Compressed sampling inequalities by Tchakaloff's theorem *

M. Vianello [1]
Dept. of Mathematics, University of Padova (Italy)

October 1, 2015

Abstract

We show that a discrete version of Tchakaloff's theorem on the existence of positive algebraic cubature formulas, entails that the information required for multivariate polynomial approximation can be suitably compressed.

2000 AMS subject classification: 41A10, 65D32.

Keywords: Tchakaloff's theorem, multivariate polynomial inequalities, compressed sampling.

1 Introduction

The purpose of this note is to give a sound theoretical foundation to a compression technique, recently developed in [15], to reduce the sampling cardinality in multivariate polynomial least-squares and multivariate polynomial meshes (on the notion of polynomial mesh see, e.g., [4, 6, 9]).

We show that such a foundation can be obtained in a simple way by the well-known Tchakaloff's theorem, a deep result of cubature theory, that ensures existence of positive algebraic cubature formulas of low cardinality. Originally stated and proved by Tchakaloff [16] for compactly supported absolutely continuous measures with respect to the Lebesgue measure, it has then been generalized to arbitrary (even discrete) measures with finite moments; cf., e.g., [1, 14].

Here is a quite general version of the theorem, taken from [14, Thm. 1].

*Supported the ex-60% funds and by the biennial projects CPDA124755, CPDA143275 of the University of Padova, and by the INdAM GNCS.

[1] e-mail: *marcov@math.unipd.it*

Theorem 1 *Let μ be a positive measure with compact support Ω in $\mathbb{R}^d$ and let k be a fixed positive integer. Then there are $s \leq dim(\mathbb{P}_k^d)$ points $\{\xi_j\}$ in Ω and positive real numbers $\{\lambda_j\}$ such that*

$$\int_{\mathbb{R}^d} p(x)\, d\mu = \sum_{j=1}^{s} \lambda_j\, p(\xi_j) \tag{1}$$

for all $p \in \mathbb{P}_k^d$.

2 Compressed sampling

We use now a discrete instance of Tchakaloff's theorem, to prove polynomial inequalities for compressed sampling.

Proposition 1 *Let $X = \{x_1, \dots, x_M\}$ be a $\mathbb{P}_n^d$-determining finite subset of $\mathbb{R}^d$, with $M = card(X) > N = dim(\mathbb{P}_{2n}^d)$, and let $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_M) \in \mathbb{R}^M$ be positive weights. Then there exist a subset $\mathcal{A} = \{a_1, \dots, a_m\} \subset X$ with $m = card(\mathcal{A}) \leq N$, and positive weights $\boldsymbol{w} = (w_1, \dots, w_m) \in \mathbb{R}^m$, such that, denoting by $\mathcal{L}_n f \in \mathbb{P}_n^d$ the weighted least-squares polynomial approximation on $\mathcal{A}$ to a function f defined on X, i.e.*

$$\|f - \mathcal{L}_n f\|_{\ell^2_{\boldsymbol{w}}(\mathcal{A})} = \min_{p \in \mathbb{P}_n^d} \|f - p\|_{\ell^2_{\boldsymbol{w}}(\mathcal{A})} , \tag{2}$$

the following estimate holds

$$\|\mathcal{L}_n f\|_{\ell^2_{\boldsymbol{\sigma}}(X)} \leq \sqrt{\mu(X)}\, \|f\|_{\ell^\infty(A)} , \quad \mu(X) = \sum_{i=1}^{M} \sigma_i , \tag{3}$$

where m, $\mathcal{A}$ and $\boldsymbol{w}$ are independent of f.

Proof. Take $\Omega = X$ and $\mu = \sum_{i=1}^{M} \sigma_i \delta_{x_i}$ (the discrete measure with support X and point masses $\{\sigma_i\}$). By Theorem 1, applied for degree $k = 2n$, there exist a subset $\mathcal{A} = \{a_1, \dots, a_m\} \subset X$ with $m = \mathrm{card}(\mathcal{A}) \leq N$, and positive weights $\boldsymbol{w} = (w_1, \dots, w_m)$, such that

$$\|p\|^2_{\ell^2_{\boldsymbol{\sigma}}(X)} = \sum_{i=1}^{M} \sigma_i p^2(x_i) = \sum_{j=1}^{m} w_j p^2(a_j) = \|p\|^2_{\ell^2_{\boldsymbol{w}}(\mathcal{A})} \tag{4}$$

for all $p \in \mathbb{P}_n^d$. Observe that, since X is $\mathbb{P}_n^d$-determining (i.e., polynomials vanishing there vanish everywhere), by (4) also $\mathcal{A}$ is $\mathbb{P}_n^d$-determining.

Therefore, the weighted least-squares operator $\mathcal{L}_n : \ell^2_{\boldsymbol{w}}(\mathcal{A}) \to \mathbb{P}_n^d$ in (2) is well-defined. Taking in (4) the polynomial $p = \mathcal{L}_n f$ (i.e., the extension of $\mathcal{L}_n f$ to X), and using the Pythagorean theorem in $\ell^2_{\boldsymbol{w}}(\mathcal{A})$ ($\mathcal{L}_n f$ being an orthogonal projection), and the fact that $\sum w_j = \mu(X)$ by (1), we get the chain of inequalities

$$\|\mathcal{L}_n f\|^2_{\ell^2_{\boldsymbol{\sigma}}(X)} = \|\mathcal{L}_n f\|^2_{\ell^2_{\boldsymbol{w}}(\mathcal{A})} \le \|f\|^2_{\ell^2_{\boldsymbol{w}}(\mathcal{A})} \le \sum_{j=1}^{m} w_j \, \|f\|^2_{\ell^\infty(\mathcal{A})} = \mu(X)\, \|f\|^2_{\ell^\infty(A)} \;,$$

that is (3). $\square$

Estimate (3) has a number of significant consequences in sampling theory for multivariate polynomial approximation. For example, consider the following basic error estimate for the weighted least-squares polynomial approximation on $(X, \boldsymbol{\sigma})$ of a function $f \in C(K)$, say $L_n f$, where $K \supseteq X$ is a compact subset of $\mathbb{R}^d$,

$$\|f - L_n f\|_{\ell^2_{\boldsymbol{\sigma}}(X)} = \min_{p \in \mathbb{P}^d_n} \|f - p\|_{\ell^2_{\boldsymbol{\sigma}}(X)} \le \|f - p^*_n\|_{\ell^2_{\boldsymbol{\sigma}}(X)}$$

$$\le \sqrt{\mu(X)}\, \|f - p^*_n\|_{\ell^\infty(X)} \le \sqrt{\mu(X)}\, \|f - p^*_n\|_{L^\infty(K)} = \sqrt{\mu(X)}\, E_n(f;K) \;, \tag{5}$$

where p^*_n is the best uniform approximation polynomial for f on K and $E_n(f;K)$ the corresponding minimum error (for example K could be the closure of the convex hull of the points). If K is a Jackson compact, (5) estimates the approximation quality through the regularity of f, cf. [13].

On the other hand, consider the weighted least-squares polynomial approximation on $(\mathcal{A}, \boldsymbol{w})$ defined in (2). Since $\mathcal{L}_n$ is a projection operator, i.e., $\mathcal{L}_n p = p$ for all $p \in \mathbb{P}^d_n$, we can write the chain of inequalities

$$\|f - \mathcal{L}_n f\|_{\ell^2_{\boldsymbol{\sigma}}(X)} \le \|f - p^*_n\|_{\ell^2_{\boldsymbol{\sigma}}(X)} + \|\mathcal{L}_n(f - p^*_n)\|_{\ell^2_{\boldsymbol{\sigma}}(X)}$$

$$\le \sqrt{\mu(X)}\, \left(\|f - p^*_n\|_{\ell^\infty(X)} + \|f - p^*_n\|_{\ell^\infty(\mathcal{A})}\right) \le 2\sqrt{\mu(X)}\, E_n(f;K) \;, \tag{6}$$

which shows that the $\ell^2_{\boldsymbol{\sigma}}(X)$ reconstruction error of a continuous function f by the "compressed" least-squares operator $\mathcal{L}_n$ has a natural bound, with the same magnitude of that appearing in (5) for the original "complete" least-squares operator L_n. "Compressed" means that $\mathcal{L}_n f$ can be constructed by sampling f at less points than those required by $L_n f$, and this becomes particularly significant when $N \ll \text{card}(X)$ and/or the sampling process is difficult or costly.

Another interesting situation occurs in the standard unweighted instance, $\boldsymbol{\sigma} = (1, 1, \dots, 1)$, when $X = X_n$ is a (weakly) admissible polynomial mesh. We recall that a *Weakly Admissible Mesh* (WAM) is a sequence of finite subsets of a multidimensional $\mathbb{P}^d$-determining compact set, say $X_n \subset K \subset \mathbb{R}^d$ (or $\mathbb{C}^d$), which are *norming sets* for total-degree polynomial subspaces,

$$\|p\|_{L^\infty(K)} \le C_n \, \|p\|_{\ell^\infty(X_n)} \;, \;\; \forall p \in \mathbb{P}^d_n \;, \tag{7}$$

and both C_n and $\text{card}(X_n)$ increase at most polynomially with n (necessarily $\text{card}(X_n) \ge \dim(\mathbb{P}^d_n)$ since X_n is $\mathbb{P}^d_n$-determining). The positive number C_n

is called the "*constant*" of the WAM. When $C_n \equiv C$ does not depend on n we speak of an *Admissible Mesh* (AM), sometimes also called "polynomial mesh" in the literature [9, 11]. An AM is said *optimal* if $\mathrm{card}(X_n) = \mathcal{O}(n^d)$.

The notion of WAM has been introduced in the seminal paper [6] and since then it has emerged as a powerful tool in multivariate polynomial approximation. We quote among their properties that WAMs are preserved by affine transformations, can be constructed incrementally by finite union and product, and are stable under small perturbations on Markov compacts. Moreover, WAMs are well-suited for uniform least-squares approximation, and for polynomial interpolation at suitable extremal subsets, which are approximate versions of Fekete and Leja points. Concerning various theoretical and computational features of WAMs, and their role in multivariate polynomial approximation, we refer the reader, e.g., to [2, 4, 9, 11] and to the references therein.

Now, if X_n is a WAM, with

$$\mathrm{card}(X_n) > N = \dim(\mathbb{P}^d_{2n}) = \binom{2n+d}{d} = \frac{2^d}{d!}\,n^d + \mathcal{O}(n^{d-1}) \;, \qquad (8)$$

Proposition 1 entails that there exists a sequence of subsets $\mathcal{A}_n \subset X_n$, $\mathrm{card}(\mathcal{A}_n) \leq N$, such that

$$\begin{aligned} \|\mathcal{L}_n f\|_{L^\infty(K)} &\leq C_n\, \|\mathcal{L}_n f\|_{\ell^\infty(X_n)} \leq C_n\, \|\mathcal{L}_n f\|_{\ell^2(X_n)} \\ &\leq C_n\, \sqrt{\mathrm{card}(X_n)}\, \|f\|_{\ell^\infty(\mathcal{A}_n)} \;, \end{aligned} \qquad (9)$$

which shows immediately, for example, that we can estimate the uniform norm of the compressed least-squares operator as

$$\|\mathcal{L}_n\| = \sup_{f\neq 0} \frac{\|\mathcal{L}_n f\|_{L^\infty(K)}}{\|f\|_{L^\infty(K)}} \leq C_n\, \sqrt{\mathrm{card}(X_n)} \;. \qquad (10)$$

On the other hand, taking $f = p \in \mathbb{P}^d_n$ in (9), we get the following

Proposition 2 *If the compact set $K \subset \mathbb{R}^d$ has a (Weakly) Admissible Mesh X_n with $card(X_n) > dim(\mathbb{P}^d_{2n})$, then it has also a Weakly Admissible Mesh $\mathcal{A}_n \subset X_n$, with constant $C'_n = C_n\,\sqrt{card(X_n)}$ and $card(\mathcal{A}_n) \leq dim(\mathbb{P}^d_{2n})$.*

Existence proofs for WAMs and AMs typically resort to geometric properties of the compact sets (cf., e.g., [3, 4, 5, 6, 8, 9, 12]). Few general results are known. One concern Fekete interpolation points (points that maximize the absolute value of Vandermonde determinants), that always exist and form a WAM with $\mathrm{card}(X_n) = C_n = \dim(\mathbb{P}^d_n) = \mathcal{O}(n^d)$, but are very difficult to compute (and known analytically essentially only in dimension one, interval or complex circle).

Another general class has been found (non constructively) in [2, Prop. 23], where it is shown that any $\mathbb{P}^d$-determining compact set of $\mathbb{C}^d$ has a "near optimal" AM with $C_n \equiv C$ and $\mathrm{card}(X_n) = \mathcal{O}((n\log n)^d)$. Proposition 2 entails then immediately the following

Corollary 1 *Let $K \subset \mathbb{R}^d$ be compact and $\mathbb{P}^d$-determining. Then K has a Weakly Admissible Mesh $\mathcal{A}_n$ with card$(\mathcal{A}_n) \leq$ dim$(\mathbb{P}^d_{2n})$ and constant $C'_n = \mathcal{O}((n \log n)^{d/2})$.*

Proposition 2 could be relevant from the point of view of applications, since it guarantees the existence of low cardinality WAMs within high-cardinality WAMs. For example, given a simple polygon with ν sides in $\mathbb{R}^2$, it is simple to construct by minimal triangulation and finite union of triangles an optimal AM for it, with $C_n \equiv C = 2$ and $\text{card}(X_n) = 4(\nu-2)n^2 + \mathcal{O}(n)$; cf. [5]. Proposition 2 says that we could extract from X_n a WAM $\mathcal{A}_n$, with constant $C'_n = \mathcal{O}(n)$ and $\text{card}(\mathcal{A}_n) \leq 2n^2 + \mathcal{O}(n)$, thus reducing the cardinality essentially by a factor $2(\nu-2)$, which is substantial when ν is large (notice that there are other known WAMs for polygons, with constant $C_n = \mathcal{O}(\log^2 n)$ and $\text{card}(X_n) = (\nu-2)n^2 + \mathcal{O}(n)$, cf. [8]).

In the spirit of the considerations above, we can also state the following general result, which again is an immediate consequence of Proposition 2.

Corollary 2 *Let $K \subset \mathbb{R}^d$ be compact and $\mathbb{P}^d$-determining. If K has an optimal Admissible Mesh X_n (i.e., $C_n \equiv C$ and card$(X_n) = \mathcal{O}(n^d)$) with card$(X_n) >$ dim$(\mathbb{P}^d_{2n})$, then it has also a Weakly Admissible Mesh $\mathcal{A}_n \subset X_n$ with card$(\mathcal{A}_n) \leq$ dim$(\mathbb{P}^d_{2n})$ and constant $C'_n = \mathcal{O}(n^{d/2})$.*

Unfortunately, the known proofs of generalized Tchakaloff's theorem are not constructive, so it cannot be used directly for the compression of least-squares or of polynomial meshes. Nevertheless, it provides a foundation to the computational method studied in [15]. Indeed, the computation of the nodes $\{a_j\}$ and positive weights $\{w_j\}$ (cf. Proposition 1) is there formulated as the problem of finding a sparse nonnegative solution to the underdetermined Vandermonde-like linear system (consider column vectors)

$$V^t \boldsymbol{z} = \boldsymbol{\alpha} \ , \ \ V = (v_{ij}) = (\phi_j(x_i)) \ , \ \ \boldsymbol{\alpha} = V^t \boldsymbol{\sigma} \ , \tag{11}$$

$i = 1, \ldots, M$, $j = 1, \ldots, N$, where $span\{\phi_1, \ldots, \phi_N\} = \mathbb{P}^d_{2n}$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is the vector of $d\mu$-moments of the polynomial basis $\{\phi_j\}$,

$$\alpha_j = \int_X \phi_j(x)\, d\mu = \sum_{i=1}^{M} \sigma_i\, \phi_j(x_i) \ .$$

Now, by the discrete version of *Tchakaloff's theorem*, Proposition 1 ensures that *a nonnegative solution with at least $M - N$ zero components exists.* Therefore, we can solve the underdetermined system (11) via the NNLS (*Non Negative Least Squares*) quadratic programming problem

$$\|\boldsymbol{\alpha} - V^t \boldsymbol{z}\|_2 = \min \|\boldsymbol{\alpha} - V^t \boldsymbol{u}\|_2 \ , \ \ \boldsymbol{u} \in \mathbb{R}^M \ , \ \ \boldsymbol{u} \geq \boldsymbol{0} \ , \tag{12}$$

for which several numerical algorithms are known, for example the active set optimization algorithm by Lawson and Hanson [10] that computes a sparse solution. The nonzero components of the solution vector $\boldsymbol{z}$ correspond to the positive weights $\boldsymbol{w}$ and allow to extract a subset $\mathcal{A} \subset X$ with card$(\mathcal{A}) \leq N$.

We refer the reader to [15] for the implementation details, as well as for an analysis of the effect of the approximation error (a very small but nonzero residual $\varepsilon = \|\boldsymbol{\alpha} - V^t \boldsymbol{z}\|_2$) on the compressed sampling inequalities (6), (9) and (10). A compression code for WAMs, based on Proposition 2 and (12), has been inserted in the Matlab package [7].

References

[1] C. Bayer and J. Teichmann, The proof of Tchakaloff's theorem, Proc. Amer. Math. Soc. 134 (2006), 3035–3040.

[2] T. Bloom, L. Bos, J.-P. Calvi and N. Levenberg, Polynomial interpolation and approximation in $\mathbb{C}^d$, Ann. Polon. Math. 106 (2012), 53–81.

[3] L. Bos, J.-P. Calvi, N. Levenberg, A. Sommariva and M. Vianello, Geometric Weakly Admissible Meshes, Discrete Least Squares Approximations and Approximate Fekete Points, Math. Comp. 80 (2011), 1601–1621.

[4] L. Bos, S. De Marchi, A. Sommariva and M. Vianello, Weakly Admissible Meshes and Discrete Extremal Sets, Numer. Math. Theory Methods Appl. 4 (2011), 1–12.

[5] L. Bos and M. Vianello, Low cardinality admissible meshes on quadrangles, triangles and disks, Math. Inequal. Appl. 15 (2012), 229–235.

[6] J.P. Calvi and N. Levenberg, Uniform approximation by discrete least squares polynomials, J. Approx. Theory 152 (2008), 82–100.

[7] S. De Marchi, F. Piazzon, A. Sommariva and M. Vianello, `WAM`: Matlab package for polynomial fitting and interpolation on Weakly Admissible Meshes (software available online at: `http://www.math.unipd.it/~marcov/CAAsoft.html`).

[8] M. Gentile, A. Sommariva and M. Vianello, Polynomial interpolation and cubature over polygons, J. Comput. Appl. Math. 235 (2011), 5232–5239.

[9] A. Kroó, On optimal polynomial meshes, J. Approx. Theory 163 (2011), 1107–1124.

[10] C.L. mLawson and R.J. Hanson, Solving least squares problems, Classics in Applied Mathematics 15, SIAM, Philadelphia, 1995.

[11] F. Piazzon and M. Vianello, Small perturbations of polynomial meshes, Appl. Anal. 92 (2013), 1063–1073.

[12] F. Piazzon and M. Vianello, Constructing optimal polynomial meshes on planar starlike domains, Dolomites Res. Notes Approx. DRNA 7 (2014), 22–25.

[13] W. Pleśniak, Multivariate Jackson Inequality, J. Comput. Appl. Math. 233 (2009), 815–820.

[14] M. Putinar, A note on Tchakaloff's theorem, Proc. Amer. Math. Soc. 125 (1997), 2409-2414.

[15] A. Sommariva and M. Vianello, Compression of multivariate discrete measures and applications, Numer. Funct. Anal. Optim. 36 (2015), 1198–1223.

[16] V. Tchakaloff, Formules de cubature mécaniques à coefficients nonnégatifs, Bull. Sci. Math. 81 (1957), 123–134.