Z does not satisfy $CB(i+1)$. Moreover, for all j, $G_j(P_j) \neq 0$ and $G_j(P_k) = 0$ for any $k \neq j$. Therefore, in the exact sequence

$$0 \longrightarrow H^0(J_Z(i+1)) \longrightarrow H^0(\mathcal{O}_{\mathbb{P}^n}(i+1)) \xrightarrow{ev_Z(i+1)} H^0(\mathcal{O}_Z(i+1))$$

we have that $\mathrm{im}(ev_Z(i+1)) = H^0(\mathcal{O}_Z(i+1))$, i.e., $ev_Z(i+1)$ is a surjective map, which implies (6). □

The following result, due to Geramita, Kreuzer, and Robbiano, gives a strong bound on the Hilbert function of sets with a Cayley–Bacharach property.

THEOREM 4.8 (Geramita, Kreuzer, and Robbiano [GKR93]). *If a finite set $Z \subset \mathbb{P}^n$ satisfies $CB(i_Z)$, then we have*

(i) $Dh_Z(0) + Dh_Z(1) + \cdots + Dh_Z(j) \leq Dh_Z(i_Z + 1 - j) + \cdots + Dh_Z(i_Z + 1)$, *for any j with $0 \leq j \leq i_Z + 1$;*
(ii) $Dh_Z(0) + Dh_Z(1) + \cdots + Dh_Z(j) \leq Dh_Z(k - j) + \cdots + Dh_Z(k)$, *for any j, k with $0 \leq j \leq k \leq i_Z + 1$.*

PROOF. See Corollary 3.7 part (b) and (c) of [GKR93]. □

In order to create a link between Cayley–Bacharach properties and identifiability of symmetric tensors, we need to extend Theorem 4.8 by replacing i_Z with any integer i such that Z satisfies $CB(i)$.

THEOREM 4.9. *If a finite set $Z \subset \mathbb{P}^n$ satisfies $CB(i)$, then for any j such that $0 \leq j \leq i+1$ we have*

$$Dh_Z(0) + Dh_Z(1) + \cdots + Dh_Z(j) \leq Dh_Z(i+1-j) + \cdots + Dh_Z(i+1).$$

PROOF. We proceed by induction on the residual part $h_Z^1(i)$, which is defined as:

$$h_Z^1(i) = \ell(Z) - h_Z(i) = \ell(Z) - \sum_{j=0}^{i} Dh_Z(j) = \sum_{j=i+1}^{i_Z+1} Dh_Z(j).$$

If $h_Z^1(i) = 1$, then $\ell(Z) = h_Z(i) + 1$. Thus, by Remark 2.15 we must have $\ell(Z) = h_Z(i+1)$, i.e., $i = i_Z$ and we conclude by Theorem 4.8 part (i).

Next, we assume that the theorem is true for all $1 \le h^1_Z(i) \le e$, and prove it for $e+1$. If Z satisfies $CB(i_Z)$, then we can conclude by Theorem 4.8 part (ii). So assume that Z does not satisfy $CB(i_Z)$. In this case notice that there exists $P \in Z$ such that $Z_P = Z \setminus \{P\}$ satisfies $CB(i)$. Indeed, if for any $P \in Z$ the set Z_P does not satisfy $CB(i)$, then by Lemma 4.7 we have that $h^1_Z(i+1) = 0$, that is $i = i_Z$, which is a contradiction. Fix then a point $P \in Z$ such that Z_P satisfies $CB(i)$. Since Z satisfies $CB(i)$, by (4) we have that

$$h_{Z_P}(j) = h_Z(j) \quad \text{and} \quad Dh_{Z_P}(j) = Dh_Z(j)$$

for $j \le i$. Therefore,

$$h^1_{Z_P}(i) = \ell(Z_P) - \sum_{j=0}^{i} Dh_{Z_P}(j) = \ell(Z) - 1 - \sum_{j=0}^{i} Dh_Z(j) = h^1_Z(i) - 1,$$

so that by applying the induction hypothesis to Z_P, we get

$$Dh_{Z_P}(0) + Dh_{Z_P}(1) + \cdots + Dh_{Z_P}(j) \le Dh_{Z_P}(i+1-j) + \cdots + Dh_{Z_P}(i+1)$$

for every j such that $0 \le j \le i+1$. From (4) and $Dh_{Z_P}(i+1) \le Dh_Z(i+1)$, by Lemma 2.12 (viii), we get the conclusion. □

5. Castelnuovo's Lemma revisited

In order to apply the Cayley–Bacharach property to the identifiability of symmetric tensors, we need an *ad hoc* extension of the classical Castelnuovo Lemma for finite sets in $\mathbb{P}^n$.

Lemma 5.1 (Castelnuovo, [GH78] page 531). *Let $Z \subset \mathbb{P}^n$ be a finite set such that:*

(i) $\ell(Z) \ge 2n+3$*;*
(ii) *Z is in LGP;*
(iii) *Z imposes $2n+1$ conditions to quadrics, i.e., $h_Z(2) = 2n+1$.*

Then, the quadrics containing Z intersect in a rational normal curve, which thus contains Z.

In particular, we need to weaken the hypothesis, by assuming only that some subset of Z is in LGP. We do that in several steps.

Proposition 5.2. *Let $Z \subset \mathbb{P}^n$ be a finite set with a subset $A \subset Z$ of cardinality $2n+1$ in LGP. Then $h_Z(2) \ge 2n+1$.*

Proof. Since A is in LGP, then A imposes independent conditions to quadrics. Therefore $h_A(2) = 2n+1$. As $A \subset Z$, Lemma 2.12 (viii) concludes the proof. □

By means of Proposition 5.2, Castelnuovo's Lemma can be rephrased as follows:

Lemma 5.3. *Let $Z \subset \mathbb{P}^n$ be a finite set such that:*

(i) $\ell(Z) \geq 2n+3$;
(ii) *Z is in LGP;*
(iii) *Z imposes at most $2n+1$ conditions to quadrics, i.e., $h_Z(2) \leq 2n+1$.*

Then, $h_Z(2) = 2n+1$ and the quadrics containing Z intersect in a rational normal curve, which thus contains Z.

Based on Lemma 5.3, we can prove the following extension:

Lemma 5.4. *Let $Z \subset \mathbb{P}^n$ be a finite set such that:*

(i) $\ell(Z) \geq 2n+3$;
(ii) $h_Z(2) \leq 2n+1$;
(iii) *there exists $A \subset Z$ such that $\ell(A) = 2n+1$ and A is in LGP.*

Then, Z is in LGP and it is contained in a rational normal curve.

Proof. First, let us assume that $\ell(Z) = 2n+3$ and let us set $Z = A \cup \{P, Q\}$. We claim that $W = A \cup \{P\}$ is in LGP. Indeed, if this is not the case, there are subsets of W, of cardinality at most $n+1$, which are not linearly independent. This implies that one can find $V \subset W$ and a hyperplane H such that $\ell(V) = n+1$ and $V \subset H$. Since A is in LGP, then, necessarily, V contains P and n points of A. Thus we can renumber the points of $A = \{P_1, \ldots, P_{2n+1}\}$ so that $P_i \in H$ if and only if $i \leq n$. For any $j \in \{n+1, \ldots, 2n\}$, we can find a hyperplane $H_j \subset \mathbb{P}^n$ such that $\{P_{n+1}, \ldots, P_j\} \subset H_j$ and $\{P_{j+1}, \ldots, P_{2n+1}\} \not\subset H_j$, because A is in LGP. It follows that the quadric $Q_j = H \cup H_j$ contains $\{P, P_1, \ldots, P_j\}$ and misses $\{P_{j+1}, \ldots, P_{2n+1}\}$. In particular, if we set $V_j = \{P, P_1, \ldots, P_j\}$, then $h^0(J_{V_j}(2)) > h^0(J_{V_{j+1}}(2))$, which implies:

$$h_{V_j}(2) < h_{V_{j+1}}(2). \tag{7}$$

We notice that $h_{V_j}(1) = n+1$, for V_j contains at least $n+1$ points of A, which is in LGP. Therefore $Dh_{V_j}(1) = n$. Moreover, since $\ell(V_{n+1}) = n+2 = h_{V_{n+1}}(1) + 1$, by Remark 2.15 we get that $h_{V_{n+1}}(2) = \ell(V_{n+1})$. Thus, the h-vector of V_{n+1} is $(1, n, 1)$. By induction on j, we show that

$$h_{V_j}(2) = j+1, \quad \forall j \geq n+1. \tag{8}$$

Indeed, the claim holds for $j = n+1$. If (8) holds for $j \geq n+1$, then by (7) we have that $h_{V_{j+1}}(2) > j+1$. Since, by definition, $h_{V_{j+1}}(2) \leq \ell(V_{j+1}) = j+2$, necessarily it has to be the case that $h_{V_{j+1}}(2) = j+2$, as desired. By applying (8) to $j = 2n+1$, we get that

$$h_W(2) = h_{V_{2n+1}}(2) = 2n+2,$$

which contradicts assumption (ii) via Lemma 2.12 (viii), as $W \subset Z$. So, W is in LGP.

Now assume that Z is not in LGP. Then, as above, there exists $V \subset Z$ and a hyperplane $H \subset \mathbb{P}^n$ such that $\ell(V) = n+1$ and $Q \in V \subset H$. Then there exists $U \subset Z$ such that $\ell(U) = 2n+2$ and U is not in LGP. Since $h_U(2) \leq 2n+1 = h_Z(2)$, and U contains $2n+1$ points of W which is in LGP, then we get a contradiction by arguing as above.

Finally, assume that $\ell(Z) > 2n+3$. Notice that we have just proved the existence of $Z_0 \subset Z$ such that $\ell(Z_0) = 2n+3$ and Z_0 is in LGP. In particular, $h_{Z_0}(2) \leq 2n+1$. By Lemma 5.3, $h_{Z_0}(2) = 2n+1$ and Z_0 lies in a rational normal curve Γ, which is the intersection of all quadrics containing Z_0. Since $h_Z(2) \leq 2n+1$, then equality holds and $H^0(J_{Z_0}(2)) = H^0(J_Z(2))$. Thus Z itself is contained in Γ, which, by Remark 2.5, implies that Z is in LGP. □

Example 5.5. The previous formulation of Castelnuovo's Lemma is sharp, in the sense that the existence of a subset of cardinality $2n$ in LGP is not enough to guarantee that a set Z of $2n+3$ points in $\mathbb{P}^n$, with $h_Z(2) = 2n+1$, is contained in a rational normal curve.

Namely, take $n = 3$ and take a smooth quadric $Q \subset \mathbb{P}^3$, a set of 4 general points $Q_1, \ldots, Q_4$ on Q, a general line $L \subset Q$ of type $(0,1)$ and a set of 5 general points $P_1, \ldots, P_5$ on L. The set $Z = \{Q_1, \ldots, Q_4, P_1, \ldots, P_5\}$ has cardinality $9 = 2n+3$ and contains the subset $A = \{Q_1, \ldots, Q_4, P_1, P_2\}$ of cardinality $6 = 2n$ which is in LGP, due to the generality in the choice of the points. Since 4 general points of Q lie in two linearly independent divisors of type $(2,1)$, then there are 3 independent quadrics of $\mathbb{P}^3$ containing Z, i.e., $h_Z(2) = 7 = 2n+1$. The set Z, however, cannot lie in a rational normal curve, for it contains 5 points on a line.

6. Application to the identifiability of quartics

In this section, we apply the previous results on the geometry of finite sets to decompositions of quartic polynomials with the purpose of reaching beyond the range of Kruskal's criterion of identifiability.

Throughout this section, we consider $d = 4$ and let $T \in S^4\mathbb{C}^{n+1}$ be a homogeneous polynomial. Let $A \subset \mathbb{P}^n$ be a finite set that computes T, such that:

(i) $\ell(A) = 2n+1$, i.e., A has one point more than Theorem 3.2 allows; see Remark 3.3;
(ii) A is in LGP;
(iii) A satisfies the minimality property.

Let $B \subset \mathbb{P}^n$ be *another* finite set that computes T with $\ell(B) \leq 2n+1$. Without loss of generality we can also assume that B satisfies the minimality property. In particular, $T \in \langle v_4(A) \rangle \cap \langle v_4(B) \rangle$. Our target is to find criteria that exclude the existence of B, so that T has rank $2n+1$ and is identifiable.

In the following, we analyze the geometry of the union

$$Z = A \cup B.$$

Clearly, $\ell(Z) \le \ell(A) + \ell(B)$, with equality if $A \cap B$ is empty.

It is a straightforward fact that

$$\langle v_4(Z) \rangle = \langle v_4(A) \rangle + \langle v_4(B) \rangle.$$

so that by using Grassmann's formula, we have that

$$\dim(\langle v_4(Z) \rangle) = \dim(\langle v_4(A) \rangle) + \dim(\langle v_4(B) \rangle) - \dim(\langle v_4(A) \rangle \cap \langle v_4(B) \rangle). \tag{9}$$

From Remark 2.4, we then find that

$$\dim(\langle v_4(Z) \rangle) = \ell(A) + \ell(B) - 2 - \dim(\langle v_4(A) \rangle \cap \langle v_4(B) \rangle). \tag{10}$$

Since $A \cap B$ is a proper subset of A, then, by the minimality assumption on A, we get that $T \notin \langle v_4(A \cap B) \rangle$. Thus,

$$\dim(\langle v_4(A) \rangle \cap \langle v_4(B) \rangle) > \dim(\langle v_4(A \cap B) \rangle). \tag{11}$$

From (10), (11) and since $\dim(\langle v_4(A \cap B) \rangle) = \ell(A \cap B) - 1$, it follows that

$$\begin{aligned} \dim(\langle v_4(Z) \rangle) &< \ell(A) + \ell(B) - 2 - \dim(\langle v_4(A \cap B) \rangle) \\ &= \ell(A) + \ell(B) - \ell(A \cap B) - 1 = \ell(Z) - 1. \end{aligned} \tag{12}$$

We notice that

$$h_Z(4) \le \ell(Z) - 1.$$

Indeed by the inequality (12), the dimension of $\langle v_4(Z) \rangle$, which is, by Remark 2.10, $h_Z(4) - 1$, cannot be $\ell(Z) - 1$. Thus, by Proposition 2.13 and Lemma 2.12 (v):

$$Dh_Z(5) > 0. \tag{13}$$

We set now, as usual,

$$h^1_Z(4) = \ell(Z) - h_Z(4) = \sum_{j=5}^{\infty} Dh_Z(4).$$

By means of (3) and part (iv) of Lemma 2.12, we have that

$$\begin{aligned} \dim(\langle v_4(Z) \rangle) &= h_Z(4) - 1 = \ell(Z) - 1 - h^1_Z(4) \\ &= \ell(A) + \ell(B) - \ell(A \cap B) - 1 - h^1_Z(4). \end{aligned} \tag{14}$$

Since $\ell(A) = \dim(\langle v_4(A)\rangle) + 1$ and $\ell(B) = \dim(\langle v_4(B)\rangle) + 1$, then, comparing (14) and (9) we get that

$$\dim(\langle v_4(A)\rangle \cap \langle v_4(B)\rangle) = \ell(A \cap B) - 1 + h^1_Z(4). \tag{15}$$

Proposition 6.1. *We cannot have that $\ell(Z) \leq \ell(A) + 1 = 2n + 2$.*

Proof. From Proposition 5.2 and our assumptions, we know that $\ell(A) = h_A(2)$, i.e., quadrics separate the points of A. Thus $h_Z(2) \geq h_A(2) = 2n + 1$. It follows that if $\ell(Z) \leq 2n + 2$, then $h_Z(2) \geq \ell(Z) - 1$, which implies, by Remark 2.15, that $h_Z(3) = \ell(Z)$; thus, $Dh_Z(4) = 0$. This fact contradicts Proposition 2.13, since (13) holds. □

We use the previous arguments and Lemma 5.4 to prove the next crucial result.

Theorem 6.2. *Let $T \in S^4\mathbb{C}^{n+1}$ be a homogeneous polynomial. Let $A \subset \mathbb{P}^n$ be a finite set that computes T, such that $\ell(A) = 2n + 1$, A is in LGP, and A satisfies the minimality property. If there exists another subset $B \subset \mathbb{P}^n$ computing T with $\ell(B) \leq 2n + 1$, then the set $Z = A \cup B$ is contained in a rational normal curve of $\mathbb{P}^n$.*

Proof. First, note that Z satisfies Lemma 5.4 (iii) and, because of Proposition 6.1 it also satisfies Lemma 5.4 (i). Assume that Z has the Cayley–Bacharach property $CB(4)$. By applying Theorem 4.9 with $i = 4$ and $j = 2$ we get

$$Dh_Z(0) + Dh_Z(1) + Dh_Z(2) \leq Dh_Z(3) + Dh_Z(4) + Dh_Z(5). \tag{16}$$

Since Z contains $n + 1$ points of A in LGP, then $\langle Z\rangle = \mathbb{P}^n$, so that $Dh_Z(1) = n$. We claim that also (ii) of Lemma 5.4 holds. Indeed, if this is not the case, then $Dh_Z(2) > n$ and thus, by (16) and Lemma 2.12 part (v),

$$\ell(Z) = \sum_j Dh_Z(j) \geq \sum_{j=0}^{5} Dh_Z(j) > 4n + 2.$$

On the other hand, $\ell(Z) \leq \ell(A) + \ell(B) \leq 2\ell(A) = 4n + 2$, which leads to a contradiction. Therefore, all the assumptions of Lemma 5.4 hold, concluding the proof for this case.

It remains to prove that with our assumptions Z necessarily has the property $CB(4)$. Assume that Z does not satisfy $CB(4)$. Then, there exists $P \in Z$ such that $h^0(J_{Z\setminus\{P\}}(4)) > h^0(J_Z(4))$, i.e., $h_{Z\setminus\{P\}}(4) < h_Z(4)$. As $\ell(Z\setminus\{P\}) = \ell(Z) - 1$, by Lemma 2.12 (v), we have $Dh_{Z\setminus\{P\}}(j) < Dh_Z(j)$ for some $j \in \{1, 2, 3, 4\}$. Since $Dh_{Z\setminus\{P\}}(j) \leq Dh_Z(j)$ for all j, it follows that $Dh_{Z\setminus\{P\}}(j) = Dh_Z(j)$ for $j \geq 5$, i.e., by (3), $h^1_Z(4) = h^1_{Z\setminus\{P\}}(4)$.

Assume that $A \cap B = \emptyset$. If $P \in A$, then, by (15), $\dim(\langle v_4(A\setminus\{P\})\rangle \cap \langle v_4(B)\rangle) = \dim(\langle v_4(A)\rangle \cap \langle v_4(B)\rangle)$, so $\langle v_4(A\setminus\{P\})\rangle \cap \langle v_4(B)\rangle = \langle v_4(A)\rangle \cap \langle v_4(B)\rangle$.

Thus $A \backslash \{P\}$ computes T, which contradicts the hypothesis of minimality of A. Therefore, $P \in B \backslash A$. But now we can repeat the previous argument for A and $B \backslash \{P\}$, and we get that B is not minimal, which is a contradiction.

Hence, we can assume that $A \cap B \neq \emptyset$. We claim that

$$\langle v_4(A) \rangle \cap \langle v_4(B \backslash A) \rangle \neq \emptyset. \tag{17}$$

Indeed, let $s = \ell(A \cap B)$. We can renumber the elements of $v_4(A)$ and $v_4(B)$ in a way such that, for both sets, the first s comprise $v_4(A \cap B)$, i.e.,

$$v_4(A \cap B) = \{v_4(P_1), \ldots, v_4(P_s)\}.$$

Note that $\langle v_4(A \cap B) \rangle$ is a proper subset of $\langle v_4(A) \rangle \cap \langle v_4(B) \rangle$, since, for example, $T \in (\langle v_4(A) \rangle \cap \langle v_4(B) \rangle) \backslash \langle v_4(A \cap B) \rangle$, as A and B are minimal for T. Therefore,

$$\begin{aligned} T &= \alpha_1 v_4(P_1) + \cdots + \alpha_s v_4(P_s) + \alpha_{s+1} v_4(P_{s+1}) + \cdots + \alpha_{\ell(A)} v_4(P_{\ell(A)}) \\ &= \beta_1 v_4(P_1) + \cdots + \beta_s v_4(P_s) + \beta_{s+1} v_4(Q_{s+1}) + \cdots + \beta_{\ell(B)} v_4(Q_{\ell(B)}) \end{aligned}$$

with $\alpha_j, \beta_k \in \mathbb{C} \backslash \{0\}$ for any j, k and $P_j \in A \backslash B$, $Q_k \in B \backslash A$ for $j \in \{s+1, \ldots, \ell(A)\}$ and $k \in \{s+1, \ldots, \ell(B)\}$. It turns out that

$$\sum_{i=1}^{s} (\alpha_i - \beta_i) v_4(P_i) + \sum_{i=s+1}^{\ell(A)} \alpha_i v_4(P_i) - \sum_{i=s+1}^{\ell(B)} \beta_i v_4(Q_i) = 0.$$

Thus, the tensor

$$T' = \sum_{i=1}^{s} (\alpha_i - \beta_i) v_4(P_i) + \sum_{i=s+1}^{\ell(A)} \alpha_i v_4(P_i) = \sum_{i=s+1}^{\ell(B)} \beta_i v_4(Q_i) \tag{18}$$

is an element of $\langle v_4(A) \rangle \cap \langle v_4(B \backslash A) \rangle$, which implies (17). Now, if A is minimal for T', then we have two finite sets $A' = A$ and $B' = B \backslash A$ computing T' and such that $A' \cap B' = \emptyset$. Thus, by replacing T with T' and by arguing as in the case $A \cap B = \emptyset$, we get a contradiction because $A' \cup B' = A \cup B$ does not satisfy $CB(4)$. Therefore, we can assume that A is not minimal for T', i.e., there exists a proper subset A' of A such that $T' \in \langle v_4(A') \rangle$. Then some of the P_i's with $i \in \{1, \ldots, s\}$ does not appear in the decomposition of T', say P_1. So there exists $\gamma_i \in \mathbb{C} \backslash \{0\}$, such that

$$T' = \sum_{i=1}^{s} (\alpha_i - \beta_i) v_4(P_i) + \sum_{i=s+1}^{\ell(A)} \alpha_i v_4(P_i) = \gamma_2 v_4(P_2) + \cdots + \gamma_{\ell(A)} v_4(P_{\ell(A)}).$$

Since, by Remark 2.4, $v_4(P_1), \ldots, v_4(P_{\ell(A)})$ are linearly independent, it follows that $\alpha_1 = \beta_1$, $\gamma_i = \alpha_i - \beta_i$ for $i \in \{2, \ldots, s\}$ and $\gamma_i = \alpha_i$ for $i \in \{s+1, \ldots, \ell(A)\}$.

Therefore, by (18),

$$T' = \sum_{i=2}^{\ell(A)} \gamma_i v_4(P_i) = \sum_{i=s+1}^{\ell(B)} \beta_i v_4(Q_i),$$

so that T' has two different decompositions A', B' with, respectively, $\ell(A') = \ell(A) - 1 = 2n$ and $\ell(B') = \ell(B) - s \leq 2n$ summands. As $\ell(A'), \ell(B') \leq 2n$, this contradicts Theorem 3.2. □

As a consequence of Theorem 6.2 we get the following result.

THEOREM 6.3. *Fix a homogeneous polynomial $T \in S^4\mathbb{C}^{n+1}$ for which there exists a finite, minimal set A that computes T, such that $\ell(A) = 2n+1$ and A is in LGP. Then, the existence of a second set $B \subset \mathbb{P}^n$ that computes T and $\ell(B) \leq 2n+1$ implies that $\ell(B) = 2n+1$ and both A, B belong to a rational normal curve of $\mathbb{P}^n$.*

If we want to understand the identifiability of quadrics of rank $2n+1$, we should thus study the case where a minimal set that computes T lies in a rational normal curve.

PROPOSITION 6.4. *Fix a homogeneous polynomial $T \in S^4\mathbb{C}^{n+1}$ for which there exists a finite, minimal set A that computes T, such that $\ell(A) = 2n+1$ and A is contained in a rational normal curve $\Gamma \subset \mathbb{P}^n$. Then, there exists a positive dimensional family of finite sets $A_t \subset \Gamma$ such that*

(i) *$\ell(A_t) = 2n+1$;*
(ii) *$A_0 = A$;*
(iii) *$T \in \langle v_4(A_t) \rangle$, i.e., each A_t computes T.*

PROOF. Notice that, by Remark 2.5, A is in LGP. The curve Γ is the image of a Veronese map $\Gamma = v_n(\mathbb{P}^1)$, thus $\Gamma' = v_4(\Gamma) = v_{4n}(\mathbb{P}^1)$ is a rational normal curve in $\mathbb{P}^{4n}$. Therefore, T belongs to the secant variety $\sigma_{2n+1}(\Gamma')$. By Remark 2.6, $\sigma_{2n+1}(\Gamma')$ covers $\mathbb{P}^{4n}$. If $\Sigma_{2n+1}(\Gamma')$ denotes the abstract $(2n+1)$-secant variety of Γ', then the fibre of the $(2n+1)$-secant map

$$\pi_{2n+1} : \Sigma_{2n+1}(\Gamma') \to \mathbb{P}^{4n}$$

at T has dimension $\dim(\Sigma_{2n+1}(\Gamma')) - 4n = 2(2n+1) - 1 - 4n = 1$. □

REMARK 6.5. With the above notation, the tangent lines to Γ' at the points of $v_4(A)$ span a space of dimension at most $4n$. Therefore, the tangent spaces to $v_4(\mathbb{P}^n)$ at the same points span a space of dimension at most $(2n+1)(n+1) - 2$.

Summarizing, for quartics T in $\mathbb{P}^n$ which are computed by a set $A \subset \mathbb{P}^n$ in LGP and cardinality at most $2n+1$, we have that either:

(i) T has rank $\ell(A)$ and is identifiable; or
(ii) T has rank $2n+1$ and is computed by a 1-dimensional family of sets of cardinality $2n+1$ that includes A.

One can use *Terracini's test*, introduced in [COV17b, Lemma 6.5], to decide which of these two cases occurs.

Remark 6.6. The case $n = 2$, i.e., ternary quartics, is quite peculiar. Theorem 6.2 and Proposition 6.4 also apply in this case. However, note that any set A of cardinality 5 in $\mathbb{P}^2$ is contained in a rational normal curve, i.e., in a conic (the conic is irreducible when A is in LGP). Consequently, by our result, a general ternary quartic of rank $5 = 2n+1$ has infinitely many decompositions.

This is, in fact, a classic result [P03, C91], see also [IK99]; indeed, the 5-secant variety of the 4-Veronese embedding of $\mathbb{P}^2$ is *defective*: its dimension is smaller than the expected value. This forces general quartics of rank 5 to have infinitely many decompositions.

6.1. *The algorithm*

In an abuse of notation, let $v_4 : \mathbb{C}^{n+1} \to \mathbb{C}^{\binom{n+4}{4}}$ be the 4-fold symmetric tensor product. Then, $[v_4(\mathbf{m})] = v_4([\mathbf{m}])$. Given a length-$r$ symmetric tensor rank decomposition of a quartic

$$T = \sum_{i=1}^{r} v_4(P_i)$$

in the form of the collection of points $A = \{P_i = [\mathbf{m}_i]\}_{i=1}^{r} \subset \mathbb{P}^n$, we can apply the following algorithm for verifying that the given decomposition of T is identifiable:

S1. If $r > 2n+1$, the criterion cannot be applied.
S2. If $r < 2n+1$, use the reshaped Kruskal criterion from [COV17a, Section 6.2].
S3. If $r = 2n+1$, perform the next tests:
   1) *minimality test*: check that $\dim\langle v_4(\mathbf{m}_1), \ldots, v_4(\mathbf{m}_r)\rangle = r$;
   2) *Kruskal's test*: check that A is in LGP;
   3) *Terracini's test*: check that $\dim\langle \mathrm{T}_{\mathbf{m}_1} v_4(\mathbb{C}^{n+1}), \ldots, \mathrm{T}_{\mathbf{m}_r} v_4(\mathbb{C}^{n+1})\rangle = 2n^2 + 3n + 1$.

   If all these tests are successful, then T is of rank r and is r-identifiable.

An implementation of this algorithm is included in the ancillary Macaulay2 file `identifiabilityS4Cn.m2` to arXiv:1712.04211.

We note that the new criterion for $r = 2n+1$ is effective in the sense of [COV17b]. Indeed, quartics with $r = 2n+1$ are always generically r-identifiable [B05], and it is easy to verify that the conditions in tests 1, 2, and 3 are not satisfied precisely on a Zariski-closed strict subvariety of the r-secant variety of $v_4(\mathbb{P}^n)$.

6.2. *Examples*

We present some examples of identifiable and unidentifiable Waring decompositions in the original case $r = 2n + 1$.

An identifiable example. Consider $n = 4$ and $r = 2n + 1 = 9$. We generated a random collection of 9 points $A = \{P_i = [\mathbf{m}_i]\}_{i=1}^{9}$ in Macaulay2, where the vectors $\mathbf{m}_i \in \mathbb{N}^5$ had the following values in our experiment

$$M = [\mathbf{m}_i]_{i=1}^{9} = \begin{bmatrix} 0 & 1 & 1 & -3 & -5 & 2 & -1 & 2 & -1 \\ -2 & -1 & 2 & 0 & 1 & 2 & -4 & 3 & 1 \\ 2 & 0 & 5 & 1 & 4 & -5 & -1 & -3 & 4 \\ 1 & -5 & -1 & 3 & -2 & 3 & 5 & 2 & -3 \\ 1 & -3 & -2 & -5 & -4 & 3 & -2 & 1 & 4 \end{bmatrix}.$$

The minimality test shows that $\dim\langle v_4(A)\rangle = \operatorname{rank}([\mathbf{m}_i \otimes \mathbf{m}_i \otimes \mathbf{m}_i \otimes \mathbf{m}_i]_{i=1}^{9}) = 9$, which is as required. We then compute the rank of all 126 subsets of 5 columns of M. They are all of rank 5, so that the Kruskal rank is 5 and A is in LGP. Finally, we compute a basis B_i of the affine cone over the tangent space of the Veronese variety $v_4(\mathbb{P}^n)$ at one of the points in the cone over $v_4(P_i)$, and then compute $\dim\langle B_1, \ldots, B_9\rangle$. The computation reveals that it equals $45 = 2 \cdot 4^2 + 3 \cdot 4 + 1$, so that Terracini's test is also successful. We can conclude that $T = \sum_{i=1}^{9} v_4(\mathbf{m}_i)$ is (complex) identifiable.

A variation of Derksen's example. Consider the Vandermonde matrix

$$M = [\mathbf{m}_i]_{i=1}^{r} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_r \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_r^2 \\ \vdots & \vdots & \vdots & \\ \lambda_1^n & \lambda_2^n & \cdots & \lambda_r^n \end{bmatrix}.$$

Let $A = \{P_i = [\mathbf{m}_i]_i\}$ be the corresponding collection of points in $\mathbb{P}^n$. By construction, each of the points is in the image of $v_n(\mathbb{P}^1)$, i.e., they lie on a rational normal curve in $\mathbb{P}^n$. Hence, if $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ are pairwise distinct then A is in LGP. Taking $r = 2n + 1$, then $T = \sum_{i=1}^{r} v_4(\mathbf{m}_i)$ is not symmetric identifiable by Proposition 6.4, even though A is LGP.

We applied the criterion to the case $n = 4$, $r = 9$, and $\lambda_i = i - 1$ for $i = 1, \ldots, 9$. Running the algorithm, we find that both the minimality test and the Kruskal test are successful, consistent with the theory in [D13]. Our theory predicts that Terracini's test must fail, because T can now only be unidentifiable if there is a rational normal curve passing through the points A. Performing the computation, we find that the tangent spaces only span a space of dimension 44, one less than expected. Hence, Terracini's test fails, and we cannot conclude that T is identifiable.

ACKNOWLEDGMENTS. The first and second author were supported by the Italian PRIN 2015 – Geometry of Algebraic Varieties (B16J15002000005). The third author was supported by a Postdoctoral Fellowship of the Research Foundation–Flanders (FWO). The authors would like to thank Juan Migliore, Giorgio Ottaviani, and Maria Evelina Rossi for several fruitful discussions on the topics of the present research. We also thank the anonymous reviewer for remarks that helped us to further clarify the presented results.

REFERENCES

[AMR09] E. S. ALLMAN - C. MATIAS - J. A. RHODES, *Identifiability of parameters in latent structure models with many observed variables*, Ann. Stat. 37 (2009), 3099–3132.

[AGHKT14] A. ANANDKUMAR - R. GE - D. HSU - S. M. KAKADE - M. TELGARSKY, *Tensor decompositions for learning latent variable models*, J. Machine Learn. Res. 15 (2014), 2773–2832.

[B05] E. BALLICO, *On the weak non-defectivity of Veronese embeddings of projective spaces*, Central Eur. J. Math. 3 (2005), 183–187.

[BC12] E. BALLICO - L. CHIANTINI, *A criterion for detecting the identifiability of symmetric tensors of size three*, Diff. Geom. Applic. 30 (2012), 233–237.

[BC13] E. BALLICO - L. CHIANTINI, *Sets computing the symmetric tensor rank*, Mediterranean J. Math. 10 (2013), 643–654.

[BGM94] A. M. BIGATTI - A. V. GERAMITA - J. MIGLIORE, *Geometric consequences of extremal behavior in a theorem of Macaulay*, Trans. Amer. Math. Soc. 346 (1994), 203–235.

[C91] J. E. CAMPBELL, *Note on the maximum number of arbitrary points which can be double points on a curve, or surface, of any degree*, Messenger Math. XXI (1891–1892), 158–164.

[CC06] L. CHIANTINI - C. CILIBERTO, *On the concept of k-secant order of a variety*, J. London Math. Soc. 73 (2006), 436–454.

[COV17a] L. CHIANTINI - G. OTTAVIANI - N. VANNIEUWENHOVEN, *On generic identifiability of symmetric tensors of subgeneric rank*, Trans. Amer. Math. Soc. 369 (2017), 4021–4042.

[COV17b] L. CHIANTINI - G. OTTAVIANI - N. VANNIEUWENHOVEN, *Effective criteria for specific identifiability of tensors and forms*, SIAM J. Matrix Anal. Appl. 38 (2017), 656–681.

[D13] H. DERKSEN, *Kruskal's uniqueness inequality is sharp*, Linear Alg. Applic. 438 (2013), 708–712.

[GH78] P. GRIFFITHS - J. HARRIS, *Principles of Algebraic Geometry*, Wiley Interscience, New York NY (1978).

[GKR93] A. V. GERAMITA - M. KREUZER - L. ROBBIANO, *Cayley-Bacharach schemes and their canonical modules*, Trans. Amer. Math. Soc. 339 (1993), 443–452.

[H92] J. HARRIS, *Algebraic Geometry, a First Course*, Graduate Texts in Math., Springer, Berlin – New York NY (1992).

[IK99] A. IARROBINO - V. KANEV, *Power Sums, Gorenstein Algebras, and Determinantal Loci*, Lecture Notes in Mathematics 1721, Springer, Berlin – New York NY (1999).

[K77] J. B. KRUSKAL, *Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Linear Algebra Appl. 18 (1977), 95–138.

[MMS17] A. Massarenti - M. Mella - G. Staglianó, *Effective identifiability criteria for tensors and polynomials*, J. Symbolic Comput. 87 (2018), 227–237.

[P03] F. Palatini, *Sulla rappresentazione delle forme ternarie medinate la somma di potenze di forme lineari*, Rend. Accad. Lincei 12 (1903), 378–384.

[S17] Y. Shitov, *A counterexample to Comon's conjecture*, Preprint arXiv:1705.08740 (2017).

[SKLV17] V. Sharan - S. Kakade - P. Liang - G. Valiant, *Learning overcomplete HMMs*, Preprint arXiv:1711.02309 (2017).

Received 12 December 2017,
and in revised form 1 February 2018.

Elena Angelini
Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche
Università di Siena
Via Roma, 56
53100 Siena, Italy
elena.angelini@unisi.it

Luca Chiantini
Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche
Università di Siena
Via Roma, 56
53100 Siena, Italy
luca.chiantini@unisi.it

Nick Vannieuwenhoven
Department of Computer Science
KU Leuven
Celestijnenlaan 200A
3001 Leuven, Belgium
nick.vannieuwenhoven@cs.kuleuven.be