Caratheodory-Tchakaloff Least Squares

Federico Piazzon, Alvise Sommariva and Marco Vianello
Department of Mathematics, University of Padova (Italy)
Email: fpiazzon, alvise, marcov@math.unipd.it

***Abstract*—We discuss the Caratheodory-Tchakaloff (CATCH) subsampling method, implemented by Linear or Quadratic Programming, for the compression of multivariate discrete measures and polynomial Least Squares.**

I. Subsampling for discrete measures

Tchakaloff theorem, a cornerstone of quadrature theory, substantially asserts that for every compactly supported measure there exists a positive algebraic quadrature formula with cardinality not exceeding the dimension of the exactness polynomial space (restricted to the measure support). Originally proved by V. Tchakaloff in 1957 for absolutely continuous measures [21], it has then be extended to any measure with finite polynomial moments, cf. e.g. [6].

In this paper we focus on polynomial Least Squares (LS), that are ultimately orthogonal projections with respect to a discrete measure. In Section II and III we show that the original sampling set can be replaced by a smaller one, keeping practically invariant the LS approximation estimates.

We begin by stating a discrete version of Tchakaloff theorem, in its full generality, whose proof is based on Caratheodory theorem on conic linear combinations.

Theorem 1: Let μ be a multivariate discrete measure supported at a finite set $X = \{x_i\} \subset \mathbb{R}^d$, with correspondent positive weights (masses) $\boldsymbol{\lambda} = \{\lambda_i\}$, $i = 1, \dots, M$, and let $S = span(\phi_1, \dots, \phi_L)$ a finite dimensional space of d-variate functions defined on $K \supseteq X$, with $N = dim(S|_X) \leq L$.

Then, there exist a quadrature formula with nodes $T = \{t_j\} \subseteq X$ and positive weights $\boldsymbol{w} = \{w_j\}$, $1 \leq j \leq m \leq N$, such that

$$I_\mu(f) = \sum_{i=1}^{M} \lambda_i\, f(x_i) = \sum_{j=1}^{m} w_j\, f(t_j)\ ,\ \ \forall f \in S|_X\ . \quad (1)$$

Proof. Let $\{\psi_1, \dots, \psi_N\}$ be a basis of $S|_X$, and $V = (v_{ij}) = (\psi_j(x_i)) \in \mathbb{R}^{M \times N}$ the Vandermonde-like matrix of the basis computed at the support points. If $M > N$ (otherwise there is nothing to prove), existence of a positive quadrature formula for μ with cardinality not exceeding N can be immediately translated into existence of a nonnegative solution with at most N nonvanishing components to the underdetermined linear system

$$V^t \boldsymbol{u} = \boldsymbol{b} = V^t \boldsymbol{\lambda}\ ,\ \ \boldsymbol{u} \geq \boldsymbol{0}\ , \quad (2)$$

where $V^t \boldsymbol{\lambda} = \{I_\mu(\psi_j)\}$ is the vector of μ-moments of the basis $\{\psi_j\}$. Existence then holds by the well-known Caratheodory theorem [4] applied to the columns of V^t, which asserts that a conic (i.e., with positive coefficients) combination of any number of vectors in $\mathbb{R}^N$ can be rewritten as a conic combination of at most N (linearly independent) of them [5, §3.4.4]. □

Since the discrete version of Tchakaloff theorem is a direct consequence of Caratheodory theorem, we may term such an approach *Caratheodory-Tchakaloff subsampling*, and the corresponding nodes (with associated weights) a set of Caratheodory-Tchakaloff (CATCH) points.

The idea of reduction/compression of a finite measure by Tchakaloff or directly Caratheodory theorem recently arose in different contexts, for example in a probabilistic setting [12], as well as in univariate [9] and multivariate numerical quadrature [1], [17], [20], with applications to multivariate polynomial inequalities and Least Squares approximation [20], [23]. In many situations CATCH subsampling can produce a high Compression Ratio, namely when $N \ll M$ like for example in polynomial Least Squares approximation [20] or in QMC (Quasi-Monte Carlo) integration [1] or in particle methods [12],

$$C_{ratio} = \frac{M}{m} \geq \frac{M}{N} \gg 1\ , \quad (3)$$

so that the efficient computation of CATCH points and weights becomes a relevant task.

Now, unlike the proof of the general Tchakaloff theorem, that of the discrete version can be made constructive, since Caratheodory theorem itself has a constructive proof (cf., e.g., [5, §3.4.4]). On the other hand, such a proof does not give directly an efficient implementation. Nevertheless, there are at least two reasonably efficient approaches to solve the problem.

The first, adopted for example in [9] (univariate) and [20] (multivariate) in the framework of polynomial spaces, rests on *Quadratic Programming*, namely on the classical *Lawson-Hanson active set method* for *NonNegative Least Squares* (NLLS). Indeed, we may think to solve the quadratic minimum problem

$$\text{NNLS}:\quad \begin{cases} \min \|V^t \boldsymbol{u} - \boldsymbol{b}\|_2 \\ \boldsymbol{u} \geq \boldsymbol{0} \end{cases} \quad (4)$$

which exists by Theorem 1 and can be computed by standard NNLS solvers based on the Lawson-Hanson method [11], which seeks a sparse solution. Then, the nonvanishing components of such a solution give the weights $\boldsymbol{w} = \{w_j\}$ as well as the indexes of the nodes $T = \{t_j\}$ within X. A variant of the Lawson-Hanson method is implemented in the Matlab native function `lsqnonneg`, with a recent optimization in [18].

The second approach is based instead on *Linear Programming* via the classical *simplex method*. Namely, we may think to solve the linear minimum problem

$$\text{LP}: \quad \begin{cases} \min \boldsymbol{c}^t \boldsymbol{u} \\ V^t \boldsymbol{u} = \boldsymbol{b} \,, \quad \boldsymbol{u} \geq \boldsymbol{0} \end{cases} \tag{5}$$

where the constraints identify a polytope (the feasible region) in $\mathbb{R}^M$ and the vector $\boldsymbol{c}$ is chosen to be linearly independent from the rows of V^t (i.e., it is not the restriction to X of a function in S), so that the objective functional is not constant on the polytope. To this aim, if $X \subset K$ is determining on a supspace $T \supset S$ on K, i.e. a function in T vanishing on X vanishes everywhere on K, then it is sufficient to take $\boldsymbol{c} = \{g(x_i)\}$, $1 \leq i \leq M$, where the function $g|_K$ belongs to $T|_K \setminus S|_K$. For example, working with polynomials it is sufficient to take a polynomial of higher degree on K with respect to those in $S|_K$.

Observe that in our setting the feasible region is nonempty, since $\boldsymbol{b} = V^t \boldsymbol{\lambda}$, and we are interested in any basic feasible solution, i.e., in any vertex of the polytope, that has at least $M - N$ vanishing components. As it is well-known, the solution of the Linear Programming problem is a vertex of the polytope that can be computed by the simplex method (cf., e.g., [5]). Again, the nonvanishing components of such a vertex give the weights $\boldsymbol{w} = \{w_j\}$ as well as the indexes of the nodes $T = \{t_j\}$ within X. This approach was adopted for example in [17] as a basic step to compute, when it exists, a multivariate algebraic Gaussian quadrature formula.

Even though both, the active set method for (4) and the simplex method for (5), have theoretically an exponential complexity (worst case analysis), as it is well-known their practical behavior is quite satisfactory, since the average complexity turns out to be polynomial in the dimension of the problems (observe that in the present setting we deal with dense matrices); cf., e.g., [8, Ch. 9]. It is worth quoting here the extensive theoretical and computational results recently presented in the Ph.D. dissertation [22], where Caratheodory reduction of a discrete measure is implemented by Linear Programming, claiming an experimental average cost of $\mathcal{O}(N^{3.7})$.

A different combinatorial algorithm (Recursive Halving Forest), based on the SVD, is also there proposed to compute a basic feasible solution and compared with the best Linear Programming solvers, claiming an experimental average cost of $\mathcal{O}(N^{2.6})$. The methods are essentially applied to the reduction of Cartesian tensor cubature measures.

In our implementation of CATCH subsampling [15], we have chosen to work with the Octave native Linear Programming solver `glpk` and the Matlab native Quadratic Programming solver `lsqnonneg`, that are suitable for moderate size problems, like those typically arising with polynomial spaces ($S = S_\nu = \mathbb{P}_\nu^d$) in dimension $d = 2, 3$ and small/moderate degree of exactness ν. On large size problems, like those typically arising in higher dimension and/or high degree of exactness, the solvers discussed in [22] could become necessary.

Now, since we may expect that the underdetermined system (2) is not satisfied exactly by the computed solution, due to finite precision arithmetic and by the effect of an error tolerance in the iterative algorithms, namely that there is a nonzero moment residual

$$\|V^t \boldsymbol{u} - \boldsymbol{b}\|_2 = \varepsilon > 0 \,, \tag{6}$$

it is then worth studying the effect of such a residual on the accuracy of the quadrature formula. We can state and prove an estimate still in the general discrete setting of Theorem 1.

Proposition 1: [14] *Let the assumptions of Theorem 1 be satisfied, let $\boldsymbol{u}$ be a nonnegative vector such that (6) holds, where V is the Vandermonde-like matrix at X corresponding to a μ-orthonormal basis $\{\psi_k\}$ of $S|_X$, and let $(T, \boldsymbol{w})$ be the quadrature formula corresponding to the nonvanishing components of $\boldsymbol{u}$. Moreover, let $1 \in S$ (i.e., S contains the constant functions).*

Then, for every function f defined on X, the following error estimate holds

$$\left| I_\mu(f) - \sum_{j=1}^{m} w_j \, f(t_j) \right| \leq C_\varepsilon E_S(f; X) + \varepsilon \|f\|_{\ell^2_{\boldsymbol{\lambda}}(X)} \,, \tag{7}$$

where $C_\varepsilon = 2 \left(\mu(X) + \varepsilon \sqrt{\mu(X)} \right)$ *and*

$$E_S(f; X) = \min_{\phi \in S} \|f - \phi\|_{\ell^\infty(X)} \,. \tag{8}$$

It is worth observing that the assumption $1 \in S$ is quite natural, being satisfied for example in the usual polynomial and trigonometric spaces. From this point of view, we can also stress that sparsity cannot be ensured by the standard Compressive Sensing approach to underdetermined systems, such as the Basis Pursuit algorithm that minimizes $\|u\|_1$ (cf., e.g., [7]), since if $1 \in S$ then $\|u\|_1 = \mu(X)$ is constant.

Moreover, we notice that if $K \supset X$ is a compact set, then

$$E_S(f; X) \leq E_S(f; K) \,, \quad \forall f \in C(K) \,. \tag{9}$$

If S is a polynomial space (as in the sequel) and K is a "Jackson compact", $E_S(f; K)$ can be estimated by the regularity of f via multivariate Jackson-like theorems; cf. [16].

II. Caratheodory-Tchakaloff Least Squares

The case where $(X, \boldsymbol{\lambda})$ is itself a quadrature/cubature formula for some measure on $K \supset X$, that is the compression (or reduction) of such formulas, has been till now the main application of Caratheodory-Tchakaloff subsampling, in the classical framework of algebraic formulas as well as in the probabilistic/QMC framework; cf. [9], [17], [20] and [1], [12], [22]. In this survey, we concentrate on another relevant application, that is the *compression of multivariate polynomial Least Squares*.

Let us consider the total-degree polynomial framework, that is $S = S_\nu = \mathbb{P}_\nu^d(K)$, the space of d-variate real polynomials with total-degree not exceeding ν, restricted to $K \subset \mathbb{R}^d$, a compact set or a compact (subset of a) manifold. Let us define for notational convenience $E_n(f) = E_{\mathbb{P}_n^d(K)}(f;K) = \min_{p\in\mathbb{P}_n^d(K)} \|f-p\|_{L^\infty(K)}$, for $f \in C(K)$.

Discrete LS approximation by total-degree polynomials of degree at most n on $X \subset K$ is ultimately an orthogonal projection of a function f on $\mathbb{P}_n^d(X)$, with respect to the scalar product of $\ell^2(X)$, namely

$$\|f - \mathcal{L}_n f\|_{\ell^2(X)} = \min_{p\in\mathbb{P}_n^d(X)} \|f-p\|_{\ell^2(X)} \;. \qquad (10)$$

Recall that $\|g\|^2_{\ell^2(X)} = \sum_{i=1}^M g^2(x_i) = I_\mu(g^2)$ for every function g defined on X, where μ is the discrete measure supported at X with unit masses $\boldsymbol{\lambda} = (1,\dots,1)$.

Taking $p^* \in \mathbb{P}_n^d(X)$ such that $\|f - p^*\|_{\ell^\infty(X)}$ is minimum (the polynomial of best uniform approximation of f in $\mathbb{P}_n^d(X)$), we get immediately the classical LS error estimate

$$\|f - \mathcal{L}_n f\|_{\ell^2(X)} \le \|f - p^*\|_{\ell^2(X)} \le \sqrt{M}\, E_n(f) \;, \qquad (11)$$

where $M = \mu(X) = \mathrm{card}(X)$. In terms of the Root Mean Square Error (RMSE), an indicator widely used in the applications, we have

$$\mathrm{RMSE}_X(\mathcal{L}_n f) = \frac{1}{\sqrt{M}} \|f - \mathcal{L}_n f\|_{\ell^2(X)} \le E_n(f) \;. \qquad (12)$$

Now, if $M > N_{2n} = \dim(\mathbb{P}_{2n}^d(X))$ (we stress that here polynomials of degree $2n$ are involved), by Theorem 1 there exist $m \le N_{2n}$ Caratheodory-Tchakaloff (CATCH) points $T_{2n} = \{t_j\}$ and weights $\boldsymbol{w} = \{w_j\}$, $1 \le j \le m$, such that the following basic ℓ^2 identity holds for every $p \in \mathbb{P}_n^d(X)$

$$\|p\|^2_{\ell^2(X)} = \sum_{i=1}^M p^2(x_i) = \sum_{j=1}^m w_j\, p^2(t_j) = \|p\|^2_{\ell^2_{\boldsymbol{w}}(T_{2n})} \;. \qquad (13)$$

Notice that the CATCH points $T_{2n} \subset X$ are $\mathbb{P}_n^d(X)$-determining, i.e., a polynomial of degree at most n vanishing there vanishes everywhere on X, or in other terms $\dim(\mathbb{P}_n^d(T_{2n})) = \dim(\mathbb{P}_n^d(X))$, or equivalently any Vandermonde-like matrix with a basis of $\mathbb{P}_n^d(X)$ at T_{2n} has full rank. This also entails that, if X is $\mathbb{P}_n^d(K)$-determining, then such is T_{2n}.

Consider the $\ell^2_{\boldsymbol{w}}(T_{2n})$ LS polynomial $\mathcal{L}_n^c f$, namely

$$\|f - \mathcal{L}_n^c f\|_{\ell^2_{\boldsymbol{w}}(T_{2n})} = \min_{p\in\mathbb{P}_n^d(X)} \|f-p\|_{\ell^2_{\boldsymbol{w}}(T_{2n})} \;. \qquad (14)$$

Notice that $\mathcal{L}_n^c$ is a *weighted* least squares operator; reasoning as in (12) and observing that $\sum_{j=1}^m w_j = M$ since $1 \in \mathbb{P}_n^d$, we get immediately

$$\|f - \mathcal{L}_n^c f\|_{\ell^2_{\boldsymbol{w}}(T_{2n})} \le \sqrt{M}\, E_n(f) \;. \qquad (15)$$

On the other hand, we can also write the following estimates

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

$$\|\mathcal{L}_n^c(p^*-f)\|_{\ell^2(X)} = \|\mathcal{L}_n^c(p^*-f)\|_{\ell^2_{\boldsymbol{w}}(T_{2n})} \le \|f-p^*\|_{\ell^2_{\boldsymbol{w}}(T_{2n})} \;,$$

where we have used the basic ℓ^2 identity (13), the fact that $\mathcal{L}_n^c p^* = p^*$ and that $\mathcal{L}_n^c f$ is an orthogonal projection. By the estimates above we get eventually

$$\mathrm{RMSE}_X(\mathcal{L}_n^c f) \le 2E_n(f) \;, \qquad (16)$$

which shows the most relevant feature of the "compressed" least squares operator $\mathcal{L}_n^c$ at the CATCH points (CATCHLS), namely that *the LS and compressed CATCHLS RMSE estimates (12) and (16) have substantially the same size*.

This fact, in particular the appearance of the factor 2 in the estimate for the compressed operator, is reminiscent of *hyperinterpolation theory* [19]. Indeed, what we are constructing here is a sort of hyperinterpolation in a fully discrete setting. Roughly summarizing, hyperinterpolation ultimately approximates a (weighted) L^2 projection on $\mathbb{P}_n^d$ by a discrete weighted ℓ^2 projection, via a quadrature formula of exactness degree $2n$. Similarly, here we are approximating a ℓ^2 projection on $\mathbb{P}_n^d$ by a weighted ℓ^2 projection with a smaller support, again via a quadrature formula of exactness degree $2n$.

The estimates above are valid by the theoretical exactness of the quadrature formula. In order to take into account a nonzero moment residual as in (6), we state and prove the following

Proposition 2: [14] *Let μ be the discrete measure supported at X with unit masses $\boldsymbol{\lambda} = (1,\dots,1)$, let $\boldsymbol{u}$ be a nonnegative vector such that (6) holds, where V is the orthogonal Vandermonde-like matrix at X corresponding to a μ-orthonormal basis $\{\psi_k\}$ of $\mathbb{P}_{2n}^d(X)$, and let $(T_{2n}, \boldsymbol{w})$ be the quadrature formula corresponding to the nonvanishing components of $\boldsymbol{u}$. Then the following polynomial inequality hold for every $p \in \mathbb{P}_n^d(X)$*

$$\|p\|_{\ell^2(X)} \le \sqrt{M}\,\mathcal{A}_M(\varepsilon)\,\|p\|_{\ell^\infty(T_{2n})} \;,$$

$$\mathcal{A}_M(\varepsilon) = \left(1 - \varepsilon\sqrt{M}\right)^{-1/2} \left(1 + \varepsilon/\sqrt{M}\right)^{1/2} \;, \qquad (17)$$

provided that $\varepsilon\sqrt{M} < 1$.

Corollary 1: [14] *Let the assumptions of Proposition 2 be satisfied. Then the following error estimate holds for every $f \in C(K)$*

$$\|f - \mathcal{L}_n^c f\|_{\ell^2(X)} \le (1 + \mathcal{A}_M(\varepsilon))\,\sqrt{M}\,E_n(f) \;. \qquad (18)$$

Remark 1: Observe that $\mathcal{A}_M(\varepsilon) \to 1$ as $\varepsilon \to 0$, and quantitatively, $\mathcal{A}_M(\varepsilon) \approx 1$ for $\varepsilon\sqrt{M} \ll 1$. Then we can write the approximate estimate

$$\mathrm{RMSE}_X(\mathcal{L}_n^c f) \lesssim (2 + \varepsilon\sqrt{M}/2)\,E_n(f) \;,\;\; \varepsilon\sqrt{M} \ll 1 \;, \qquad (19)$$

i.e., we substantially recover (16), as well as the size of (12), with a mild requirement on the moment residual error (6).

An example of reconstruction of two bivariate functions with different regularity by LS and CATCHLS on a nonstandard domain (union of four disks) is displayed in Table 1 and Figure 1, where X is a low-discrepancy point set, namely the about 5600 Halton points of the domain taken from 10000 Halton points of the minimal surrounding rectangle. Polynomial Least Squares on low-discrepancy point sets have been recently studied for example in [13], in the more general framework of Uncertainty Quantification.

We have implemented CATCH subsampling by NonNegative Least Squares (via the `lsqnonneg` Matlab native function) and by Linear Programming (via the `glpk` Octave native

TABLE I: Cardinality m, Compression Ratio, moment residual and RMSE$_X$ by LS and CATCHLS for the Gaussian $f_1(\rho) = \exp(-\rho^2)$ and the power function $f_2(\rho) = (\rho/2)^5$, $\rho = \sqrt{x^2 + y^2}$, where X is the Halton point set of Figure 1.

deg n	5	10	15	20
N_{2n}	66	231	496	861
NNLS: m	66	231	493	838
LP: m	66	231	493	837
$C_{ratio} = M/m$	85	24	11	6.7
NNLS: residual ε	8.3e-14	3.4e-13	1.1e-12	2.7e-12
LP: residual ε	3.6e-14	3.9e-14	7.7e-14	7.3e-14
cpu ratio: NNLS/LP	0.34	0.23	0.64	0.97
f_1: LS	7.4e-03	4.5e-05	2.0e-07	1.7e-10
NNLS-CATCHLS	7.8e-03	4.6e-05	2.0e-07	1.7e-10
LP-CATCHLS	9.6e-03	5.2e-05	2.0e-07	1.7e-10
f_2: LS	6.7e-03	6.0e-05	6.7e-06	1.3e-06
NNLS-CATCHLS	7.0e-03	6.1e-05	6.8e-06	1.3e-06
LP-CATCHLS	8.3e-03	7.2e-05	6.7e-06	1.3e-06

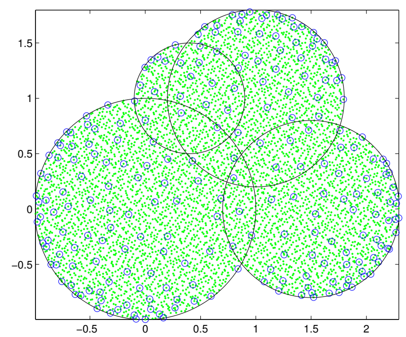

Fig. 1: Extraction of 231 points for CATCHLS ($n = 10$) from $M \approx 5600$ Halton points on the union of 4 disks by NNLS: $C_{ratio} = M/m \approx 24$.

function). In the Linear Programming approach, one has to choose a vector $\boldsymbol{c}$ in the target functional. Following [17], we have taken $\boldsymbol{c} = \left\{x_i^{2n+1} + y_i^{2n+1}\right\}$, where $X = \{(x_i, y_i)\}$, $1 \le i \le M$, i.e., the vector $\boldsymbol{c}$ corresponds to the polynomial $x^{2n+1} + y^{2n+1}$ evaluated at X. There are two reasons for this choice. The first is that (only) in the univariate case, as proved in [17], it leads to $2n + 1$ Gaussian quadrature nodes. The second is that the polynomial $x^{2n+1} + y^{2n+1}$ is not in the polynomial space of exactness, and thus we avoid that $\boldsymbol{c}^t\boldsymbol{u}$ be constant on the polytope defined by the constraints (recall, for example, that for $\boldsymbol{c}^t = (1, \dots, 1)$ we have $\boldsymbol{c}^t\boldsymbol{u} = \sum u_i = M$).

Observe in Table 1 that the CATCH points, computed by NNLS and LP, give a compressed LS operator with practically the same RMSEs as we had sampled at the original points, with remarkable Compression Ratios. The moment residuals appear more stable with LP, but are in any case extremely small with both solvers. In this example, NNLS turn out to be more efficient than LP.

We stress that the compression procedure is function independent, thus we can preselect the re-weighted CATCH sampling sites on a given region, and then apply the compressed CATCHLS formula to different functions. This approach to polynomial Least Squares could be very useful in applications where the sampling process is difficult or costly, for example to place a small/moderate number of accurate sensors on some region of the earth surface, for the measurement and reconstruction of a scalar or vector field.

III. From the discrete to the continuum

In what follows we study situations where the sampling sets are discrete models of "continuous" compact sets, in the framework of polynomial approximation. In particular, we have in mind the case where K is the *closure of a bounded open subset* of $\mathbb{R}^d$ (or of a bounded open subset of a lower-dimensional manifold in the induced topology, such as a subarc of the circle in $\mathbb{R}^2$ or a subregion of the sphere in $\mathbb{R}^3$). The so-called "Jackson compacts", that are compact sets where a Jackson-like inequality holds, are of special interest, since there the best uniform approximation error $E_n(f)$ can be estimated by the regularity of f; cf. [16].

Such a connection with the continuum has already been exploited in the previous sections, namely on the right-hand side of the LS error estimates, e.g. in (12) and (18). Now, to get a connection also on the left-hand side, we should give some structure to the discrete sampling set X. We shall work within the theory of *polynomial meshes*, introduced in [3] and later developed by various authors; cf., e.g., [2], [3], [10] and the references therein.

We recall that a *weakly admissible polynomial mesh* of a compact set K (or of a compact subset of a manifold) in $\mathbb{R}^d$ (or $\mathbb{C}^d$, we restrict here to the real case), is a sequence of finite subsets $X_n \subset K$ such that

$$\|p\|_{L^\infty(K)} \le C_n \|p\|_{\ell^\infty(X_n)} \;,\; \forall p \in \mathbb{P}_n^d(K) \;, \qquad (20)$$

where $C_n = \mathcal{O}(n^\alpha)$, $M_n = \text{card}(X_n) = \mathcal{O}(N^\beta)$, with $\alpha \ge 0$, and $\beta \ge 1$. Indeed, since X_n is automatically $\mathbb{P}_n^d(K)$-determining, then $M_n \ge N = \dim(\mathbb{P}_n^d(K)) = \dim(\mathbb{P}_n^d(X_n))$. In the case where $\alpha = 0$ (i.e., $C_n \le C$) we speak of an *admissible polynomial mesh*, and such a mesh is termed *optimal* when $\text{card}(X_n) = \mathcal{O}(N)$.

Polynomial meshes have interesting computational features (cf. [2]), e.g. can be extended by algebraic transforms, finite union and product, contain computable near optimal interpolation sets, and are near optimal for uniform LS approximation, namely [3, Thm. 1]

$$\|\mathcal{L}_n\| = \sup_{f \in C(K),\, f \neq 0} \frac{\|\mathcal{L}_n f\|_{L^\infty(K)}}{\|f\|_{L^\infty(K)}} \le C_n \sqrt{M_n} \;, \qquad (21)$$

where $\mathcal{L}_n$ is the $\ell^2(X_n)$-orthogonal projection operator $C(K) \to \mathbb{P}_n^d(K)$. From (21) we get in a standard way the uniform error estimate for $f \in C(K)$

$$\|f - \mathcal{L}_n f\|_{L^\infty(K)} \le \left(1 + C_n \sqrt{M_n}\right) E_n(f) \;. \qquad (22)$$

These properties show that polynomial meshes are good models of multivariate compact sets, in the context of polynomial approximation. Unfortunately, several computable meshes, even optimal meshes, have high cardinality already for $d = 2$ or $d = 3$, e.g. on Markov compact sets [3, Thm. 5], on polygons/polyhedra with many vertices, or star-shaped domains with smooth boundary [10]. As already observed, in the applications of LS approximation it is very important to reduce the sampling cardinality, especially when the sampling process is difficult or costly. Thus we may think to apply CATCH subsampling to polynomial meshes, in view of CATCHLS approximation, as in the previous section. In particular, it results that we can substantially keep the uniform approximation features of the polynomial mesh. We give the main result in the following

Proposition 3: [14] *Let X_n be a polynomial mesh (cf. (20)) and let the assumptions of Proposition 2 be satisfied with $X = X_n$.*

Then, the following estimate hold for the uniform norm of the CATCHLS operator

$$\|\mathcal{L}_n^c\| \leq \mathscr{C}_n(\varepsilon)\sqrt{M_n}\;, \quad \mathscr{C}_n(\varepsilon) = C_n\,\mathcal{A}_{M_n}(\varepsilon)\;, \tag{23}$$

provided that $\varepsilon\sqrt{M_n} < 1$, where $\mathcal{L}_n^c f$ is the Least Squares polynomial at the Caratheodory-Tchakaloff points $T_{2n} \subseteq X_n$. Moreover, for every $p \in \mathbb{P}_n^d(K)$

$$\|p\|_{L^\infty(K)} \leq \mathscr{C}_n(\varepsilon)\sqrt{M_n}\,\|p\|_{\ell^\infty(T_{2n})}\;. \tag{24}$$

By Proposition 3, we have that the (estimate of) the uniform norm of the CATCHLS operator has substantially the same size of (21), and the the $2n$-degree CATCH points of a polynomial mesh are a polynomial mesh, as long as $\varepsilon\sqrt{M_n} \ll 1$.

In the example of Figure 2 we see that the CATCHLS operator norm is close to the LS operator norm, as we could expect from (21) and (23), which however turn out to be large overestimates of the actual norms.

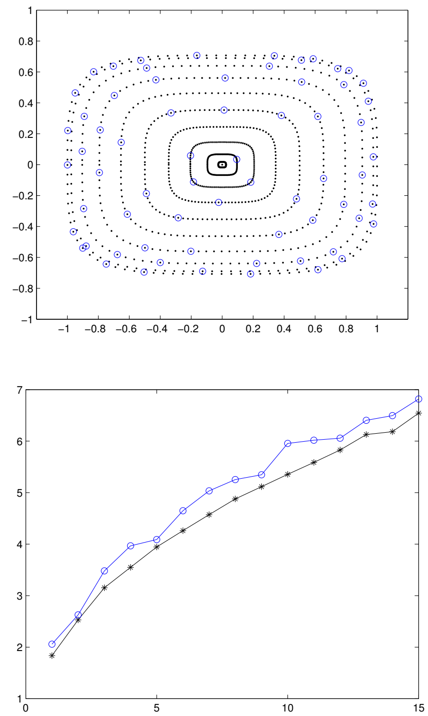

Fig. 2: Top: polynomial mesh and extracted CATCH points on a smooth convex set ($n = 5$, $C_{ratio} = 971/66 \approx 15$); bottom: numerically evaluated LS ($*$) and CATCHLS ($\circ$) uniform operator norms, for degree $n = 1, \ldots, 15$.

Acknowledgments

Work partially supported by the DOR funds and the biennial projects CPDA143275 and BIRD163015 of the University of Padova, and by the GNCS-INdAM. This research has been accomplished within the RITA (Research ITalian network on Approximation).

References

[1] L. Bittante, S. De Marchi and G. Elefante, A new quasi-Monte Carlo technique based on nonnegative least-squares and approximate Fekete points, Numer. Math. Theory Methods Appl. 9 (2016), 609–632.
[2] 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.
[3] J.P. Calvi and N. Levenberg, Uniform approximation by discrete least squares polynomials, J. Approx. Theory 152 (2008), 82–100.
[4] C. Caratheodory, ber den Variabilittsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen, Rend. Circ. Mat. Palermo 32 (1911), 193–217.
[5] M. Conforti, G. Cornuéjols and G. Zambelli, Integer programming, Graduate Texts in Mathematics 271, Springer, Cham, 2014.
[6] R.E. Curto and L.A. Fialkow, A duality proof of Tchakaloff's theorem. J. Math. Anal. Appl. 269 (2002), 519–532.
[7] S. Foucart and H. Rahut, A Mathematical Introduction to Compressive Sensing, Birkhäuser, 2013.
[8] I. Griva, S.G. Nash and A. Sofer, Linear and Nonlinear Optimization, 2nd Edition, SIAM, 2009.
[9] D. Huybrechs, Stable high-order quadrature rules with equidistant points, J. Comput. Appl. Math. 231 (2009), 933–947.
[10] A. Kroó, On optimal polynomial meshes, J. Approx. Theory 163 (2011), 1107–1124.
[11] C.L. Lawson and R.J. Hanson, Solving least squares problems. Revised reprint of the 1974 original, SIAM, Philadelphia, 1995.
[12] C. Litterer and T. Lyons, High order recombination and an application to cubature on Wiener space, Ann. Appl. Probab. 22 (2012), 1301–1327.
[13] G. Migliorati and F. Nobile, Analysis of discrete least squares on multivariate polynomial spaces with evaluations at low-discrepancy point sets, J. Complexity 31 (2015), 517–542.
[14] F. Piazzon, A. Sommariva and M. Vianello, Caratheodory-Tchakaloff Subsampling, Dolomites Res. Notes Approx. DRNA 10 (2017), 5–15.
[15] F. Piazzon, A. Sommariva and M. Vianello, CATCH: code for Caratheodory-Tchakaloff compression of multivariate discrete measures, www.math.unipd.it/∼marcov/CAAsoft.html.
[16] W. Pleśniak, Multivariate Jackson Inequality, J. Comput. Appl. Math. 233 (2009), 815–820.
[17] E.K. Ryu and S.P. Boyd, Extensions of Gauss quadrature via linear programming, Found. Comput. Math. 15 (2015), 953–971.
[18] M. Slawski, Nonnegative least squares: comparison of algorithms (paper and code), https://sites.google.com/site/slawskimartin/code.
[19] I.H. Sloan, Interpolation and Hyperinterpolation over General Regions, J. Approx. Theory 83 (1995), 238–254.
[20] A. Sommariva and M. Vianello, Compression of multivariate discrete measures and applications, Numer. Funct. Anal. Optim. 36 (2015), 1198–1223.
[21] V. Tchakaloff, Formules de cubatures mécaniques à coefficients non négatifs. (French) Bull. Sci. Math. 81 (1957), 123–134.
[22] M. Tchernychova, "Caratheodory" cubature measures, Ph.D. dissertation in Mathematics (supervisor: T. Lyons), University of Oxford, 2015.
[23] M. Vianello, Compressed sampling inequalities by Tchakaloff's theorem, Math. Inequal. Appl. 19 (2016), 395–400.