



**HAL**  
open science

# Sparse Moments of Univariate Step Functions and Allele Frequency Spectra

Zvi Rosen, Georgy Scholten, Cynthia Vinzant

► **To cite this version:**

Zvi Rosen, Georgy Scholten, Cynthia Vinzant. Sparse Moments of Univariate Step Functions and Allele Frequency Spectra. Vietnam Journal of Mathematics, 2022, 50 (2), pp.523-544. 10.1007/s10013-022-00548-5 . hal-04360559

**HAL Id: hal-04360559**

**<https://hal.science/hal-04360559>**

Submitted on 28 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SPARSE MOMENTS OF UNIVARIATE STEP FUNCTIONS AND ALLELE FREQUENCY SPECTRA

ZVI ROSEN, GEORGY SCHOLTEN, AND CYNTHIA VINZANT

ABSTRACT. We study the univariate moment problem of piecewise-constant density functions on the interval  $[0, 1]$  and its consequences for an inference problem in population genetics. We show that, up to closure, any collection of  $n$  moments is achieved by a step function with at most  $n - 1$  breakpoints and that this bound is tight. We use this to show that any point in the  $n$ th coalescence manifold in population genetics can be attained by a piecewise constant population history with at most  $n - 2$  changes. Both the moment cones and the coalescence manifold are projected spectrahedra and we describe the problem of finding a nearest point on them as a semidefinite program.

Given a finite collection  $A \subset \mathbb{N}$ , we consider the convex cone  $M(A)$  of all moments  $(m_a)_{a \in A}$  of the form  $m_a = \int x^a d\mu$  where  $\mu$  is a nonnegative Borel measure on the unit interval  $[0, 1]$ . For consecutive moments  $A = \{0, 1, 2, \dots, d\}$ , this is a classical object in analysis and real algebraic geometry. The problem of determining membership in the cone  $M(A)$  is known as the truncated Hausdorff moment problem. See, for example, [2, 4, 5, 8].

In this paper we study moments coming from piecewise-constant density functions with the idea of minimizing the number of pieces needed. Formally, we consider the set  $M_k(A)$  as the closure of the set of moments  $(m_a)_{a \in A}$  where  $m_a = \int_0^1 x^a f(x) dx$  and  $f$  is a nonnegative step function with at most  $k$  discontinuities. Our main theorem is the following:

**Main Theorem 1.**  $M_k(A) = M(A)$  if and only if  $k \geq |A| - 1$ .

This is the content of Theorem 1.5 and Corollary 1.12. The proof involves studying the convex algebraic boundary of these cones and in particular showing that they are simplicial (Corollary 1.9). When restricting to the moments of monotone density functions, only half as many break points are needed (see Propositions 2.4 and 2.5).

**Main Theorem 2.** *Every  $A$ -moment vector of a monotone density function is the limit of  $A$ -moments of monotone step functions with  $\leq k$  breakpoints if and only if  $k \geq \lfloor |A|/2 \rfloor$ .*

One of our motivations for studying this problem came from its relation to the coalescence manifold studied by [7]. The coalescence manifold  $\mathcal{C}_{n,k}$ , formally defined in Section 3, is a set of summary statistics in population genetics, derived from observing  $n$  genomes with a population history consisting of  $k + 1$  different population sizes. Our last main theorem, appearing as Theorem 3.5, is that the coalescence manifold  $\mathcal{C}_{n,k}$  coincides with an affine section of the moments  $M_k(A)$  for  $A = \{0, 2, \dots, \binom{n}{2} - 1\}$ .

**Main Theorem 3.** *The coalescence manifold  $\mathcal{C}_{n,k}$  is the intersection of  $M_k(A)$  with the affine hyperplane of points with coordinate sum equal to one for  $A = \{0, 2, \dots, \binom{n}{2} - 1\}$ . That is,  $\mathcal{C}_{n,k} = \{(m_a)_{a \in A} \in M_k(A) : \sum_{a \in A} m_a = 1\}$ .*

The authors in [7] show that the manifold  $\mathcal{C}_{n,k}$  stabilizes at  $k = 2n - 2$ , i.e.  $\mathcal{C}_{n,2n-2} = \mathcal{C}_{n,k}$  for all  $k \geq 2n - 2$ . Together, the main theorems above improve this bound by a factor of two, showing that the coalescence manifolds stabilize at  $k = n - 2$  and this bound is tight.

The connection with the moment problem also provides a description of  $\mathcal{C}_{n,n-2}$  as the projection of a spectrahedron. The problem of finding the nearest point in  $\mathcal{C}_{n,n-2}$  to a given point in  $\mathbb{R}^{n-1}$  can then be formulated as a semidefinite program.

This paper is organized as follows. In Section 1, we introduce formal definitions of the moment sets  $M_k(A)$ , study their convex and algebraic structure, and prove Theorem 1. In Section 2, we analyze analogous questions for moment problems coming from monotone step functions. The definitions and connections with the coalescence manifold  $\mathcal{C}_{n,k}$  are given in Section 3. Semidefinite descriptions of these sets are discussed in Section 4. Finally we end with a discussion of open problems surrounding these interesting sets in Section 5.

**Acknowledgements.** Author Rosen thanks Dr. Yun S. Song for introducing him to this topic. Authors Scholten and Vinzant were partially supported by NSF-DMS grants #1620014 and #1943363. This material is based upon work directly supported by the National Science Foundation Grant No. DMS-1926686, and indirectly supported by the National Science Foundation Grant No. CCF-1900460.

## 1. MOMENTS OF STEP FUNCTIONS

For  $k \in \mathbb{N}$ , let  $S_k$  denote the set of nonnegative step functions on  $[0, 1]$  of the form

$$(1) \quad f = y_1 \mathbf{1}_{[0, s_1]} + \sum_{i=2}^{k+1} y_i \mathbf{1}_{(s_{i-1}, s_i]},$$

where  $0 = s_0 < s_1 < \dots < s_k < s_{k+1} = 1$  and  $y_1, \dots, y_{k+1} \in \mathbb{R}_{\geq 0}$ . Note that:

- (1)  $S_k$  is invariant under nonnegative scaling,
- (2)  $S_k \subseteq S_\ell$  when  $k \leq \ell$ , and
- (3)  $S_k + S_\ell$ , defined as  $\{f + g \mid f \in S_k, g \in S_\ell\}$ , is a subset of  $S_{k+\ell}$ .

Elements of  $S_k$  define nonnegative measures on  $[0, 1]$ . We will be interested in the possible moments of these measures. Given a finite collection  $A \subset \mathbb{N}$ , we define to be the Euclidean closure of the set moments given by density functions in  $S_k$ :

$$M_k(A) = \overline{\left\{ \left( \int_0^1 x^a f(x) dx \right)_{a \in A} : f \in S_k \right\}}.$$

One important case is that of consecutive moments  $A = \{0, 1, \dots, d\}$ . For any finite collection  $A \subset \mathbb{N}$ , the moment cone  $M_k(A)$  can be expressed, up to closure, as the image of  $M_k(\{0, 1, \dots, \max(A)\})$  under the coordinate projection  $\pi_A : \mathbb{R}^{\max(A)+1} \rightarrow \mathbb{R}^A$  given by  $\pi_A(m_0, \dots, m_{\max(A)}) = (m_a)_{a \in A}$ .

**Remark 1.1.** By linearity of the integral, we see that  $M_k(A)$  inherits many properties of  $S_k$ . That is,  $M_k(A)$  is invariant under nonnegative scaling,  $M_k(A) \subseteq M_\ell(A)$  when  $k \leq \ell$  and  $M_k(A) + M_\ell(A) \subseteq M_{k+\ell}(A)$  (here, in the sense of the Minkowski sum), as desired.

We will be interested in comparing this to the full moment cone:

$$M(A) = \overline{\left\{ \left( \int_0^1 x^a d\mu \right)_{a \in A} : \mu \text{ is a nonnegative Borel measure on } [0, 1] \right\}}.$$

The cone  $M(A)$  is dual to the convex cone of univariate polynomials supported on  $A$  that are nonnegative on  $[0, 1]$ , as will be discussed below in Proposition 1.6.

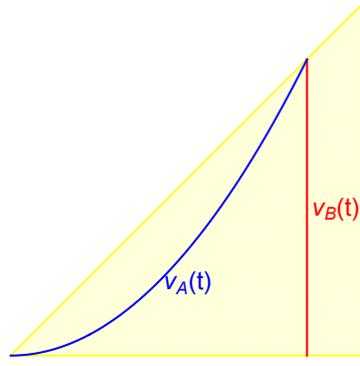


FIGURE 1. The cone  $M(A)$  along with the curve segments  $v_A([0, 1])$  and  $v_B([0, 1])$  for  $A = \{1, 2\}$  and  $B = \{0, 1\}$ .

When  $0 \in A$ , the closure in the definition of  $M(A)$  is not necessary, and the extreme rays of  $M(A)$  are come from point evaluations. That is, we can write  $M(A)$  as the conical hull of the image of  $[0, 1]$  under the corresponding moment map:

$$M(A) = \text{conicalHull} \{v_A(t) : t \in [0, 1]\} \quad \text{where} \quad v_A(t) = (t^a)_{a \in A}.$$

See, for example, [8, Prop. 10.5].

When  $0 \notin A$ , this equality only holds up to closure, as the curve parametrized by  $v_A(t)$  includes the origin. In this case,  $M(A) = \overline{\text{conicalHull} \{v_A(t) : t \in [0, 1]\}}$ . As we will see below, then we can still write  $M(A)$  as the conical hull of a curve segment. Specifically,  $M(A) = \text{conicalHull} \{v_B(t) : t \in [0, 1]\}$  where  $B = \{a - \min(A) : a \in A\}$ .

**Lemma 1.2.** *If  $A \subset \mathbb{N}$  is finite and  $B = \{a - \min(A) : a \in A\}$ , then  $M(A) = M(B)$ .*

*Proof.* For  $t \in (0, 1]$ , the point  $v_A(t)$  can be rewritten as  $t^{\min(A)} v_B(t)$ , a scalar multiple of  $v_B(t)$ . It follows that the conical hulls of  $\{v_A(t) : t \in (0, 1]\}$  and  $\{v_B(t) : t \in (0, 1]\}$  are equal. We observe that the extreme ray  $v_B(0)$  of  $M(B)$  can be attained in the closure of  $M(A)$  as the limit of the moment of the step function  $f = \epsilon^{-(\min(A)+1)} \mathbf{1}_{[0, \epsilon]}$  as  $\epsilon$  goes to zero:

$$\lim_{\epsilon \rightarrow 0} \int_0^1 x^a f(x) dx = \lim_{\epsilon \rightarrow 0} \epsilon^{-(\min(A)+1)} \int_0^\epsilon x^a dx = \lim_{\epsilon \rightarrow 0} \frac{\epsilon^{a-\min(A)}}{a+1} = \begin{cases} \frac{1}{a+1} & \text{if } a = \min(A) \\ 0 & \text{otherwise.} \end{cases}$$

It follows that  $v_B(0)$  belongs to  $M(A)$ . Since  $M(B)$  can be written as the union of the cone over  $v_A(t)$  for  $t \in (0, 1]$  and the ray over  $v_B(0)$ , the equality between the two cones ensues:

$$M(A) = \overline{\text{conicalHull} \{v_A(t) : t \in [0, 1]\}} = M(B). \quad \square$$

**Example 1.3.** Consider  $A = \{1, 2\}$  and  $B = \{0, 1\}$ . Then

$$M(B) = \text{conicalHull}\{(1, t) : t \in [0, 1]\} = \overline{\text{conicalHull}\{(t, t^2) : t \in [0, 1]\}} = M(A).$$

Here we see the need for taking closures when  $0 \notin A$ . The point  $(1, 0) = v_B(0)$  is not contained in the the conical hull of the curve segment  $\{(t, t^2) : t \in [0, 1]\}$  but is contained in its closure. See Figure 1. In this case, the boundary of  $M(A) = M(B)$  consists of scalar multiples of  $v_B(0)$  and  $v_B(1)$ , both of which belong to  $M_1(A)$ , by Proposition 1.4 below. Arguments below will then show that  $M(A) = M_1(A)$ .

**Proposition 1.4.** *Let  $A \subset \mathbb{N}$  be finite and let  $B = \{a - \min(A) : a \in A\}$ . The points  $v_B(0)$  and  $v_B(1)$  belong to  $M_1(A)$  and for every  $t \in (0, 1)$ ,  $v_B(t)$  belongs to  $M_2(A)$ .*

*Proof.* For  $0 < t < 1$ , and  $0 < \epsilon < 1 - t$  consider the step function  $f = \epsilon^{-1}t^{-\min(A)}\mathbf{1}_{(t, t+\epsilon]}$  in  $S_2$ . By continuity, the integral  $\int_0^1 x^a f(x) dx$  limits to  $t^{a-\min(A)}$  as  $\epsilon \rightarrow 0$ , thus  $M_2(A)$  contains the limit point  $v_B(t) = (t^b)_{b \in B}$ . Similarly, the limit as  $\epsilon \rightarrow 0$  of the  $A$ -moment vectors of step functions  $f = (\min(A) + 1)\epsilon^{-1-\min(A)}\mathbf{1}_{[0, \epsilon]}$  and  $f = \epsilon^{-1}\mathbf{1}_{(1-\epsilon, 1]}$  in  $S_1$  are  $v_B(0)$  and  $v_B(1)$ , respectively. Therefore these vectors belong to  $M_1(A)$ .  $\square$

A corollary of this statement is that  $M_k(A) = M(A)$  for  $k = 2|A|$ . By Carathéodory's Theorem, any point in  $M(A)$  is in the conical hull of at most  $|A|$  points of the form  $(t^a)_{a \in A}$  where  $t \in [0, 1]$ , each of which belongs to  $M_2(A)$  by Proposition 1.4. By Remark 1.1, the sum of  $|A|$  elements from  $M_2(A)$  belongs to  $M_{2|A|}(A)$ , giving  $M(A) \subseteq M_{2|A|}(A)$ . In fact,  $M_k(A)$  fills out the whole moment cone much sooner:

**Theorem 1.5.** *If  $k \geq |A| - 1$ ,  $M_k(A) = M(A)$ .*

The proof of this theorem relies on understanding the points on the boundary of  $M(A)$ .

**Proposition 1.6.** *Let  $A \subset \mathbb{N}$  be finite with  $0 \in A$ . If  $\mathbf{m} = (m_a)_{a \in A}$  belongs to the Euclidean boundary of  $M(A)$ , then any representing measure  $\mu$  on  $[0, 1]$  with  $m_a = \int x^a d\mu$  has finite support. Specifically, the support of  $\mu$  is a subset of the roots contained in  $[0, 1]$  of a polynomial nonnegative on  $[0, 1]$  and of the form  $p(x) = \sum_{a \in A} p_a x^a$ . The vector  $\mathbf{m}$  is a conic combination of the vectors  $v_A(r)$  where  $r$  ranges over the roots of  $p$ .*

*Proof.* Let  $\ell : \mathbb{R}^A \rightarrow \mathbb{R}$  be a linear function  $\ell(\mathbf{v}) = \sum_{a \in A} p_a v_a$  defining a supporting hyperplane of  $M(A)$  at  $\mathbf{m}$ . That is,  $\ell(\mathbf{v}) \geq 0$  for all  $\mathbf{v} \in M(A)$  and  $\ell(\mathbf{m}) = 0$ . Consider the polynomial  $p(x) = \ell(v_A(x)) = \sum_{a \in A} p_a x^a$ . Since  $v_A(t) \in M(A)$  for all  $t \in [0, 1]$ ,  $p$  is nonnegative on  $[0, 1]$ . Furthermore, for any measure  $\mu$  with moments  $\mathbf{m}$ ,

$$\int p(x) d\mu = \sum_{a \in A} p_a m_a = \ell(\mathbf{m}) = 0.$$

The measure  $\mu$  is nonnegative and the polynomial  $p$  is nonnegative on  $[0, 1]$ . From this we see that the support of the measure  $\mu$  must be contained in the (finite) set of roots  $R$  of  $p(x)$ . Specifically,  $\mu = \sum_{r \in R} w_r \delta_r$  for some  $w_r \in \mathbb{R}_{\geq 0}$ ; therefore,  $\mathbf{m} = \sum_{r \in R} w_r v_A(r)$ .  $\square$

*Proof of Theorem 1.5.* First, consider a point  $\mathbf{m}$  in the boundary of  $M(A)$ . By Lemma 1.2,  $M(A) = M(B)$  where  $B = \{a - \min(A) : a \in A\}$ , and so  $\mathbf{m}$  also belongs to the boundary of  $M(B)$ . By Proposition 1.6,  $\mathbf{m}$  is the vector of  $B$ -moments of a measure  $\mu$  supported on the roots of a nonnegative polynomial on  $[0, 1]$  of the form  $p(x) = \sum_{a \in B} p_a x^a$ . Let  $b$  be the number of distinct roots of  $p$  in the set  $\{0, 1\}$  and  $i$  be the number of distinct roots in of  $p$  in the open interval  $(0, 1)$ . Then  $\mathbf{m}$  is in the conical hull of the  $b + i$  points given by  $v_B(r)$  where  $r$  ranges over these roots. By Proposition 1.4,  $\mathbf{m}$  belongs to  $M_k(A)$  for  $k = b + 2i$ .

By Descartes' rule of signs, the number of positive roots of  $p$ , counting multiplicity, is at most the number of sign changes in the list of coefficients  $\{p_a\}_{a \in B}$ . If  $p_0 \neq 0$ , then  $p$  has at most  $|B| - 1$  roots in  $\mathbb{R}_{>0}$ . If  $p_0 = 0$ , then  $p$  is the sum of at most  $|B| - 1$  nonzero terms and its number of roots in  $\mathbb{R}_{>0}$  must be smaller or equal to  $|B| - 2$ . Note that every root of  $p$  in  $(0, 1)$  must have even multiplicity greater or equal to 2. All together this gives  $b + 2i \leq |B| - 1 = |A| - 1$ .

Now consider  $\mathbf{m}$  in the interior of  $M(A)$ . Let  $\mathbf{c} = (1/(a+1))_{a \in A} \in M_0(A)$  denote the vector obtained by integrating against the constant step function of height one. Let  $\lambda^*$  be the maximum value of  $\lambda \in \mathbb{R}$  for which  $\mathbf{m} - \lambda \mathbf{c}$  belongs to  $M(A)$ . From  $\mathbf{m} \in M(A)$ , we see that  $\lambda^* \geq 0$ . Moreover, since  $M(A)$  is pointed,  $-\mathbf{c}$  does not belong to  $M(A)$ , meaning that for sufficiently large  $\lambda$ ,  $\mathbf{m} - \lambda \mathbf{c}$  does not belong to  $M(A)$ . Since  $M(A)$  is closed, it follows that such a maximum  $\lambda^*$  must exist.

The point  $\mathbf{m} - \lambda^* \mathbf{c}$  belongs to the boundary of  $M(A)$ . By the arguments above,  $\mathbf{m} - \lambda^* \mathbf{c}$  belongs to  $M_k(A)$  for  $k \geq |A|-1$ . Since  $\mathbf{c} \in M_0(A)$  and

$$\mathbf{m} = (\mathbf{m} - \lambda^* \mathbf{c}) + \lambda^* \mathbf{c},$$

the point  $\mathbf{m}$  also belongs to  $M_k(A)$  for  $k \geq |A|-1$ . □

**Remark 1.7.** It follows from the proof of Theorem 1.5 that for all  $k \geq 0$ ,  $M_k(A)$  is star convex with respect to the point  $\mathbf{c} = (1/(a+1))_{a \in A}$ , the  $A$ -moment of the constant function. Indeed, since  $\mathbf{c}$  belongs to  $M_0(A)$ ,  $\lambda \mathbf{c} + M_k(A) \subseteq M_k(A)$  for all  $\lambda \geq 0$ .

We can go further in characterizing the facial structure of the boundary of  $M(A)$ . Through a connection to Schur polynomials, we can deduce linear independence among sets of points from the curve of the correct size.

**Proposition 1.8.** *For a collection  $A$  of integers  $0 = a_1 < a_2 < \dots < a_n$  and any real values  $0 \leq r_1 < r_2 < \dots < r_n \leq 1$ , the determinant of the matrix  $\mathcal{S}_A$  is strictly positive, where*

$$\mathcal{S}_A(\mathbf{r}) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ r_1^{a_2} & r_2^{a_2} & \dots & r_n^{a_2} \\ \vdots & \vdots & & \vdots \\ r_1^{a_n} & r_2^{a_n} & \dots & r_n^{a_n} \end{pmatrix}.$$

*Proof.* By the bialternant formula for Schur polynomials, the determinant of the matrix  $\mathcal{S}_A(\mathbf{r})$  can be expressed as

$$(2) \quad \det(\mathcal{S}_A) = \left( \prod_{1 \leq i < j \leq n} (r_j - r_i) \right) \mathfrak{s}_\lambda(r_1, \dots, r_n) \text{ for } \lambda = (a_n - (n-1), a_{n-1} - (n-2), \dots, a_1),$$

where  $\mathfrak{s}_\lambda(x_1, \dots, x_n)$  denotes the Schur polynomial associated to the partition  $\lambda$ . By definition, the Schur polynomial  $\mathfrak{s}_\lambda(x_1, \dots, x_n)$  is the sum of monomials  $\mathbf{x}^T$  over all semistandard Young tableaux  $T$  of shape  $\lambda$ . One can observe, either from expanding the determinant of  $\mathcal{S}_A(r)$  along the first column, or by filling out a semistandard Young Tableau of shape  $\lambda$  without using the number 1, that  $x_1$  does not appear in all the monomials of the determinant of  $\mathcal{S}_A$ . It follows that  $\det(\mathcal{S}_A)$  is strictly positive for any  $0 \leq r_1 < r_2 < \dots < r_n \leq 1$ . □

**Corollary 1.9.** *All proper faces of  $M(A)$  are simplicial.*

*Proof.* By Lemma 1.2, we can assume that  $0 \in A$ . Recall that  $M(A)$  is the conical hull over the curve segment  $\{v_A(t) : t \in [0, 1]\}$  and any proper face  $F$  of this cone can be expressed as the conical hull of some points  $v_A(r_1), \dots, v_A(r_k)$  where  $0 \leq r_1 < r_2 < \dots < r_k \leq 1$ . If  $k > \dim(F)$ , then there is a subset of these points of size  $\dim(F) + 1 \leq n$ , which necessarily lie in  $F$  and are therefore linearly dependent, contradicting Proposition 1.8. Therefore  $k = \dim(F)$  and  $F$  is simplicial. □

This lemma lets us assign an *index* to points on the boundary of  $M(A)$ , following [8, Ch. 10.2]. Let  $\mathbf{m}$  be a point on the boundary of  $M(A)$ . By Corollary 1.9, there is a *unique* representation of  $\mathbf{m}$  as  $\sum_{j=1}^k w_j v_A(r_j)$  where  $0 \leq r_1 < \dots < r_k \leq 1$  and  $w_1, \dots, w_k \in \mathbb{R}_{>0}$ . We define the **index** of  $\mathbf{m}$ , denoted  $\text{ind}(\mathbf{m})$ , to be  $b + 2i$  where  $b = \#\{j : r_j \in \{0, 1\}\}$  and  $i = \#\{j : r_j \in (0, 1)\}$ . By Proposition 1.4, any point  $\mathbf{m}$  on the boundary of  $M(A)$  belongs to  $M_{\text{ind}(\mathbf{m})}(A)$ .

To prove the converse, we must rule out the possibility that  $\mathbf{m} \in M_k(A)$  for  $k < \text{ind}(\mathbf{m})$ . In other words, it is impossible to approach a point  $\mathbf{m}$  on the boundary of  $M(A)$  with moment vectors of step functions with fewer breakpoints than expected.

**Lemma 1.10.** *Let  $\mathbf{m}$  be a point on the boundary of  $M(A)$ . For  $k < \text{ind}(\mathbf{m})$ ,  $\mathbf{m} \notin M_k(A)$ . That is, if  $\mathbf{m} \in M_k(A)$ , then  $\text{ind}(\mathbf{m}) \leq k$ .*

*Proof.* Note that for any non-zero point  $\mathbf{m}$  in  $M(A)$ ,  $m_0 > 0$  and so we can rescale  $\mathbf{m}$  to have  $m_0 = 1$ . We will write  $M_k(A) \cap \{m_0 = 1\}$  as the image of a compact polytope under a polynomial map and check that any point  $\mathbf{m}$  in the image of this map and the boundary of  $M(A)$  has index  $\leq k$ .

Any function  $f \in S_k$  can be written as  $f = y_1 \mathbf{1}_{[0, s_1]} + \sum_{i=2}^{k+1} y_i \mathbf{1}_{(s_{i-1}, s_i]}$  for some values  $0 = s_0 < s_1 < \dots < s_k < s_{k+1} = 1$  and  $y_i \geq 0$  for all  $i$ . We now introduce transformed  $w$ -coordinates by letting  $w_i = y_i(s_i - s_{i-1})$  denote the area  $\int_{s_{i-1}}^{s_i} f(x) dx$ . The corresponding moment in  $M_k(A)$  is given by the image of the point  $(\mathbf{s}, \mathbf{w}) = (s_1, \dots, s_k, w_1, \dots, w_{k+1})$  under the polynomial map

$$(3) \quad \mu_A(\mathbf{s}, \mathbf{w}) = \left( \sum_{i=1}^{k+1} y_i \frac{s_i^{a+1} - s_{i-1}^{a+1}}{a+1} \right)_{a \in A} = \left( \sum_{i=1}^{k+1} w_i \frac{(s_i^a + s_i^{a-1} s_{i-1} + \dots + s_{i-1}^a)}{a+1} \right)_{a \in A}.$$

Note that the constraint that  $m_0 = 1$  translates into  $\sum_i w_i = 1$ . Consider the polytope

$$(4) \quad P = \left\{ (\mathbf{s}, \mathbf{w}) \in \mathbb{R}^k \times \mathbb{R}^{k+1} \text{ such that } 0 \leq s_1 \leq \dots \leq s_k \leq 1, w_i \geq 0, \sum_{i=1}^{k+1} w_i = 1 \right\},$$

which is a product of two simplices of dimension  $k$ . The moments of step functions  $f \in S_k$  with  $\int f(x) dx = 1$  is the image under  $\mu_A$  of the set of points  $(\mathbf{s}, \mathbf{w}) \in P$  with distinct  $0 < s_1 < \dots < s_k < 1$ . Its closure is  $M_k(A) \cap \{m_0 = 1\}$ , which necessarily coincides with the image of  $P$  under  $\mu_A$ , as the image of a compact set under a continuous map is closed.

If  $w_i > 0$  and  $s_{i-1} < s_i$  for some  $i$ , then  $\mu_A(\mathbf{s}, \mathbf{w})$  has a representing measure whose support includes the interval  $(s_{i-1}, s_i]$  and is therefore not finite. Then by Proposition 1.6,  $\mathbf{m}$  belongs to the interior of  $M(A)$ .

Suppose the point  $\mathbf{m}$  belongs to  $M_k(A)$ . Then  $\mathbf{m} = \mu_A(\mathbf{s}, \mathbf{w})$  for some  $(\mathbf{s}, \mathbf{w}) \in P$ . Let  $I$  denote the collection of indices  $1 \leq i \leq k$  for which  $w_i > 0$ . If  $\mathbf{m}$  belongs to the boundary of  $M(A)$ ,  $s_{i-1} = s_i$  for all  $i \in I$ . Then

$$\mathbf{m} = \mu_A(\mathbf{s}, \mathbf{w}) = \sum_{i \in I} w_i v_A(s_i).$$

We can bound  $\text{ind}(\mathbf{m})$  by bounding the number of distinct values of  $s_i$  that appear. For each  $i \in I$  with  $s_i \in (0, 1)$ ,  $s_i$  equals  $s_{i-1}$ , hence there are at least two indices  $j$  in  $\{1, \dots, k\}$  for which  $s_j = s_i$ . Trivially, if  $s_i \in \{0, 1\}$ , there is at least one  $j \in \{1, \dots, k\}$  such that  $s_j = s_i$ .

Together, these show that

$$\text{ind}(\mathbf{m}) = \#\{s_i \in \{0, 1\} : i \in I\} + 2 \cdot (\#\{s_i \in (0, 1) : i \in I\}) \leq k. \quad \square$$

**Lemma 1.11.** *The intersection of  $M_k(A)$  with the Euclidean boundary of  $M(A)$  is a semi-algebraic set of dimension  $\leq k$ .*

*Proof.* By Lemma 1.10, the intersection of  $M_k(A)$  with the Euclidean boundary of  $M(A)$  is the set of boundary points of index  $\leq k$ . We can parametrize this as the union of the semialgebraic sets:

$$\bigcup_{\sigma \in \{0,1\}^2} \left\{ \sum_{j=1}^{\ell} w_j v_A(r_j) + w_{\ell+\sigma_1} v_A(0) + w_{\ell+\sigma_1+\sigma_2} v_A(1) : \mathbf{r} \in (0, 1)^\ell, \mathbf{w} \in (\mathbb{R}_{>0})^{\ell+\sigma_1+\sigma_2} \right\},$$

where in each set,  $\ell$  is chosen so that  $2\ell + \sigma_1 + \sigma_2 \leq k$ . Here we use  $\mathbf{r}$  to denote the vector  $(r_j)_j$  and  $\mathbf{w}$  for the vector  $(w_j)_j$ . Note that each set is the image of  $(0, 1)^n \times (\mathbb{R}_{>0})^m$  under a polynomial map where  $n + m \leq k$  and therefore has dimension  $\leq k$ .  $\square$

**Corollary 1.12.** *If  $k < |A| - 1$ ,  $M_k(A) \neq M(A)$ .*

*Proof.* The cone  $M(A)$  is full-dimensional in  $\mathbb{R}^{|A|}$ , in consequence, the cone's boundary is a hypersurface of dimension  $|A| - 1$ . By Lemma 1.11, the dimension of the intersection of  $M_k(A)$  with the boundary of  $M(A)$  has dimension  $\leq k$ , so for  $k < |A| - 1$ , this cannot be the entire boundary of  $M(A)$ .  $\square$

**Example 1.13.** For  $A = \{0, 2, 5\}$ , Theorem 1.5 and Corollary 1.12 imply that  $M_k(A) = M_2(A)$  for all  $k \geq 2$  but not for  $k = 1$ . Affine transformations of their intersections with the affine hyperplane  $\{m_0 = 1\}$  are shown in Figure 2. See also [7, Figure 6]. The intersection of  $M_1(A)$  with the boundary of  $M_2(A)$  consists of just two rays, which appear as points in the hyperplane  $\{m_0 = 1\}$ . The set  $M_1(A)$  consists of moments of functions with just one breakpoint. Step functions with one breakpoint and total mass one can be parametrized by  $f = \frac{w}{s} \mathbf{1}_{[0,s]} + \frac{1-w}{1-s} \mathbf{1}_{(s,1]}$  for  $s \in (0, 1)$  and  $w \in [0, 1]$ . Note that fixing  $w$  and taking the limit as  $s \rightarrow 0$  gives a weighted sum of a point mass at zero and the constant function  $w\delta_0 + (1 - w)\mathbf{1}_{(0,1]}$ . Similarly  $s \rightarrow 1$  gives  $w\mathbf{1}_{[0,1]} + (1 - w)\delta_1$ .

For  $s = w \in [0, 1]$ , the corresponding step function is constant, i.e.  $f = \mathbf{1}_{[0,1]}$  and the moment map sends this line segment to a single point. However, away from this line, the moment map is a homeomorphism to its image in  $M_1(\{0, 2, 5\})$ .

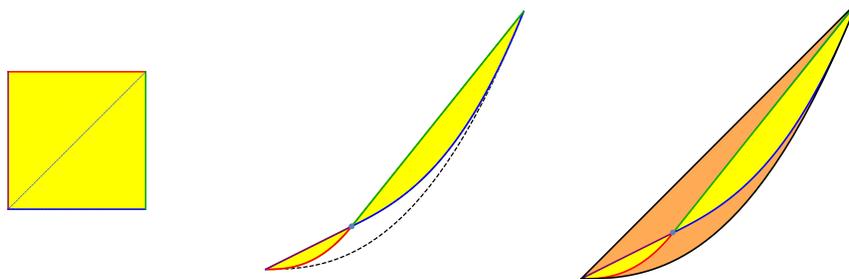


FIGURE 2. The parameter space of  $M_1(A)$  and  $M_1(A), M_2(A)$  for  $A = \{0, 2, 5\}$ .

**Example 1.14.** Consider  $A = \{0, 2, 5, 9\}$ . To visualize the moment sets  $M_k(A)$ , we consider their intersections with the affine hyperplane  $\{m_0 = 1\}$ . Affine transformations of these intersections are shown in Figure 3. Note that the step functions with at most one breakpoint and total mass one can be written as  $\lambda \mathbf{1}_{[0,1]} + (1 - \lambda) \frac{1}{s} \mathbf{1}_{[0,s]}$  or  $\lambda \mathbf{1}_{[0,1]} + (1 - \lambda) \frac{1}{1-s} \mathbf{1}_{(s,1]}$  where  $\lambda \in [0, 1]$ . The result is a two-dimensional surface in the plane  $\{m_0 = 1\}$ . The set  $M_2(A)$  is full-dimensional, but does not fill up all of  $M(A)$ . As promised by Lemma 1.11, the intersection  $M_2(A)$  with the boundary of  $M(A)$  has dimension  $\leq 2$ , so its image in  $\{m_0 = 1\}$  has dimension  $\leq 1$ . Indeed, we see this intersection is given by the curve parametrized by  $(t^2, t^5, t^9)$  for  $t \in [0, 1]$  and the line segment between its end points  $(0, 0, 0)$  and  $(1, 1, 1)$ . Finally, by Theorem 1.5,  $M_3(A)$  is the full cone  $M(A)$ . Points on the boundary of  $M(A)$  have index  $\leq 3$ , and so have one of the two forms  $w_0 v_A(0) + w_r v_A(r)$  or  $w_1 v_A(1) + w_r v_A(r)$  where  $r \in [0, 1]$ ,  $w_0, w_1, w_r \in \mathbb{R}_{\geq 0}$ .

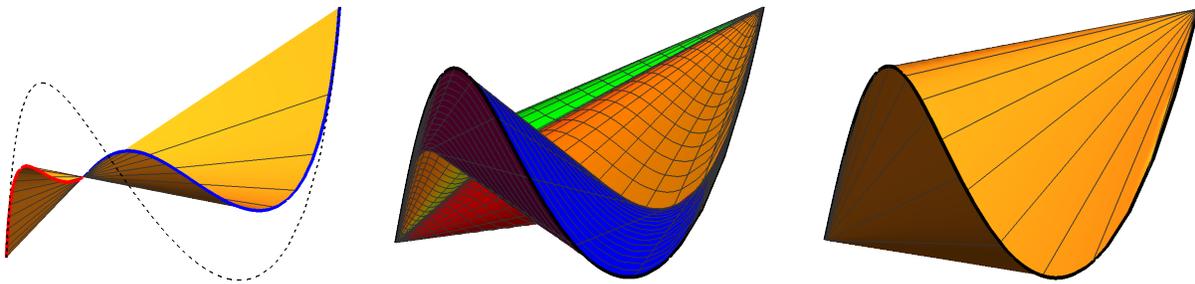


FIGURE 3. The sets  $M_1(A)$ ,  $M_2(A)$ ,  $M_3(A)$  in  $\{m_0 = 1\}$  for  $A = \{0, 2, 5, 9\}$ .

We will now take a closer look at this example, paying close attention to the boundary of the domain and its image in  $M_2(A)$ . The domain is the set of step functions with two breakpoints, so it lives in the product of two 2-simplices

$$P = \{(s_1, s_2) : 0 \leq s_1 \leq s_2 \leq 1\} \times \{(w_1, w_2, w_3) \in \mathbb{R}_{\geq 0}^3 : w_1 + w_2 + w_3 = 1\}.$$

Here the  $s$  variables parametrize the two breakpoints and the  $w$  variables parametrize the proportion of mass in each piece. So the domain is four-dimensional while the image has dimension three. Therefore the generic fiber of the moment map has dimension one.

We can however obtain a generically finite-to-one map by restricting to facets of the polytope  $P$ . The boundary is composed of six triangular prisms given by  $s_1 = 0$ ,  $s_2 = 1$ ,  $s_1 = s_2$ ,  $w_1 = 0$ ,  $w_2 = 0$ , and  $w_3 = 0$ . We can visualize this by way of a Schlegel diagram via one of its facets, seen in Figure 4.

In observing the moment map restricted to the boundary, we make a number of observations. There are four 2-faces that the moment map collapses to a curve, namely the faces given by  $0 = s_1 = s_2$ ,  $s_1 = s_2 = 1$ ,  $w_1 = w_2 = 0$ , and  $w_2 = w_3 = 0$ . In addition to these 2-faces, the intersections of the facets of  $P$  with the hypersurfaces given by  $y_1 = y_2$  and  $y_2 = y_3$  sometimes drop dimension under the moment map  $\mu_A$ . In the  $(\mathbf{s}, \mathbf{w})$  variables, these correspond to surfaces  $(s_2 - s_1)w_1 = s_1 w_2$  and  $(1 - s_2)w_2 = s_2 w_3$ , respectively. For example, in each of the facets  $s_1 = 0$  and  $w_1 = 0$ , the equation  $(1 - s_2)w_2 = s_2 w_3$  cuts out a surface whose image under  $\mu_A$  is a curve. For the face  $s_1 = 0$ , the moments of this surface collapse to the line segment connecting the constant population and the point mass at 0 and for  $w_1 = 0$ , the image of this surface collapses to the curve segment of moments of step functions of a single step with  $w_1 = 0$ . Similarly, the faces  $s_2 = 1$  and  $w_3 = 0$  each

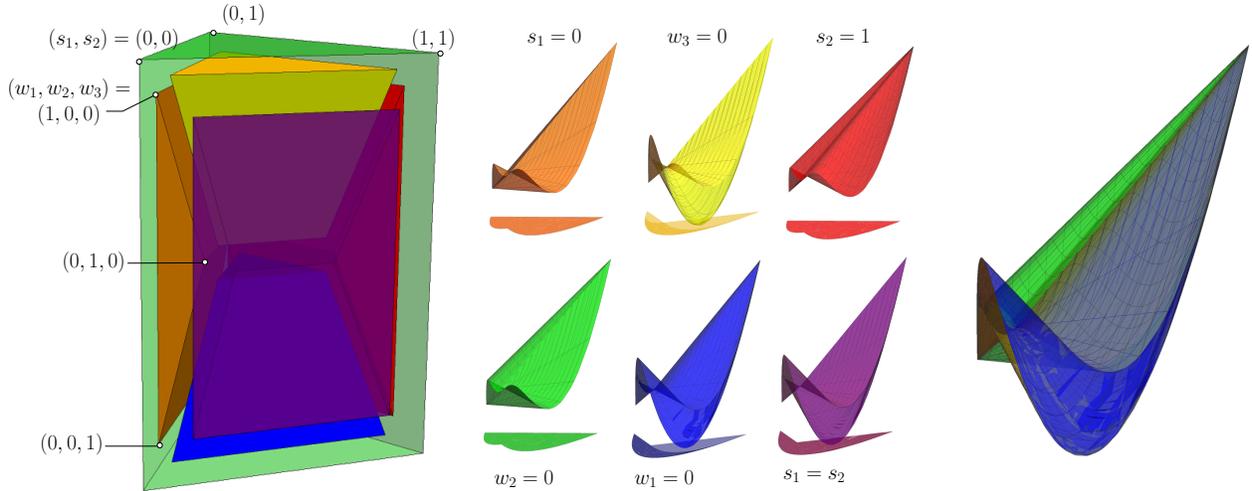


FIGURE 4. Schlegel diagram for the boundary of  $P$ , together with the image of each facet under the moment map.

contain a two-dimensional surface cut out by  $s_1 w_2 = (1 - s_1) w_1$  whose moments collapse to a line segment—from the constant population to the point mass at 1 for  $s_2 = 1$  and a curve segment of moments of a single step with  $w_3 = 0$ .

Aside from these subsurfaces, the map on the boundary  $\partial P$  is locally nondegenerate. Interestingly, the images of these facets can overlap in full-dimensional sets. One consequence is that the fibers of the moment map can be disconnected.

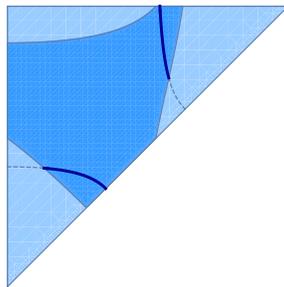


FIGURE 5. A disconnected fiber of  $\mu_A$  for  $A = \{0, 2, 5, 9\}$ .

For example, the point  $(m_0, m_2, m_5, m_9) = (1, 0.164, 0.054, 0.031)$  belongs to  $M_2(A)$  for  $A = \{0, 2, 5, 9\}$ . Its fiber under  $\mu_A$  is a curve in the four-dimensional polytope  $P$  from (4). Figure 5 shows the  $(s_1, s_2)$ -coordinates of this curve. In particular, this fiber has at least two connected components. In a lighter shade is the two-dimensional fiber of the point  $(m_0, m_2, m_5) = (1, 0.164, 0.054)$  under the corresponding map for  $\{0, 2, 5\}$ .

## 2. INCREASING AND DECREASING STEP FUNCTIONS

In this section, we study the moment cones of non-negative monotone functions on the unit interval  $[0, 1]$ . We define the increasing and decreasing moment cones

$$M^\uparrow(A) = \overline{\left\{ \left( \int_0^1 x^a f(x) dx \right)_{a \in A} : f \text{ is nonnegative and } \textit{increasing} \text{ on } [0, 1] \right\}} \text{ and}$$

$$M^\downarrow(A) = \overline{\left\{ \left( \int_0^1 x^a f(x) dx \right)_{a \in A} : f \text{ is nonnegative and } \textit{decreasing} \text{ on } [0, 1] \right\}}.$$

Recall that if a function  $f : [0, 1] \rightarrow \mathbb{R}$  is monotone, then it is automatically Borel-measurable. As in the non-monotone case, all of these moment vectors can be achieved as a limit of moments of step functions with a bounded number of steps. For  $k \in \mathbb{N}$ , let  $S_k^\uparrow$  denote the set of nonnegative, *increasing* step functions on  $[0, 1]$  with at most  $k$  discontinuities. Similarly, let  $S_k^\downarrow$  denote the analogous set of *decreasing* step functions. This corresponds to requiring  $y_1 \leq y_2 \leq \dots \leq y_{k+1}$  or  $y_1 \geq y_2 \geq \dots \geq y_{k+1}$  in (1).

Similarly, for finite  $A \subset \mathbb{N}$ , we consider the  $A$ -moments of these step functions,

$$M_k^\square(A) = \overline{\left\{ \left( \int_0^1 x^a f(x) dx \right)_{a \in A} : f \in S_k^\square \right\}} \text{ for } \square \in \{\uparrow, \downarrow\}.$$

Just as with  $M_k(A)$ , we see that the set  $M_k^\square(A)$  is invariant under nonnegative scaling,  $M_k^\square(A) \subseteq M_\ell^\square(A)$  when  $k \leq \ell$  and  $M_k^\square(A) + M_\ell^\square(A) \subseteq M_{k+\ell}^\square(A)$ .

As in the non-monotone case, we can understand the cones  $M^\square(A)$  as the conical hull of curve segments.

**Definition 2.1.** We define maps  $\gamma_A^\uparrow$  and  $\gamma_A^\downarrow$  from  $[0, 1]$  to  $\mathbb{R}^A$  where, for  $t \in [0, 1]$ ,  $\gamma_A^\uparrow(t)$  and  $\gamma_A^\downarrow(t)$  are the  $A$ -moment vectors of the step functions  $(1/(1-t))\mathbf{1}_{(t,1]}$  and  $(1/t^{\min(A)+1})\mathbf{1}_{[0,t]}$ , respectively. For every  $a \in A$ , the  $a$ th coordinate of these maps are given by

$$\left( \gamma_A^\uparrow(t) \right)_a = \frac{1}{1-t} \int_t^1 x^a dx = \frac{1}{a+1} \sum_{i=0}^a t^i$$

$$\text{and } \left( \gamma_A^\downarrow(t) \right)_a = \frac{1}{t^{\min(A)+1}} \int_0^t x^a dx = \frac{1}{a+1} t^{a-\min(A)}.$$

We observe that  $\gamma_A^\uparrow(0) = \gamma_A^\downarrow(1) = (1/(a+1))_{a \in A}$  corresponds to the moment vector of constant function  $\mathbf{1}_{[0,1]}$ . The other end points correspond to point masses. Specifically,  $\gamma_A^\uparrow(1) = v_A(1)$  is the moment vector of a point mass at  $t = 1$  and  $\gamma_A^\downarrow(0) = \frac{1}{\min(A)+1} v_B(0)$  for  $B = \{a - \min(A) : a \in A\}$  corresponds to a point mass at  $t = 0$ .

**Remark 2.2.** The conical hull over  $\{\gamma_A^\square(t) : t \in [0, 1]\}$  is closed because this curve is compact and does not contain the origin. Indeed, for  $\square = \uparrow$ , the  $a$ th coordinate of  $\gamma_A^\uparrow(t)$  is  $\geq (1/a+1)$  for all  $t$ . For  $\square = \downarrow$ , the  $\min(A)$ -th coordinate of  $\gamma_A^\downarrow(t)$  is identically  $1/(\min(A)+1)$ .

**Lemma 2.3.** For  $\square \in \{\uparrow, \downarrow\}$ , the cone  $M^\square(A)$  equals the conical hull of  $\{\gamma_A^\square(t) : t \in [0, 1]\}$ .

*Proof.* Since  $M^\square(A)$  is a convex cone containing the point  $\gamma_A^\square(t)$  for all  $t$ , it automatically contains the conical hull of this curve.

For the other direction, consider a monotone function  $f : [0, 1] \rightarrow \mathbb{R}$ . We can construct a sequence of step functions  $f_n$  converging uniformly to  $f$  on  $[0, 1]$ . For example, we may take

$f_n = \sum_{i=1}^n \frac{M}{n} \mathbf{1}_{T_i}$  where  $M \in \{f(0), f(1)\}$  is the maximal value of  $f$  on  $[0, 1]$  and  $\mathbf{1}_{T_i}$  is the indicator function of  $T_i = \{x \in [0, 1] : f(x) \geq iM/n\}$ . That is  $f_n(x) = \frac{M}{n} \cdot \lfloor \frac{n}{M} f(x) \rfloor$ . Note that  $|f_n - f| \leq M/n$  and so  $f_n$  converges uniformly to  $f$  on  $[0, 1]$ . It follows that for any  $a$ ,  $x^a f_n$  converges uniformly to  $x^a f$  and so the integral  $\int_0^1 x^a f_n(x) dx$  converges to  $\int_0^1 x^a f(x) dx$ .

Note that the set  $T_i$  defined above has the form  $(s_i, 1]$  or  $[s_i, 1]$  if  $f$  is increasing and  $[0, s_i]$  or  $[0, s_i)$  if  $f$  is decreasing for some  $s_i \in [0, 1]$ . The moment vector of  $f_n$  therefore is a conic combination of the points  $\gamma_A^\square(s_i)$  for the appropriate  $\square \in \{\uparrow, \downarrow\}$ . Taking  $n \rightarrow \infty$  shows that the moment vector of  $f$  belongs to the closure of the conical hull of  $\{\gamma_A^\square(t) : t \in [0, 1]\}$ .

Therefore the moment cone  $\{(\int_0^1 x^a f(x) dx)_{a \in A} : f \text{ nonnegative and increasing on } [0, 1]\}$  belongs to the closure of the conical hull of  $\{\gamma_A^\uparrow(t) : t \in [0, 1]\}$ . By definition,  $M^\uparrow(A)$  is the closure of this set and so also belongs to the closure of this conical hull. Similarly  $M^\downarrow(A)$  belongs to the closure of the conical hull of  $\{\gamma_A^\downarrow(t) : t \in [0, 1]\}$ . By Remark 2.2, both of these conical hulls are already closed.  $\square$

**Proposition 2.4.** *If  $k \geq \lfloor \frac{|A|}{2} \rfloor$ , then we have  $M_k^\uparrow(A) = M^\uparrow(A)$  and  $M_k^\downarrow(A) = M^\downarrow(A)$ .*

*Proof.* Our proof proceeds similarly to that of Theorem 1.5. Let  $\mathbf{m}$  be a point of the boundary of  $M^\square(A)$ . We want to express  $\mathbf{m}$  as the  $A$ -moment of an increasing step function of the fewest steps possible. Let  $\ell : \mathbb{R}^A \rightarrow \mathbb{R}$  define a supporting hyperplane of  $M^\square(A)$  at  $\mathbf{m}$ , so that  $\ell \geq 0$  on  $M^\square(A)$  and  $\ell(\mathbf{m}) = 0$ . By Lemma 2.3,  $M^\square(A)$  is the conical hull of a curve, hence  $\mathbf{m}$  will lie in the conical hull of points on this curve with  $\ell = 0$ . We use this to show that  $\mathbf{m}$  belongs to  $M_k^\square(A)$  for  $k \geq \lfloor \frac{|A|}{2} \rfloor$ .

( $\downarrow$ ) Let  $p(x) = \ell(\gamma_A^\downarrow(x)) = \sum_{a \in A} \frac{p_a}{a+1} x^{a-\min(A)}$ . The polynomial  $p$  is nonnegative on  $[0, 1]$ .

By Descartes' rule of signs,  $p$  has at most  $|A|-1$  positive roots, counting multiplicity, and if  $p_{\min(A)} = 0$ , then it has at most  $|A|-2$ . Let  $i$  denote the number of distinct roots of  $p$  in  $(0, 1)$  and  $b = 1$  if  $p(0) = 0$  and 0 otherwise. Since each interior root of  $p$  must have multiplicity  $\geq 2$ , this gives  $2i + b \leq |A|-1$ . Note that  $\gamma_A^\downarrow(t) \in M_1^\downarrow(A)$  for all  $t \in [0, 1)$  and belongs to  $M_0^\downarrow(A)$  for  $t = 1$ . Therefore  $\mathbf{m}$  belongs to  $M_k^\downarrow(A)$  for  $k = i + b \leq \frac{1}{2}(|A|-1 + b)$ . The bound follows from the integrality of  $i + b$  and  $b \in \{0, 1\}$ .

( $\uparrow$ ) Let  $p(x) = \ell(\gamma_A^\uparrow(x)) = \sum_{a \in A} \frac{p_a}{a+1} \sum_{i=0}^a x^i$ , which is a polynomial nonnegative on  $[0, 1]$ .

Again, by Descartes' rule of signs,  $p$  has at most  $|A|-1$  positive roots, counting multiplicity. If  $i$  is the number of distinct roots of  $p$  in  $(0, 1)$  and  $b = 0$  if  $p(1) = 0$  and 0 otherwise, this gives that  $2i + b \leq |A|-1$ . As before,  $\gamma_A^\uparrow(t) \in M_1^\uparrow(A)$  for all  $t \in (0, 1]$  and belongs to  $M_0^\uparrow(A)$  for  $t = 0$ . Therefore  $\mathbf{m}$  belongs to  $M_k^\uparrow(A)$  for  $k = i + b \leq \frac{1}{2}(|A|-1 + b) \leq \frac{1}{2}|A|$ .

Now consider  $\mathbf{m}$  in the interior of  $M^\square(A)$  and let  $\mathbf{c}$  be the moment vector of the constant function  $\mathbf{1}_{[0,1]}$ . Let  $\lambda^*$  be the maximum value of  $\lambda \in \mathbb{R}$  for which  $\mathbf{m} - \lambda \mathbf{c}$  belongs to  $M^\square(A)$ . Since  $\mathbf{m} \in M^\square(A)$ , we know that  $\lambda^* \geq 0$ , and for sufficiently large  $\lambda$ ,  $\mathbf{m} - \lambda \mathbf{c} \notin M^\square(A)$ . Thus  $\mathbf{m} - \lambda^* \mathbf{c}$  belongs to the boundary of  $M^\square(A)$ , which is equal to the boundary of  $M_k^\square(A)$  by the argument above. Hence,  $\mathbf{m}$  also belongs to  $M_k^\square(A)$ .  $\square$

**Proposition 2.5.** *For all  $k < \lfloor \frac{|A|}{2} \rfloor$ , the cone  $M_k^\square(A)$  is a proper subset of  $M^\square(A)$ .*

*Proof.* The cone  $M_k^\square(A) \subset \mathbb{R}^{|A|}$  is a conic combination of  $k$  points on the boundary curve  $\gamma_A^\square$ , each contributing two degrees of freedom, and the point corresponding to the image of the constant step function  $\gamma_A^\uparrow(0) = \gamma_A^\downarrow(1)$ , contributing a single degree of freedom. Therefore, the semialgebraic set  $M_k^\square(A)$  has dimension at most  $\min\{2k + 1, |A|\}$ . The cone  $M^\square(A)$  is full-dimensional in  $\mathbb{R}^{|A|}$ . Let  $n = \lfloor |A|/2 \rfloor$  so that  $|A|$  is  $2n$  or  $2n + 1$ . In either case, we observe that for  $k \leq n - 1$ , the dimension of  $M_k^\square(A)$  is less than or equal to  $2n - 1$ , hence it cannot fill up all of  $M^\square(A)$ .  $\square$

**Example 2.6.** For  $A = \{0, 2, 5, 9\}$ ,  $M_1(A)$  is a union of  $M_1^\uparrow(A)$  and  $M_1^\downarrow(A)$ , shown on the left in Figure 3. Since  $1 < 2 = \lfloor |A|/2 \rfloor$ , these sets are not full dimensional and so cannot fill up  $M^\uparrow(A)$  or  $M^\downarrow(A)$ . For  $k = 2 = \lfloor |A|/2 \rfloor$ ,  $M_2^\uparrow(A) = M^\uparrow(A)$  and  $M_2^\downarrow(A) = M^\downarrow(A)$ . These form parts of the full dimensional set  $M_2(A)$  shown in the middle of Figure 3.

### 3. CONNECTION WITH COALESCENCE MANIFOLD

The motivation for studying moments of step functions comes from the field of **population genetics**. A central problem in this area is:

**Question 3.1.** *Given a sample of  $n$  genomes from a present-day population, what inferences can be drawn regarding the history of that population?*

Our approach to the problem is to fix a function  $p(t)$  describing *effective population size* at time  $t$  before the present. We then compute, as a function of  $p$ , a vector of invariants  $\mathbf{c}$  associated to the genome sample. Understanding the relationship between  $p$  and  $\mathbf{c}$  will allow us to infer likely values of  $p$  based on measured data.

Following [1], we model the natural process of the production of a sample of  $n$  genomes as follows:

- The genealogical tree connecting  $n$  individuals will be formed by taking coalescence of each pair of lineages as a Poisson point process with rate parameter  $1/p(t)$ , where  $p(t)$  is the effective population size at time  $t$  before present. (Heuristically, looking at the previous generation and picking parents at random, there is a  $1/p(t)$  chance that two lineages will pick the same parent.)
- After the tree is specified, mutations are distributed on the tree as a Poisson point process with constant rate relative to branch length. The infinite-sites model is used, so that repeated mutation at a given site is disallowed, which is a good model for large genomes.

**Definition 3.2.** Fixing a population history, and defining the random process as above, we define random variables:

- The *sample frequency spectrum* (also known as the *site* or *allele frequency spectrum*), abbreviated SFS, is the vector of random variables  $(X_{n,b})_{b=1,\dots,n-1}$  where  $X_{n,b}$  denotes the number of mutations that are shared by exactly  $b$  out of the  $n$  individuals.
- The *coalescence vector* is the vector  $(T_{i,i})_{i=2,\dots,n-1}$  of the time at which a sample of size  $i$  has exactly  $i$  distinct lineages, i.e. the time until the first coalescence.

For a fixed population function  $p$ , taking expectations gives the population invariants  $\xi_{n,b} = \mathbb{E}[X_{n,b}]$  and  $c_i = \mathbb{E}[T_{i,i}]$ .

In practice, the SFS is more frequently discussed as a summary statistic, but the coalescence vector is simpler to use in computations. Fortunately, Polanski and Kimmel [6] proved that

they are related by a linear transformation  $A_n$ , a matrix entirely determined by sample size  $n$ . Therefore, we focus on the coalescence vectors  $(c_i)$ .

**Fact 3.3.** We make the reasonable assumption that  $p(t)$  is bounded below by 0 and bounded above by a fixed  $P$ . Applying integration by parts and change of variables to the expected value of an exponential distribution yields the following expression for  $c_i$  in terms of  $p(t)$ :

$$(5) \quad c_i(\mathbf{p}) = \int_0^\infty \tilde{p}(\tau) \exp \left[ -\binom{i}{2} \tau \right] d\tau,$$

where  $\tilde{p}(\tau) = p(R_p^{-1}(\tau))$  and  $R_p(t) = \int_0^t \frac{1}{p(x)} dx$ . Because  $0 < p(t) < P$ , the function  $R_p$  is strictly increasing and unbounded; thus, it is a bijection from  $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ , so the inverse is well-defined. We call  $\tilde{p}(\tau)$  the *transformed population history*.

The coalescence vector can thus be considered a function from the space of (bounded) population history functions to  $\mathbb{R}^{n-1}$ . Since the former space is infinite-dimensional and the latter is finite-dimensional, it is natural to restrict our attention to a finite-dimensional space of population history functions. A common choice for this, motivated by injectivity considerations in [1], is

$$\tilde{S}_k = \{\text{nonnegative step functions on } \mathbb{R}_{\geq 0} \text{ with at most } k \text{ breakpoints}\}.$$

**Definition 3.4.** Let  $n, k$  be integers with  $n \geq 2$  and  $k \geq 0$ . The *coalescence manifold*  $\mathcal{C}_{n,k}$  is the Euclidean closure of the set of vectors  $\tilde{\mathbf{c}}(\mathbf{p}) = \mathbf{c}(\mathbf{p}) / \|\mathbf{c}(\mathbf{p})\|_1$  for all  $\mathbf{p} \in \tilde{S}_k$ . Here,  $\mathbf{c}(\mathbf{p}) = (c_2(\mathbf{p}), \dots, c_n(\mathbf{p}))$  where  $c_i(\mathbf{p})$  is defined as in Equation 5.

Because the vectors are normalized to have sum one, the coalescence manifold lives in the simplex  $\Delta^{n-1}$ . Note that this definition deviates slightly from the definition in [7] by allowing  $k$  breakpoints instead of  $k$  epochs (i.e. constant intervals). This shifts the index down by one. We now connect back to the moment cones studied above.

**Theorem 3.5.** Let  $A = \{\binom{i}{2} - 1 : i = 2, \dots, n\}$ . The coalescence manifold  $\mathcal{C}_{n,k}$  equals the intersection of the cone  $M_k(A)$  with the affine hyperplane of points with coordinate sum equal to one:

$$\mathcal{C}_{n,k} = \left\{ \mathbf{m} \in M_k(A) : \sum_{a \in A} m_a = 1 \right\}.$$

Before we prove the theorem, we demonstrate two lemmas that will simplify the proof.

**Lemma 3.6.** Define  $\tilde{p}(\tau)$  as in Equation 5. Then  $p(t) \in \tilde{S}_k$  if and only if  $\tilde{p}(\tau) \in \tilde{S}_k$ .

*Proof.* Let  $0 = s_0 < \dots < s_{k-1} < s_k$  be the sequence of breakpoints of  $p(t)$ . The function  $R_p(t)$  is a monotone increasing function, so the conditions below are equivalent:

$$s_j < t \leq s_{j+1} \iff R_p(s_j) < R_p(t) \leq R_p(s_{j+1}).$$

Since  $p$  is constant on  $(s_j, s_{j+1}]$ , the transformed history  $\tilde{p}(\tau) = p(R_p^{-1}(\tau))$  is constant on  $(R_p(s_j), R_p(s_{j+1})]$ . This implies that there are still at most  $k$  breakpoints.

For the reverse direction, repeat the argument with  $R_p^{-1}$  in place of  $R_p$ .  $\square$

**Lemma 3.7.** Let  $q$  be a strictly positive step function in  $\tilde{S}_k$ . Then, there exists  $p$  in  $\tilde{S}_k$  such that  $q(\tau) = p(R_p^{-1}(\tau))$  where  $R_p(t) = \int_0^t \frac{1}{p(x)} dx$  as above.

*Proof.* Let  $Q(t) = \int_0^t q(x)dx$ . We claim the desired function is  $p(t) = q(Q^{-1}(t))$ . First, note that because  $q$  is strictly positive and takes only finitely many values, it is bounded away from zero. Therefore  $Q$  is strictly increasing and takes all values in  $[0, \infty)$ . Its inverse  $Q^{-1}$  therefore exists and is also increasing with range  $[0, \infty)$ . It follows that  $p$  takes the same values in the same order as  $q$ . In particular,  $p \in \tilde{S}_k$ .

To check that  $q(t) = p(R_p^{-1}(t))$ , we first show that  $R_p(Q(t)) = t$  for all  $t \geq 0$ . By definition,

$$R_p(Q(t)) = \int_0^{Q(t)} \frac{1}{p(x)} dx = \int_0^{Q(t)} \frac{1}{q(Q^{-1}(x))} dx = \int_0^t \frac{1}{q(w)} q(w) dw = t,$$

where the penultimate equation comes from substituting  $x = Q(w)$  and  $dx = q(w)dw$ . Since both  $Q$  and  $R_p$  are invertible, we see that  $t = Q^{-1}(R_p^{-1}(t))$  for all  $t$ . Applying  $q$  to both sides then gives the claim.  $\square$

*Proof of Theorem 3.5.* We show that the set of coalescence vectors coming from population histories in  $\tilde{S}_k$  is equal to the set of moments in  $M_k(A)$  summing to 1. The equality of the two closures is then automatic.

Assume  $p \in \tilde{S}_k$ . From Lemma 3.6,  $\tilde{p}$  is also in  $\tilde{S}_k$ . Starting with Equation 5, we substitute  $u = e^{-\tau}$  to obtain:

$$c_i(p) = \int_0^1 \tilde{p}^*(u) u^{\binom{i}{2}-1} du, \text{ where } \tilde{p}^*(u) = p(R_p^{-1}(-\ln(u))).$$

The function  $\tilde{p}^*$  is piecewise-constant on  $[0, 1]$  with at most  $k$  breakpoints, so is in  $S_k$ ; therefore, the quantity  $c_i$  is the  $(\binom{i}{2} - 1)$ -th moment of  $\tilde{p}^*$ . This implies that  $\mathbf{c}$  is in  $M_k(A)$  where  $A = \{\binom{i}{2} - 1 : i = 2, \dots, n\}$ . Normalizing  $\mathbf{c}$  is equivalent to scaling  $\tilde{p}^*$  so we may assume its sum is already equal to 1.

Conversely, up to closure, any moment vector in  $M_k(A)$  summing to 1 comes from some  $f \in S_k$ . Changing our domain to  $\mathbb{R}_{\geq 0}$  gives  $q(\tau) = f(e^{-\tau})$  in  $\tilde{S}_k$ . By Lemma 3.7, we can produce  $p \in \tilde{S}_k$  that gives transformed population history  $q$ .  $\square$

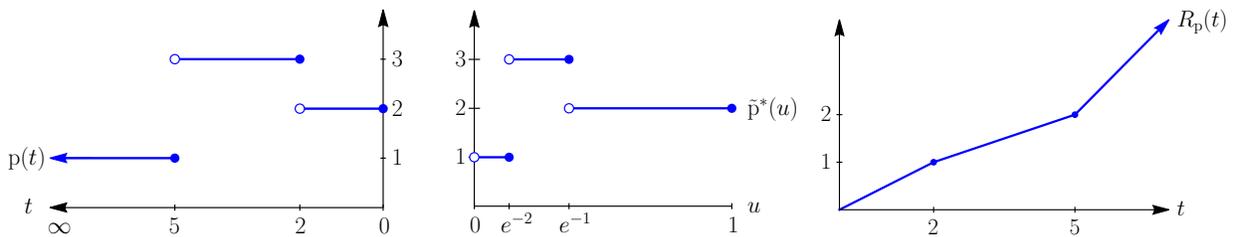


FIGURE 6. The functions  $p$ ,  $\tilde{p}^*$ , and  $R_p$  from Example 3.8.

**Example 3.8.** Consider the population function  $p(t) = p_1 \cdot \mathbf{1}_{[0, b_1)} + p_2 \cdot \mathbf{1}_{[b_1, b_2)} + p_3 \cdot \mathbf{1}_{[b_2, \infty)}$  where  $p_1, p_2, p_3, b_1, b_2 \in \mathbb{R}_{>0}$  with  $b_1 < b_2$ . The function  $R_p(t)$  is piecewise linear, given by

$$R_p(t) = \int_0^t \frac{1}{p(x)} dx = \frac{t}{p_1} \mathbf{1}_{[0, b_1)} + \left( \frac{t - b_1}{p_2} + \frac{b_1}{p_1} \right) \mathbf{1}_{[b_1, b_2)} + \left( \frac{t - b_2}{p_3} + \frac{b_2 - b_1}{p_2} + \frac{b_1}{p_1} \right) \mathbf{1}_{[b_2, \infty)}.$$

This function is unbounded and strictly increasing with  $R_p(0) = 0$ , so it has an inverse  $R_p^{-1}$  that is also increasing and unbounded on  $\mathbb{R}_{\geq 0}$ . The function  $\tilde{p}(\tau) = p(R_p^{-1}(\tau))$  is still

piecewise constant with two break points  $R_p(b_1) = b_1/p_1$  and  $R_p(b_2) = (b_2 - b_1)/p_2 + b_1/p_1$ , obtained by solving  $R_p^{-1}(\tau) = b_i$ . The  $i$ th entry of the coalescence vector is then

$$c_i = \int_0^\infty \tilde{p}(\tau) e^{\binom{i}{2}\tau} d\tau = \int_0^1 \tilde{p}^*(u) u^{\binom{i}{2}-1} du \quad \text{where} \quad \tilde{p}^*(u) = \tilde{p}(-\ln(u)).$$

The second equality comes from the change of coordinates  $u = e^{-\tau}$ . Note that  $\tilde{p}^*$  is the step function given by

$$\tilde{p}^* = p_3 \cdot \mathbf{1}_{(0, s_1]} + p_2 \cdot \mathbf{1}_{(s_1, s_2]} + p_1 \cdot \mathbf{1}_{(s_2, 1]} \quad \text{where} \quad s_1 = e^{-R_p(b_2)} \quad \text{and} \quad s_2 = e^{-R_p(b_1)}.$$

The graphs of  $p$  and  $\tilde{p}^*$  for the values  $(p_1, p_2, p_3) = (2, 3, 1)$  and  $(b_1, b_2) = (2, 5)$  are shown in Figure 6. In this case, the break points of  $\tilde{p}^*$  are  $e^{-R_p(b_2)} = e^{-2}$  and  $e^{-R_p(b_1)} = e^{-1}$ .

**Remark 3.9.** Note that because  $p(t)$  denote the population size at time  $t$  before the present, a population increasing over time corresponds to the function  $p(t)$  decreasing as a function of  $t$ , i.e.  $p_1 > p_2 > p_3$  in the example above. Note that  $p(t)$  is decreasing in  $t$  if and only if  $\tilde{p}(\tau)$  is decreasing in  $\tau$ . The parametrization  $u = e^{-\tau}$  reverses direction and so the function  $\tilde{p}^*(u)$  is then increasing as a function of  $u$ . In these coordinates,  $u = 0$  corresponds “infinitely long ago” ( $t = \infty$ ) and  $u = 1$  corresponds to the present ( $t = 0$ ). Therefore coalescence vectors of populations growing over time are moments of increasing step functions on  $[0, 1]$ .

Theorem 3.5 allows us to apply our results from  $M_k(A)$  to  $\mathcal{C}_{n,k}$ .

**Corollary 3.10.**  $\mathcal{C}_{n,n-2} = \mathcal{C}_{n,k}$  for all  $k \geq n - 2$  and  $\mathcal{C}_{n,n-3} \subsetneq \mathcal{C}_{n,n-2}$ .

*Proof.* For  $A = \left\{ \binom{i}{2} : i = 2, \dots, n \right\}$ ,  $|A|$  equals  $n - 1$ . By Theorem 1.5,  $M_k(A) = M(A)$  for all  $k \geq n - 2$ . In particular,  $M_{n-2}(A) = M_k(A)$  for all  $k \geq n - 2$ . Intersecting with the hyperplane  $\{\mathbf{m} : \sum_{a \in A} m_a = 1\}$  gives that  $\mathcal{C}_{n,n-2} = \mathcal{C}_{n,k}$  for all  $k \geq n - 2$ . By Corollary 1.12,  $M_k(A) \neq M(A)$  for  $k < |A| - 1 = n - 2$ . Hence  $M_{n-3}(A) \neq M(A)$ . Since  $M(A) = M_{n-2}(A)$ , intersecting with the hyperplane  $\{\mathbf{m} : \sum_{a \in A} m_a = 1\}$  gives that  $\mathcal{C}_{n,n-2} \neq \mathcal{C}_{n,n-3}$ .  $\square$

Affine transformations the sets  $\mathcal{C}_{5,1}$ ,  $\mathcal{C}_{5,2}$  and  $\mathcal{C}_{5,3}$  are show in Figure 3. As promised,  $\mathcal{C}_{5,3}$  is convex and  $\mathcal{C}_{5,k}$  is a strict subset for  $k < 3$ .

#### 4. CONNECTIONS WITH SEMIDEFINITE PROGRAMMING

In this section, we describe how to write the moment cone  $M(A)$  and coalescence manifold  $\mathcal{C}_{n,n-2}$  as projections of spectrahedra. This gives rise to natural algorithms for testing membership and finding nearest points in these sets based on semidefinite programming. Formally, a spectrahedron is a set of the form  $\{\mathbf{x} \in \mathbb{R}^n : A_0 + \sum_{i=1}^n x_i A_i \succeq 0\}$  where  $A_0, \dots, A_n$  are real symmetric matrices and  $X \succeq 0$  denotes that the matrix  $X$  is positive semidefinite. These are the feasible sets of semidefinite programs. See e.g. [2, Ch. 5 and 6]. Python code for computing the nearest point in  $\mathcal{C}_{n,n-2}$  to an arbitrary point in  $\mathbb{R}^{n-1}$  is available at:

<https://github.com/gescholt/DistanceToCoalescenceManifold>

**Theorem 4.1** (Theorems 10.1 and 10.2 [8]). *For any  $d \in \mathbb{Z}_+$ , the cone  $M(\{0, 1, \dots, d\})$  is a spectrahedron. If  $d = 2e$  is even, then*

$$M(\{0, 1, \dots, d\}) = \left\{ \mathbf{m} \in \mathbb{R}^{d+1} : (m_{i+j})_{0 \leq i, j \leq e} \succeq 0 \text{ and } (m_{i+j+1} - m_{i+j+2})_{0 \leq i, j \leq e-1} \succeq 0 \right\},$$

and if  $d = 2e + 1$  is odd, then

$$M(\{0, 1, \dots, d\}) = \left\{ \mathbf{m} \in \mathbb{R}^{d+1} : (m_{i+j+1})_{0 \leq i, j \leq e} \succeq 0 \text{ and } (m_{i+j} - m_{i+j+1})_{0 \leq i, j \leq e} \succeq 0 \right\}.$$

**Corollary 4.2.** *For any finite set of integers  $A \subset \mathbb{N}$ , the convex cones  $M(A)$ ,  $M^\uparrow(A)$  and  $M^\downarrow(A)$  are projections of the spectrahedron  $M(\{0, 1, \dots, \max(A)\})$ .*

*Proof.* Let  $d = \max(A)$ . Note that by definition,  $M(A)$  equals the closure of the projection of  $M(\{0, 1, \dots, d\})$  under the map  $(m_0, m_1, \dots, m_d) \mapsto (m_a)_{a \in A}$ . For  $0 \in A$ , this projection is closed and otherwise, we replace  $A$  with  $B = \{a - \min(A) : a \in A\}$  as in Lemma 1.2. By Theorem 4.1,  $M(\{0, 1, \dots, d\})$  is a spectrahedron.

More generally, consider any finite collection of polynomials  $p_1, \dots, p_n \in \mathbb{R}[x]_{\leq d}$ . We claim that the conical hull of the curve parameterized by  $\mathbf{p}(t) = (p_1(t), \dots, p_n(t))$  for  $t \in [0, 1]$  is the image of  $M(\{0, 1, \dots, d\})$  under a linear map. Specifically, consider the linear map  $\pi : \mathbb{R}^{d+1} \rightarrow \mathbb{R}^n$  taking  $(m_0, m_1, \dots, m_d)$  to  $(\sum_{j=0}^d p_{ij} m_j)_{i \in [n]}$  where  $p_i(x) = \sum_{j=0}^d p_{ij} x^j$ . For any  $t \in [0, 1]$ ,  $\mathbf{p}(t)$  equals  $\pi(v_d(t))$  where  $v_d(t) = (1, t, t^2, \dots, t^d)$ . Since  $M(\{0, 1, \dots, d\})$  is the conical hull of  $\{v_d(t) : t \in [0, 1]\}$ , the conical hull of  $\{\mathbf{p}(t) : t \in [0, 1]\}$  is the image of  $M(\{0, 1, \dots, d\})$  under  $\pi$ .

Note that the coordinates of both  $\gamma_A^\uparrow(t)$  and  $\gamma_A^\downarrow(t)$  are given by polynomials in  $t$  of degree  $\leq d$ . Then by Lemma 2.3 and the arguments above, both  $M^\uparrow(A)$  and  $M^\downarrow(A)$  can be written as the image of  $M(\{0, 1, \dots, d\})$  under a linear map.  $\square$

**Example 4.3.** For  $A = \{0, 2, 5, 9\}$ , we write  $M(A)$ ,  $M^\uparrow(A)$  and  $M^\downarrow(A)$  are projections of the spectrahedron  $M(\{0, 1, \dots, 9\})$ . By Theorem 4.1, this is given by the set of  $\mathbf{m} = (m_0, \dots, m_9)$  in  $\mathbb{R}^{10}$  for which the matrices

$$\begin{pmatrix} m_1 & m_2 & m_3 & m_4 & m_5 \\ m_2 & m_3 & m_4 & m_5 & m_6 \\ m_3 & m_4 & m_5 & m_6 & m_7 \\ m_4 & m_5 & m_6 & m_7 & m_8 \\ m_5 & m_6 & m_7 & m_8 & m_9 \end{pmatrix} \text{ and } \begin{pmatrix} m_0 - m_1 & m_1 - m_2 & m_2 - m_3 & m_3 - m_4 & m_4 - m_5 \\ m_1 - m_2 & m_2 - m_3 & m_3 - m_4 & m_4 - m_5 & m_5 - m_6 \\ m_2 - m_3 & m_3 - m_4 & m_4 - m_5 & m_5 - m_6 & m_6 - m_7 \\ m_3 - m_4 & m_4 - m_5 & m_5 - m_6 & m_6 - m_7 & m_7 - m_8 \\ m_4 - m_5 & m_5 - m_6 & m_6 - m_7 & m_7 - m_8 & m_8 - m_9 \end{pmatrix}$$

are positive semidefinite. We obtain  $M(A)$  as the image of this cone under the linear map  $\mathbf{m} \mapsto (m_0, m_2, m_5, m_9)$ . Similarly, the cones  $M^\uparrow(A)$  and  $M^\downarrow(A)$  are the images of  $M(\{0, 1, \dots, 9\})$  under the (respective) maps

$$\mathbf{m} \mapsto \left( m_0, \frac{m_0 + m_1 + m_2}{3}, \frac{1}{6} \sum_{i=0}^5 m_i, \frac{1}{10} \sum_{i=0}^9 m_i \right) \text{ and } \mathbf{m} \mapsto \left( m_0, \frac{m_2}{3}, \frac{m_5}{6}, \frac{m_9}{10} \right).$$

**Corollary 4.4.** *Testing membership any of the cones  $M(A)$ ,  $M^\uparrow(A)$  or  $M^\downarrow(A)$  is equivalent to testing the feasibility of a semidefinite program in  $\leq d + 1$  variables with two matrix constraints, each of size  $\leq d/2 + 1$ , where  $d = \max(A)$ .*

**Corollary 4.5.** *For  $k \geq n - 2$ , the coalescence manifold  $\mathcal{C}_{n,k}$  is the projection of a spectrahedron. Testing membership in  $\mathcal{C}_{n,k}$  is equivalent to testing the feasibility of a semidefinite program in  $\leq n^2/2$  variables with two matrix constraints, each of size  $\leq n^2/4$ .*

*Proof.* By Theorem 3.5 and Corollary 3.10, for all  $k \geq n - 2$ , coalescence manifold  $\mathcal{C}_{n,k}$  equals in the intersection of  $M(A)$  with the affine hyperplane given by  $\sum_{a \in A} m_a = 1$  where  $A = \{\binom{i}{2} - 1 : i = 2, \dots, n\}$ . By Corollary 4.2,  $M(A)$  is the projection of  $M(\{0, 1, \dots, d\})$  where  $d = \binom{n}{2} - 1$ . It follows that  $\mathcal{C}_{n,k}$  is the projection of the points in  $M(\{0, 1, \dots, d\})$  satisfying the affine linear equation  $\sum_{a \in A} m_a = 1$ . The intersection of a spectrahedron with an affine linear space is again a spectrahedron and so  $\mathcal{C}_{n,k}$  is the projection of a spectrahedron.

The spectrahedron  $M(\{0, 1, \dots, d\})$  is defined by two linear matrix inequalities of size  $\leq d/2 + 1 \leq n^2/4$ . There are at most  $d + 1 = \binom{n}{2} \leq n^2/2$  variables.  $\square$

Similarly, given a point  $\mathbf{p} \in \mathbb{R}^{n-1}$ , we can use a semidefinite program to find the nearest point in  $\mathcal{C}_{n,k}$  for sufficiently large  $k$ . This comes from the description of  $\mathcal{C}_{n,k}$  above and the fact that distance minimization can be phrased as a semidefinite program (see, e.g. [3]). Specifically, given  $\mathbf{x} \in \mathbb{R}^{n-1}$ , the matrix  $n \times n$  matrix  $\begin{pmatrix} \lambda & \mathbf{x} - \mathbf{p} \\ (\mathbf{x} - \mathbf{p})^T & \text{Id}_{n-1} \end{pmatrix}$  is positive semidefinite if and only if  $\|\mathbf{x} - \mathbf{p}\|_2^2 \leq \lambda$ , where  $\text{Id}_{n-1}$  denotes the  $(n-1) \times (n-1)$  identity matrix. Given a set  $S \subset \mathbb{R}^{n-1}$ , suppose that  $\lambda^*$  and  $\mathbf{x}^*$  obtain the minimum

$$\min_{\lambda \in \mathbb{R}, \mathbf{x} \in S} \lambda \quad \text{such that} \quad \begin{pmatrix} \lambda & \mathbf{x} - \mathbf{p} \\ (\mathbf{x} - \mathbf{p})^T & \text{Id}_{n-1} \end{pmatrix} \succeq 0.$$

Then  $\mathbf{x}^*$  is (one of) the nearest points in  $S$  to  $\mathbf{p}$  and the distance  $\|\mathbf{x}^* - \mathbf{p}\|_2$  is  $\sqrt{\lambda^*}$ . In particular, if the set  $S$  is the projection of a spectrahedron, then this minimization problem is a semidefinite program.

**Corollary 4.6.** *Given  $\mathbf{p} \in \mathbb{R}^{n-1}$ , the problem of finding the closest point to  $\mathbf{p}$  in  $\mathcal{C}_{n,k}$  for sufficiently large  $k$  is equivalent to solving a semidefinite program in  $\leq n^2/2$  variables with three matrices of size  $\leq n^2/4$ .*

**Example 4.7.** For  $n = 5$  and  $k \geq 3$ ,  $\mathcal{C}_{5,k}$  equals the set of points in  $M(\{0, 2, 5, 9\})$  with  $m_0 + m_2 + m_5 + m_9 = 1$ . Projecting from  $M(\{0, 1, \dots, 9\})$ , we see that

$$\mathcal{C}_{5,k} = \left\{ (m_0, m_2, m_5, m_9) \in \mathbb{R}^4 : m_0 + m_2 + m_5 + m_9 = 1 \right. \\ \left. \text{and } \exists (m_1, m_3, m_4, m_6, m_7, m_8) \in \mathbb{R}^6 \text{ such that } (m_j)_{j=0,\dots,9} \in M(\{0, 1, \dots, 9\}) \right\}$$

Let  $\mathcal{A}(\mathbf{m})$  and  $\mathcal{B}(\mathbf{m})$  denote the two  $5 \times 5$  matrices appearing in Example 4.3. Then  $M(\{0, 1, \dots, 9\})$  is the set of points  $\mathbf{m} \in \mathbb{R}^{10}$  for which  $\mathcal{A}(\mathbf{m}) \succeq 0$  and  $\mathcal{B}(\mathbf{m}) \succeq 0$ . Given a point  $\mathbf{p} = (a, b, c, d) \in \mathbb{R}^4$ , we can find the closest point in  $\mathcal{C}_{5,k}$  by solving the following semidefinite program with 10 parameters and three  $5 \times 5$  linear matrix constraints:

$$\min_{\lambda, m_0, \dots, m_9} \lambda \quad \text{such that} \quad m_0 + m_2 + m_5 + m_9 = 1, \quad \mathcal{A}(\mathbf{m}) \succeq 0, \quad \mathcal{B}(\mathbf{m}) \succeq 0, \\ \text{and} \quad \begin{pmatrix} \lambda & m_0 - a & m_2 - b & m_5 - c & m_9 - d \\ m_0 - a & 1 & 0 & 0 & 0 \\ m_2 - b & 0 & 1 & 0 & 0 \\ m_5 - c & 0 & 0 & 1 & 0 \\ m_9 - d & 0 & 0 & 0 & 1 \end{pmatrix} \succeq 0.$$

If  $(\lambda^*, \mathbf{m}^*)$  denotes the points achieving this minimum, then  $(m_0^*, m_2^*, m_5^*, m_9^*)$  is the closest point in  $\mathcal{C}_{5,k}$  to  $\mathbf{p}$  with distance  $\sqrt{\lambda^*}$ .

## 5. DISCUSSION AND OPEN QUESTIONS

One takeaway from Section 1 is that the points on the boundary of  $\mathcal{C}_{n,k}$  for  $k \geq n - 2$  correspond to moment vectors of point evaluations on  $[0, 1]$ . However these do not correspond to biologically meaningful population functions! Similarly, a point in the interior of  $\mathcal{C}_{n,k}$  can come from several different population functions, some of which are more biologically

plausible than others. One natural question from this standpoint is how to pick the right population history from the fiber of a coalescence vector.

**Question 5.1.** *Given a point  $\mathbf{m}$  in the interior of  $M_k(A)$ , how can we find the “best” step function  $f \in S_k$  with moment vector  $\mathbf{m}$ ?*

Here there is some natural flexibility in the notion of “best”. Ideally it should be biologically plausible and also easy to compute. For plausibility, it might be reasonable to try to bound or minimize the ratios  $y_{i+1}/y_i$  of consecutive population sizes. One step towards this would be to understand the structure of the fibers of the moment map  $\mu_A$ .

For  $k = 2$  and  $A = \{0, 2, 5\}$ , the  $(s_1, s_2)$ -coordinates of the fibers of some points in  $M_2(A)$  are shown below.

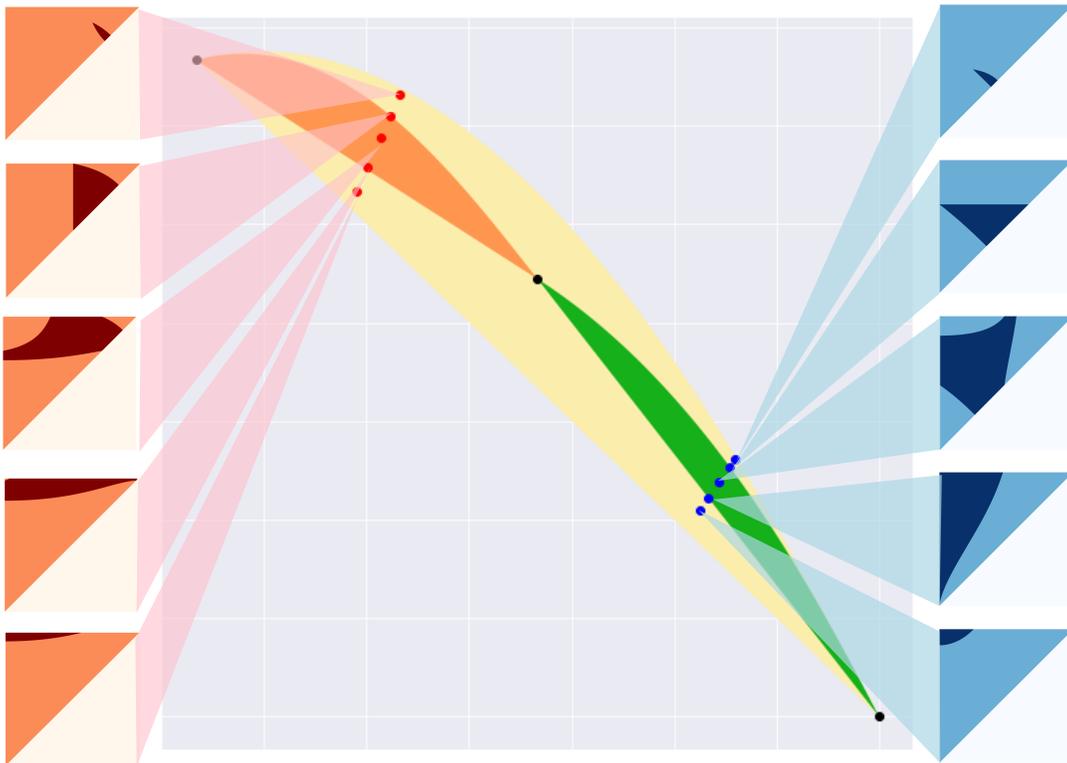


FIGURE 7. The central image depicts  $M_2(A)$  in yellow. The orange region is  $M^\uparrow(A)$  and the green region  $M^\downarrow(A)$ ; their union is  $M_1(A)$ . The triangle above each point depicts the fiber as a subset of the  $(s_1, s_2)$ -simplex.

To understand the fibers, it may also help to relate the combinatorial structure of the polytope  $P$  (which is a product of two  $k$ -dimensional simplices) to the semi-algebraic and combinatorial structure of  $M_k(A)$ . For example, the boundary of  $M_2(\{0, 2, 5, 9\})$ , seen in Figure 3, comes from some of the two-dimensional faces of the four-dimensional polytope  $P$ .

**Question 5.2.** *How does the facial structure of  $P$  relate to the algebraic boundary of  $M_k(A)$ ?*

Finally, Section 4 gives an algorithm for testing membership in  $M(A)$ , which coincides with  $M_k(A)$  for  $k \geq |A| - 1$ . It would be desirable to be able to test membership for smaller  $k$  as well.

**Question 5.3.** *Is there an effective method to test membership in  $M_k(A)$  for  $k < |A| - 1$ ?*

These sets are not convex and may have complicated semialgebraic structure (Figure 3). One possibility is the following connection to low rank matrix completion would be interesting to explore further.

Consider a step function  $f = y_1 \mathbf{1}_{[0, s_1]} + \sum_{i=2}^{k+1} y_i \mathbf{1}_{(s_{i-1}, s_i]}$  in  $S_k$ . In a slight abuse of notation, we define its derivative to be  $f' = \sum_{i=1}^k (y_{i+1} - y_i) \delta_{s_i}$ , which is a signed weighted sum of delta functions. For  $j \in A$ , let  $m'_j$  denote the  $j$ th moment of the signed measure given by  $f'$ :

$$m'_j = \int_0^1 x^j f'(x) dx = \sum_{i=1}^k (y_{i+1} - y_i) (s_i)^j.$$

One can check that for any  $j$ ,  $m'_j = f(1) - j m_{j-1}$ . In particular, we can write differences of consecutive moments of  $f'$  in terms of moments of  $f$ , namely  $m'_j - m'_{j+1} = (j+1)m_j - j m_{j-1}$ .

In the case of full moments  $A = \{0, 1, \dots, d\}$ , this lets us bound the value of  $k$  by the rank of the moment matrix corresponding to the moments of  $(x - x^2)f'(x)$ . Specifically, for  $\mathbf{m} \in \mathbb{R}^A$ , define the matrix

$$(6) \quad \mathcal{M}(\mathbf{m}) = \left( (j + \ell + 2)m_{j+\ell+1} - (j + \ell + 1)m_{j+\ell} \right)_{0 \leq j, \ell \leq \lfloor (d-1)/2 \rfloor}.$$

**Proposition 5.4.** *If for some  $f \in S_k$ ,  $m_j = \int_0^1 x^j f(x) dx$  for all  $j$ , then  $\text{rank}(\mathcal{M}(\mathbf{m})) \leq k$ .*

*Proof.* As noted above, we can rewrite the  $(j, \ell)$ th entry of  $\mathcal{M}(\mathbf{m})$  as

$$\mathcal{M}(\mathbf{m})_{j, \ell} = (j + \ell + 2)m_{j+\ell+1} - (j + \ell + 1)m_{j+\ell} = m'_{j+\ell+1} - m'_{j+\ell+2} = \sum_{i=1}^k (y_{i+1} - y_i) (s_i - s_i^2) s_i^{j+\ell}.$$

This shows that  $\mathcal{M}(\mathbf{m})$  is a sum of  $k$  rank-one matrices  $\sum_{i=1}^k (y_{i+1} - y_i) (s_i - s_i^2) v_e(s_i) v_e(s_i)^T$ , where  $e = \lfloor (d-1)/2 \rfloor$  and  $v_e(t) = (1, t, t^2, \dots, t^e)^T$ . Therefore  $\mathcal{M}(\mathbf{m})$  has rank  $\leq k$ .  $\square$

Note that if the values of  $y_i$  are increasing then this is a sum of positive semidefinite rank one matrices, in which case the rank of  $\mathcal{M}(\mathbf{m})$  will equal  $k$ , but if the values  $y_{i+1} - y_i$  have different signs, this might not be the case. Regardless, this suggests the following approach.

**Question 5.5.** *Given  $(m_a)_{a \in A} \in M(A)$ , when does the following low-rank matrix completion find the minimum  $k$  for which  $(m_a)_{a \in A}$  belongs to  $M_k(A)$ ?:*

$$\text{Minimize } \text{rank}(\mathcal{M}(\mathbf{m})) \quad \text{such that } \mathcal{A}(\mathbf{m}) \succeq 0, \mathcal{B}(\mathbf{m}) \succeq 0.$$

Here  $\mathcal{A}$  and  $\mathcal{B}$  are the matrices introduced in Theorem 4.1 and the minimization is taken over all  $\mathbf{m} \in \mathbb{R}^{\{0, 1, \dots, \max(A)\}}$  for which  $\mathbf{m}_a = m_a$  for all  $a \in A$ .

While it seems unlikely that this will always give the correct value, it would be interesting to know how far off this value might be from the true minimal value of  $k$ .

## REFERENCES

- [1] Anand Bhaskar and Yun S. Song. Descartes' rule of signs and the identifiability of population demographic models from genomic variation data. *Ann. Statist.*, 42(6):2469–2493, 2014.
- [2] Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas, editors. *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013.

- [3] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.
- [4] Jean Bernard Lasserre. *Moments, positive polynomials and their applications*, volume 1 of *Imperial College Press Optimization Series*. Imperial College Press, London, 2010.
- [5] Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008.
- [6] A. Polanski and M. Kimmel. New explicit expressions for relative frequencies of single-nucleotide polymorphisms with application to statistical inference on population growth. *Genetics*, 165(1):427–436, 2003.
- [7] Zvi S. Rosen, Anand Bhaskar, Sébastien Roch, and Y. Song. Geometry of the sample frequency spectrum and the perils of demographic inference. *Genetics*, 210:665 – 682, 2018.
- [8] Konrad Schmüdgen. *The moment problem*, volume 277 of *Graduate Texts in Mathematics*. Springer, Cham, 2017.

FLORIDA ATLANTIC UNIVERSITY, BOCA RATON, FL, USA  
*Email address:* rosenz@fau.edu

NORTH CAROLINA STATE UNIVERSITY, RALEIGH, NC, USA  
*Email address:* ghscholt@ncsu.edu

NORTH CAROLINA STATE UNIVERSITY, RALEIGH, NC, USA  
*Email address:* clvinzan@ncsu.edu