Continuous Set Packing and Near-Boolean Functions

Giovanni Rossi
Department of Computer Science and Engineering (DISI), University of Bologna, Mura Anteo Zamboni 7, 40126, Bologna, Italy

Keywords: Set Packing, Pseudo-Boolean Function, Polynomial Multilinear Extension, Local Search.

Abstract: Given a family of feasible subsets of a ground set, the packing problem is to find a largest subfamily of pairwise disjoint family members. Non-approximability renders heuristics attractive viable options, while efficient methods with worst-case guarantee are a key concern in computational complexity. This work proposes a novel near-Boolean optimization method relying on a polynomial multilinear form with variables ranging each in a high-dimensional unit simplex. These variables are the elements of the ground set, and distribute each a unit membership over those feasible subsets where they are included. The given problem is thus translated into a continuous version where the objective is to maximize a function taking values on collections of points in a unit hypercube. Maximizers are shown to always include collections of hypercube disjoint vertices, i.e. partitions of the ground set, which constitute feasible solutions for the original discrete version of the problem. A gradient-based local search in the expanded continuous domain is designed. Approximations with polynomials of bounded degree and near-Boolean coalition formation games are also finally discussed.

1 INTRODUCTION

Consider a finite set $N = \{1, \ldots, n\}$ of items to be packed into feasible subsets, where these latter constitute a family $\mathcal{F} \subseteq 2^N = \{A : A \subseteq N\}$. The problem is to find a subfamily $\mathcal{F}^* \subseteq \mathcal{F}$ of pairwise disjoint feasible subsets with largest size $|\mathcal{F}^*|$. In the weighted version, a function $w : \mathcal{F} \to \mathbb{R}_+$ identifies as optimal those such subfamilies $\mathcal{F}^*$ with maximum weight $W(\mathcal{F}^*) = \sum_{A \in \mathcal{F}^*} w(A)$. Maximizing $|\mathcal{F}^*|$ is equivalent to setting $w(A) = 1$ for all $A \in \mathcal{F}$. Thus this work proposes to use the polynomial multilinear extension, or MLE for short, of set functions (such as w) in order to evaluate families of fuzzy feasible subsets. Although unfeasible, such families shall still drive the search towards locally optimal feasible ones.

Set packing is a key combinatorial optimization problem (Korte and Vygen, 2002) extensively studied in computational complexity, where the aim is to find efficient algorithms whose output approximates optimal solutions within a provable bounded factor. In that field, the focus is placed mostly on non-approximability results for k-set packing (Trevisan, 2001), where the size of every feasible subset is no greater than some $k \ll n$ (and with unit weight for all as above). Recall that if all feasible subsets have size $k = 2$, then the problem is to find a maximal matching in a graph with vertex set N, and an efficient (i.e. with polynomial running time) algorithm capable to output an exact solution is known to exist (Papadimitriou, 1994). In fact, if $k > 2$, then k-set packing may be rephrased in terms of vertex clouring in hypergraphs, with special focus on the d-regular and k-uniform case, where every element of the ground set is present in precisely $d > 1$ feasible subsets (i.e. $|\{A : i \in A \in \mathcal{F}\}| = d$ for every $i \in N$), each of which, in turn, has size k (i.e. $|A| = k$ for every $A \in \mathcal{F}$) (Hazan et al., 2006).

Set packing aslo has important applications, among which combinatorial auctions constitute a main and lucrative example: the ground set may consist of items to be sold in bundles (or subsets) towards revenue maximization, and once bids are processed the issue may be tackled as a maximum-weight set packing problem, with maximum received bids on bundles as weights (Sandholm, 2002). Given the exponentially large size of the search space, revenue maximization often leads to use heuristics with no worst-case guarantee or, more simply, to sell each item independently but simultaneously over a sufficiently long time period (Milgrom, 2004).

The approach to set packing problems proposed in the sequel replaces standard pseudo-Boolean optimization (Boros and Hammer, 2002) with a novel near-Boolean method. While the former employs MLE to switch from $\{0,1\}$ to $[0,1]$ as the domain of each of the n variables, the proposed near-Boolean method relies on n variables ranging each in the 2^{n-1}-

Rossi, G.
Continuous Set Packing and Near-Boolean Functions.
DOI: 10.5220/0005697800840096
In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016)*, pages 84-96
ISBN: 978-989-758-173-1

set of extreme points of a unit simplex, and employs MLE to include the continuum provided by the whole simplex. The n variables correspond to the elements $i \in N$ of the ground set, while the extreme points of each simplex are indexed by those (feasible) subsets where each element is included. Then, the MLE of the resulting near-Boolean function evaluates collections of fuzzy subsets of N or, equivalently, fuzzy subfamilies of feasible subsets $A \in \mathcal{F}$. The objective function to be maximized takes thus values over the n-product of $2^{n-1}-1$-dimensional unit simplices, allowing to design a flexible gradient-based local search.

The following section comprenshively details the framework for the full-dimensional case $\mathcal{F} = 2^N$. This seems useful in general and also allows to clearly see next that by simply introducing the empty set $\emptyset$ and all n singletons $\{i\} \in 2^N$ into the family $\mathcal{F}$ of feasible subsets (with null weights $w(\emptyset) = 0 = w(\{i\}$ if $\{i\} \notin \mathcal{F}$) the whole class of set packing problems may be framed within the proposed method. The gradient-based local search differs when switching from the full-dimensional case to the lower-dimensional one $\mathcal{F} \subset 2^N$, as with the latter a cost function $c : \mathcal{F} \to \mathbb{N}$ also enters the picture, in line with greedy approaches to weighted set packing (Chandra and Halldorsson, 2001). The cost $c(A) = |\{B : B \in \mathcal{F}, B \cap A \neq \emptyset\}|$ of including a feasible subset in the packing is the number of members with which it has non-empty intersection (itself included, hence $c(A) \in \mathbb{N}, A \in \mathcal{F}$).

Note that maximum-weight set packing may be tackled through constrained maximization of standard pseudo-Boolean function $v : \{0,1\}^{|\mathcal{F}|} \to \mathbb{R}_+$

given by $v\left(x_{A_1}, \ldots, x_{A_{|\mathcal{F}|}}\right) = \sum_{1 \leq k \leq |\mathcal{F}|} x_{A_k} w(A_k)$, s.t.

$A_k \cap A_l \neq \emptyset \Rightarrow x_{A_k} + x_{A_l} \leq 1$ for all $1 \leq k < l \leq |\mathcal{F}|$,

where $x_A \in \{0,1\}$ for all $A \in \mathcal{F} = \{A_1, \ldots A_{|\mathcal{F}|}\}$. Also, v can be replaced with $xMx \simeq v$, where M is a suitable $|\mathcal{F}| \times |\mathcal{F}|$-matrix and $x = (x_{A_1}, \ldots, x_{A_{|\mathcal{F}|}})$ (Alidaee et al., 2008). An heuristic then finds a constrained maximizer x^*, while the corresponding solution is $\mathcal{F}^* = \{A : x^*_A = 1\}$. This differs from what is proposed here, in many respects, the most evident of which being that v has $|\mathcal{F}|$ constrained Boolean variables, while the expanded MLE developed below has n unconstrained near-Boolean variables.

2 FULL-DIMENSIONAL CASE

The 2^n-set $\{0,1\}^n$ of vertices of the n-dimensional unit hypercube $[0,1]^n$ corresponds one-to-one to the (power) set 2^N of subsets $A \subseteq N$ through characteristic functions $\chi_A : N \to \{0,1\}, A \in 2^N$ defined by $\chi_A(i) = 1$ if $i \in A$ and $\chi_A(i) = 0$ if $i \in N \backslash A = A^c$, while collection $\{\zeta(A, \cdot) : A \in 2^N\}$ is a linear basis of the vector space $\mathbb{R}^{2^n}$ of real-valued functions w on 2^N, where zeta function $\zeta : 2^N \times 2^N \to \mathbb{R}$ is the element of the incidence algebra (Rota, 1964b; Aigner, 1997) of Boolean lattice $(2^N, \cap, \cup)$ defined by $\zeta(A,B) = 1$ if $B \supseteq A$ and $\zeta(A,B) = 0$ if $B \not\supseteq A$. Linear combination $w(B) = \sum_{A \in 2^N} \mu^w(A)\zeta(A,B) = \sum_{A \subseteq B} \mu^w(A)$ for $B \in 2^N$ applies to any w, with Möbius inversion $\mu^w : 2^N \to \mathbb{R}$ uniquely given by ($\subset$ is strict inclusion) $\mu^w(A) =$

$$= \sum_{B \subseteq A} (-1)^{|A \backslash B|} w(B) \ \left(\text{with } \zeta(B,A) = (-1)^{|A \backslash B|}\right)$$

$$= w(A) - \sum_{B \subset A} \mu^w(B) \ \text{(recursion, with } w(\emptyset) = 0\text{)}.$$

Given this essential combinatorial *"analog of the fundamental theorem of the calculus"* (Rota, 1964b), the MLE $f^w : [0,1]^n \to \mathbb{R}$ of w takes values $w(B) =$

$$= f^w(\chi_B) = \sum_{A \in 2^N} \left(\prod_{i \in A} \chi_B(i) \right) \mu^w(A) = \sum_{A \subseteq B} \mu^w(A)$$

$$\text{on vertices, and } f^w(q) = \sum_{A \in 2^N} \left(\prod_{i \in A} q_i \right) \mu^w(A) \quad (1)$$

on any point $q = (q_1, \ldots, q_n) \in [0,1]^n$. Conventionally, $\prod_{i \in \emptyset} q_i := 1$ (Boros and Hammer, 2002, p. 157).

Let $2^N_i = \{A : i \in A \in 2^N\} = \{A_1, \ldots, A_{2^{n-1}}\}$ be the 2^{n-1}-set of subsets containing each $i \in N$. Simplex

$$\Delta_i = \left\{ \left(q_i^{A_1}, \ldots, q_i^{A_{2^{n-1}}}\right) \in \mathbb{R}^{2^{n-1}}_+ : \sum_{1 \leq k \leq 2^{n-1}} q_i^{A_k} = 1 \right\}$$

has dimension $2^{n-1} - 1$ and generic point $q_i \in \Delta_i$.

Definition 1. *A fuzzy cover* $\mathbf{q}$ *specifies a membership distribution for each* $i \in N$ *over the* 2^{n-1} *subsets containing it, i.e.* $\mathbf{q} = (q_1, \ldots, q_n) \in \Delta_N = \times_{1 \leq i \leq n} \Delta_i$.

Equivalently, $\mathbf{q} = \{q^A : \emptyset \neq A \in 2^N, q^A \in [0,1]^n\}$ is a $2^n - 1$-set whose elements $q^A = \left(q_1^A, \ldots, q_n^A\right)$ are n-vectors corresponding to non-empty subsets $A \in 2^N$ and specifying a membership q_i^A for each $i \in N$, with $q_i^A \in [0,1]$ if $i \in A$ while $q_i^A = 0$ if $i \in A^c$. Fuzzy covers being collections of points in $[0,1]^n$, and the MLE f^w of w allowing precisely to evaluate such points, the global worth $W(\mathbf{q})$ of $\mathbf{q} \in \Delta_N$ is the sum over all $q^A, A \in 2^N$ of $f^w(q^A)$ as defined by (1). That is,

$$W(\mathbf{q}) = \sum_{A \in 2^N} f^w(q^A) = \sum_{A \in 2^N} \left[\sum_{B \subseteq A} \left(\prod_{i \in B} q_i^A \right) \mu^w(B) \right],$$

$$\text{or } W(\mathbf{q}) = \sum_{A \in 2^N} \left[\sum_{B \supseteq A} \left(\prod_{i \in A} q_i^B \right) \right] \mu^w(A). \quad (2)$$

Example 2. *For $N=\{1,2,3\}$, consider w defined by $w(\{1\})=w(\{2\})=w(\{3\})=0.2$, $w(\{1,2\})=0.8$, $w(\{1,3\})=0.3$, $w(\{2,3\})=0.6$, $w(N)=0.7$. Membership distributions of elements $i=1,2,3$ over 2^N_i are $q_1\in\Delta_1, q_2\in\Delta_2, q_3\in\Delta_3$,*

$$q_1=\begin{pmatrix} q_1^1\\ q_1^{12}\\ q_1^{13}\\ q_1^N\end{pmatrix},\ q_2=\begin{pmatrix} q_2^2\\ q_2^{12}\\ q_2^{23}\\ q_2^N\end{pmatrix},\ q_3=\begin{pmatrix} q_3^3\\ q_3^{13}\\ q_3^{23}\\ q_3^N\end{pmatrix}.$$

If $\hat{q}_1^{12}=\hat{q}_2^{12}=1$, then any membership $q_3\in\Delta_3$ yields

$$\begin{aligned} W(\hat{q}_1,\hat{q}_2,q_3) &= w(\{1,2\}) \\ &+ \left(q_3^3+q_3^{13}+q_3^{23}+q_3^N\right)\mu^w(\{3\}) \\ &= w(\{1,2\})+w(\{3\})=1. \end{aligned}$$

This means that there is a continuum of fuzzy covers achieving maximum worth, i.e. 1. In order to select the one $\hat{\mathbf{q}}=(\hat{q}_1,\hat{q}_2,\hat{q}_3)$ where $\hat{q}_3^3=1$, attention must be placed only on exact ones, defined hereafter.

For any two fuzzy covers $\mathbf{q}=\{q^A:\emptyset\neq A\in 2^N\}$ and $\hat{\mathbf{q}}=\{\hat{q}^A:\emptyset\neq A\in 2^N\}$, define $\hat{\mathbf{q}}$ to be a *shrinking* of $\mathbf{q}$ if there is a subset A, with $\sum_{i\in A} q_i^A>0$ and

$$\begin{aligned} \hat{q}_i^B &= \begin{cases} q_i^B \text{ if } B\not\subseteq A \\ 0 \text{ if } B=A \end{cases} \quad \text{for all } B\in 2^N, i\in N, \\ \sum_{B\subset A}\hat{q}_i^B &= q_i^A+\sum_{B\subset A} q_i^B \text{ for all } i\in A. \end{aligned}$$

In words, a shrinking reallocates the whole membership mass $\sum_{i\in A} q_i^A>0$ from $A\in 2^N$ to all proper subsets $B\subset A$, involving all and only those elements $i\in A$ with strictly positive membership $q_i^A>0$.

Definition 3. *Fuzzy cover $\boldsymbol{q}\in\Delta_N$ is exact as long as $W(\boldsymbol{q})\neq W(\hat{\boldsymbol{q}})$ for all shrinkings $\hat{\boldsymbol{q}}$ of $\boldsymbol{q}$.*

Proprosition 4. *If $\boldsymbol{q}$ is exact, then for all $A\in 2^N$*

$$\left|\left\{i\in A: q_i^A>0\right\}\right|\in\{0,|A|\}.$$

Proof. For $\emptyset\subset A^+(\mathbf{q})=\left\{i:q_i^A>0\right\}\subset A\in 2^N$, with $\alpha=|A^+(\mathbf{q})|>1$, notice that

$f^w(q^A)=\sum_{B\subseteq A^+(\mathbf{q})}\left(\prod_{i\in B} q_i^A\right)\mu^w(B)$. Let shrinking $\hat{\mathbf{q}}$,

with $\hat{q}^{B'}=q^{B'}$ if $B'\notin 2^{A^+(\mathbf{q})}$, satisfy conditions

$$\sum_{B\in 2_i^N\cap 2^{A^+(\mathbf{q})}}\hat{q}_i^B=q_i^A+\sum_{B\in 2_i^N\cap 2^{A^+(\mathbf{q})}} q_i^B \text{ for all } i\in A^+(\mathbf{q})$$

and $\prod_{i\in B}\hat{q}_i^B=\prod_{i\in B} q_i^B+\prod_{i\in B} q_i^A$ for all $B\in 2^{A^+(\mathbf{q})}, |B|>1$.

These are $2^\alpha-1$ equations with $\sum_{1\leq k\leq\alpha} k\binom{\alpha}{k}>2^\alpha$ variables $\hat{q}_i^B, B\subseteq A^+(\mathbf{q})$. Thus there is a continuum of solutions, each providing precisely a shrinking $\hat{\mathbf{q}}$

where $\sum_{B\in 2^{A^+(\mathbf{q})}} f^w(\hat{q}^B)=f^w(q^A)+\sum_{B\in 2^{A^+(\mathbf{q})}} f^w(q^B)$.

This entails that $\mathbf{q}$ is not exact. □

Partitions (Aigner, 1997) $P=\{A_1,\ldots,A_{|P|}\}\subset 2^N$ of N are families of pairwise disjoint subsets called blocks, i.e. $A_k\cap A_l=\emptyset, 1\leq k<l\leq|P|$, with union $N=\cup_{1\leq k\leq|P|}A_k$. Any P corresponds to the collection $\{\chi_A:A\in P\}$ of those $|P|$ hypercube vertices identified by the characteristic functions of its blocks (see above). Partitions P can thus be seen as $\mathbf{p}\in\Delta_N$ where $p_i^A=1$ for all $A\in P, i\in A$, i.e. exact fuzzy covers where each $i\in N$ concentrates its whole membershisp on a unique $A\in 2_i^N$, thus justifying the following.

Definition 5. *Fuzzy partitions are exact fuzzy covers.*

Ojective function $W:\Delta_N\to\mathbb{R}$ includes among its extremizers (non-fuzzy) partitions. This expands a result from pseudo-Boolean optimization. Denote by $ex(\Delta_i)$ the 2^{n-1}-set of extreme points of Δ_i. For $\mathbf{q}\in\Delta_N, i\in N$, let $\mathbf{q}=q_i|\mathbf{q}_{-i}$, with $q_i\in\Delta_i$ as well as $\mathbf{q}_{-i}\in\Delta_{N\backslash i}=\times_{j\in N\backslash i}\Delta_j$. Then, for any w,

$$W(\mathbf{q})=\sum_{A\in 2_i^N} f^w(q^A)+\sum_{A'\in 2^N\backslash 2_i^N} f^w(q^{A'})=$$

$$=\sum_{A\in 2_i^N}\sum_{B\subseteq A\backslash i}\left(\prod_{j\in B} q_j^A\right)\left(q_i^A\mu^w(B\cup i)+\mu^w(B)\right)+$$

$$+\sum_{A'\in 2^N\backslash 2_i^N}\sum_{B'\subseteq A'}\left(\prod_{j'\in B'} q_{j'}^{A'}\right)\mu^w(B')$$

at all $\mathbf{q}\in\Delta_N$ and for all $i\in N$. Now define

$$W_i(q_i|\mathbf{q}_{-i})=\sum_{A\in 2_i^N} q_i^A\left[\sum_{B\subseteq A\backslash i}\left(\prod_{j\in B} q_j^A\right)\mu^w(B\cup i)\right],$$

$$\begin{aligned} W_{-i}(\mathbf{q}_{-i}) &= \sum_{A\in 2_i^N}\left[\sum_{B\subseteq A\backslash i}\left(\prod_{j\in B} q_j^A\right)\mu^w(B)\right]+ \\ &+\sum_{A'\in 2^N\backslash 2_i^N}\left[\sum_{B'\subseteq A'}\left(\prod_{j'\in B'} q_{j'}^{A'}\right)\mu^w(B')\right], \end{aligned}$$

$$\text{yielding } W(\mathbf{q})=W_i(q_i|\mathbf{q}_{-i})+W_{-i}(\mathbf{q}_{-i}). \quad (3)$$

Proprosition 6. *For all $\boldsymbol{q}\in\Delta_N$, there are $\underline{\boldsymbol{q}},\overline{\boldsymbol{q}}\in\Delta_N$*

$$\text{such that } \begin{cases} (i)\ W(\underline{\boldsymbol{q}})\leq W(\boldsymbol{q})\leq W(\overline{\boldsymbol{q}}) \text{ and,} \\ (ii)\ \underline{q}_i,\overline{q}^i\in ex(\Delta_i) \text{ for all } i\in N. \end{cases}$$

Proof. For $i\in N, \mathbf{q}_{-i}\in\Delta_{N\backslash i}$, define $w_{\mathbf{q}_{-i}}:2_i^N\to\mathbb{R}$ by

$$w_{\mathbf{q}_{-i}}(A)=\sum_{B\subseteq A\backslash i}\left(\prod_{j\in B} q_j^A\right)\mu^w(B\cup i). \quad (4)$$

Let $\mathbb{A}^+_{\mathbf{q}_{-i}}=\arg\max w_{\mathbf{q}_{-i}}$ and $\mathbb{A}^-_{\mathbf{q}_{-i}}=\arg\min w_{\mathbf{q}_{-i}}$, with $\mathbb{A}^+_{\mathbf{q}_{-i}}\neq\emptyset\neq\mathbb{A}^-_{\mathbf{q}_{-i}}$ at all $\mathbf{q}_{-i}$. Most importantly,

$$W_i(q_i|\mathbf{q}_i)=\sum_{A\in 2_i^N}\left(q_i^A\cdot w_{\mathbf{q}_{-i}}(A)\right)=\langle q_i,w_{\mathbf{q}_{-i}}\rangle, \quad (5)$$

where $\langle \cdot, \cdot \rangle$ denotes scalar product. Thus for given membership distributions of all $j \in N \backslash i$, global worth is affected by i's membership distribution through a scalar product. In order to maximize (or minimize) W by suitably choosing q_i for given $\mathbf{q}_{-i}$, the whole of i's membership mass must be placed over $\mathbb{A}^{+}_{\mathbf{q}_{-i}}$ (or $\mathbb{A}^{-}_{\mathbf{q}_{-i}}$), anyhow. Hence there are precisely $|\mathbb{A}^{+}_{\mathbf{q}_{-i}}| > 0$ (or $|\mathbb{A}^{-}_{\mathbf{q}_{-i}}| > 0$) available extreme points of Δ_i. The following procedure selects (arbitrarily) one of them.

ROUNDUP$(w, \mathbf{q})$

Initialize: Set $t = 0$ and $\mathbf{q}(0) = \mathbf{q}$.

Loop: While there is a $i \in N$ with $q_i(t) \notin ex(\Delta_i)$, set $t = t+1$ and:

(a) select some $A^* \in \mathbb{A}^{+}_{\mathbf{q}_{-i}(t)}$,

(b) define, for all $j \in N, A \in 2^N$,

$$q^A_j(t) = \begin{cases} q^A_j(t-1) \text{ if } j \neq i \\ 1 \text{ if } j = i \text{ and } A = A^* \\ 0 \text{ otherwise} \end{cases} .$$

Output: Set $\overline{\mathbf{q}} = \mathbf{q}(t)$.

Every change $q^A_i(t-1) \neq q^A_i(t) = 1$ (for any $i \in N, A \in 2^N_i$) induces a non-decreasing variation $W(\mathbf{q}(t)) - W(\mathbf{q}(t-1)) \geq 0$. Hence, the sought $\overline{\mathbf{q}}$ is provided in at most n iterations. Analogously, replacing $\mathbb{A}^{+}_{\mathbf{q}_{-i}}$ with $\mathbb{A}^{-}_{\mathbf{q}_{-i}}$ yields the sought minimizer $\underline{\mathbf{q}}$ (see also (Boros and Hammer, 2002, p. 163)). □

Remark 7. *For $i \in N, A \in 2^N_i$, if all $j \in A \backslash i \neq \emptyset$ satisfy $q^A_j = 1$, then (4) yields $w_{\boldsymbol{q}_{-i}}(A) = w(A) - w(A \backslash i)$, while $w_{\boldsymbol{q}_{-i}}(\{i\}) = w(\{i\})$ regardless of $\boldsymbol{q}_{-i}$.*

Corollary 8. *Some partition P satisfies $W(\boldsymbol{p}) \geq W(\boldsymbol{q})$ for all $\boldsymbol{q} \in \Delta_N$, with $W(\boldsymbol{p}) = \sum_{A \in P} w(A)$.*

Proof. Follows from propositions 4 and 6, with the above notation associating $\mathbf{p} \in \Delta_N$ to partition P. □

Defining global maximizers is clearly immediate.

Definition 9. *Fuzzy partition $\hat{\boldsymbol{q}} \in \Delta_N$ is a global maximizer if $W(\hat{\boldsymbol{q}}) \geq W(\boldsymbol{q})$ for all $\boldsymbol{q} \in \Delta_N$.*

Concerning local maximizers, consider a vector $\omega = (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n_{++}$ of strictly positive weights, with $\omega_N = \sum_{j \in N} \omega_j$, and focus on the (Nash) equilibrium (Mas-Colell et al., 1995) of the game with elements $i \in N$ as players, each strategically choosing its membership distribution $q_i \in \Delta_i$ while being rewarded with fraction $\frac{\omega_i}{\omega_N} W(q_1, \ldots, q_n)$ of the global worth attained at any strategy profile $(q_1, \ldots, q_n) = \mathbf{q} \in \Delta_N$.

Definition 10. *Fuzzy partition $\hat{\boldsymbol{q}} \in \Delta_N$ is a local maximizer if for all $q_i \in \Delta_i$ and all $i \in N$ inequality $W_i(\hat{q}_i | \hat{\boldsymbol{q}}_{-i}) \geq W_i(q_i | \hat{\boldsymbol{q}}_{-i})$ holds (see (3)).*

This definition of local maximizer entails that the *neighborhood* $\mathcal{N}(\mathbf{q}) \subset \Delta_N$ of any $\mathbf{q} \in \Delta_N$ is

$$\mathcal{N}(\mathbf{q}) = \bigcup_{i \in N} \left\{ \tilde{\mathbf{q}} : \tilde{\mathbf{q}} = \tilde{q}_i | \mathbf{q}_{-i}, \tilde{q}_i \in \Delta_i \right\}.$$

Definition 11. *The (i,A)-derivative of W at $\boldsymbol{q} \in \Delta_N$ is*

$$\partial W(\boldsymbol{q}) / \partial q^A_i = W(\overline{\boldsymbol{q}}(i,A)) - W(\underline{\boldsymbol{q}}(i,A)) =$$

$$= W_i\left(\overline{q}_i(i,A) | \overline{\boldsymbol{q}}_{-i}(i,A)\right) - W_i\left(\underline{q}_i(i,A) | \underline{\boldsymbol{q}}_{-i}(i,A)\right),$$

with $\overline{\boldsymbol{q}}(i,A) = \left(\overline{q}_1(i,A), \ldots, \overline{q}_n(i,A)\right)$ given by

$$\overline{q}^B_j(i,A) = \begin{cases} q^B_j \text{ for all } j \in N \backslash i, B \in 2^N_j \\ 1 \text{ for } j = i, B = A \\ 0 \text{ for } j = i, B \neq A \end{cases} ,$$

and $\underline{\boldsymbol{q}}(i,A) = \left(\underline{q}_1(i,A), \ldots, \underline{q}_n(i,A)\right)$ given by

$$\underline{q}^B_j(i,A) = \begin{cases} q^B_j \text{ for all } j \in N \backslash i, B \in 2^N_j \\ 0 \text{ for } j = i \text{ and all } B \in 2^N_i \end{cases} ,$$

thus $\nabla W(\boldsymbol{q}) = \{\partial W(\boldsymbol{q}) / \partial q^A_i : i \in N, A \in 2^N_i\} \in \mathbb{R}^{n2^{n-1}}$ is the (full) gradient of W at $\boldsymbol{q}$. The i-gradient $\nabla_i W(\boldsymbol{q}) \in \mathbb{R}^{2^{n-1}}$ of W at $\boldsymbol{q} = q_i | \boldsymbol{q}_{-i}$ is set function $\nabla_i W(\boldsymbol{q}) : 2^N_i \to \mathbb{R}$ defined by $\nabla_i W(\boldsymbol{q})(A) = w_{\boldsymbol{q}_{-i}}(A)$ for all $A \in 2^N_i$, where $w_{\boldsymbol{q}_{-i}}$ is given by (4).

Remark 12. *Membership distribution $\underline{q}_i(i,A)$ is the null one: its 2^{n-1} entries are all 0, hence $\underline{q}_i(i,A) \notin \Delta_i$.*

The setting obtained thus far allows to conceive searching for a local maximizer partition $\mathbf{p}^*$ from given fuzzy partition $\mathbf{q}$ as initial candidate solution, and while maintaing the whole search within the continuum of fuzzy partitions. This idea may be specified in alternative ways yielding different local search methods. One possibility is the following.

LOCALSEARCH$(w, \mathbf{q})$

Initialize: Set $t = 0$ and $\mathbf{q}(0) = \mathbf{q}$, with requirement $|\{i : q^A_i > 0\}| \in \{0, |A|\}$ for all $A \in 2^N$.

Loop 1: While $0 < \sum_{i \in A} q^A_i(t) < |A|$ for a $A \in 2^N$, set $t = t+1$ and

(a) select a $A^*(t) \in 2^N$ such that

$$\sum_{i \in A^*(t)} w_{\mathbf{q}_{-i}(t-1)}(A^*(t)) \geq \sum_{j \in B} w_{\mathbf{q}_{-j}(t-1)}(B)$$

for all $B \in 2^N$ such that $0 < \sum_{i \in B} q^B_j(t) < |B|$,

(b) for $i \in A^*(t)$ and $A \in 2^N_i$, define

$$q^A_i(t) = \begin{cases} 1 \text{ if } A = A^*(t), \\ 0 \text{ if } A \neq A^*(t), \end{cases}$$

(c) for $j \in N \backslash A^*(t)$ and $A \in 2^N_j$ with $A \cap A^*(t) = \emptyset$, define $q^A_j(t) = q^A_j(t-1) +$

$$+ \left(w(A) \sum_{\substack{B \in 2^N_j \\ B \cap A^*(t) \neq \emptyset}} q^B_j(t-1) \right) \left(\sum_{\substack{B' \in 2^N_j \\ B' \cap A^*(t) = \emptyset}} w(B') \right)^{-1}$$

(d) for $j \in N \backslash A^*(t)$ and $A \in 2^N_j$ with $A \cap A^*(t) \neq \emptyset$, define

$$q^A_j(t) = 0.$$

Loop 2: While $q^A_i(t) = 1, |A| > 1$ for a $i \in N$ and $w(A) < w(\{i\}) + w(A \backslash i)$, set $t = t+1$ and define:

$$\begin{aligned} q^{\hat{A}}_i(t) &= \begin{cases} 1 \text{ if } |\hat{A}| = 1 \\ 0 \text{ otherwise} \end{cases} \text{ for all } \hat{A} \in 2^N_i, \\ q^B_j(t) &= \begin{cases} 1 \text{ if } B = A \backslash i \\ 0 \text{ otherwise} \end{cases} \text{ for all } j \in A \backslash i, B \in 2^N_j, \\ q^{\hat{B}}_{j'}(t) &= q^{\hat{B}}_{j'}(t-1) \text{ for all } j' \in A^c, \hat{B} \in 2^N_{j'}. \end{aligned}$$

Output: Set $\mathbf{q}^* = \mathbf{q}(t)$.

Both ROUNDUP above and LOCALSEARCH yield a sequence $\mathbf{q}(0), \ldots, \mathbf{q}(t^*) = \mathbf{q}^*$ where $q^*_i \in ex(\Delta_i)$ for all $i \in N$. In the former at the end of each iteration t the novel $\mathbf{q}(t) \in \mathcal{N}(\mathbf{q}(t-1))$ is in the neighborhood of its predecessor. In the latter $\mathbf{q}(t) \notin \mathcal{N}(\mathbf{q}(t-1))$ in general, as in $|P| \leq n$ iterations of *Loop 1* a partition $\{A^*(1), \ldots, A^*(|P|)\} = P$ is generated. Selected subsets $A^*(t) \in 2^N, t = 1, \ldots, |P|$ are any of those where the sum over members $i \in A^*(t)$ of $(i, A^*(t))$-derivatives $\partial W(\mathbf{q}(t-1)) / \partial q^{A^*(t)}_i(t-1)$ is maximal. Once a block $A^*(t)$ is selected, then lines (c) and (d) make all elements $j \in N \backslash A^*(t)$ redistribute the entire membership mass currently placed on subsets $A' \in 2^N_j$ with non-empty intersection $A' \cap A^*(t) \neq \emptyset$ over those remaining $A \in 2^N_j$ such that, conversely, $A \cap A^*(t) = \emptyset$. The redistribution is such that each of these latter gets a fraction $w(A) / \sum_{B \in 2^N_j : B \cap A^*(t) = \emptyset} w(B)$ of the newly freed membership mass $\sum_{A' \in 2^N_j : A' \cap A^*(t) \neq \emptyset} q^{A'}_j(t-1)$. The subsequent *Loop 2* checks whether the partition generated by *Loop 1* may be improved by exctracting some elements from existing blocks and putting them in singleton blocks of the final output. In the limit, set function w may be such that for some element $i \in N$ global worth decreases when the element joins any subset $A \in 2^N_i, |A| > 1$, that is to say $w(A) - w(A \backslash i) - w(\{i\}) = \sum_{B \in 2^A \backslash 2^{A \backslash i} : |B| > 1} \mu^w(B) < 0$.

Proprosition 13. LOCALSEARCH$(W, \boldsymbol{q})$ *outputs a local maximizer* $\boldsymbol{q}^*$.

Proof. It is plain that the output is a partition P or, with the notation of corollary 8 above, $\mathbf{q}^* = \mathbf{p}$. Accordingly, any element $i \in N$ is either in a singleton block $\{i\} \in P$ or else in a block $A \in P, i \in A$ such that $|A| > 1$. In the former case, any membership reallocation deviating from $p^{\{i\}}_i = 1$, given memberships $p_j, j \in N \backslash i$, yields a cover (fuzzy or not) where global worth is the same as at $\mathbf{p}$, because $\prod_{j \in B \backslash i} p^B_j = 0$ for all $B \in 2^N_i \backslash \{i\}$ (see example 2 above). In the latter case, any membership reallocation q_i deviating from $p^A_i = 1$ (given memberships $p_j, j \in N \backslash i$) yields a cover which is best seen by distinguishing between $2^N_i \backslash A$ and A. Also recall that $w(A) - w(A \backslash i) = \sum_{B \in 2^A \backslash 2^{A \backslash i}} \mu^w(B)$. Again, all membership mass $\sum_{B \in 2^N_i \backslash A} q^B_i > 0$ simply collapses on singleton $\{i\}$ because $\prod_{j \in B \backslash i} p^B_j = 0$ for all $B \in 2^N_i \backslash A$. Hence $W(\mathbf{p}) - W(q_i | \mathbf{p}_{-i}) = w(A) - w(\{i\}) +$

$$- \left(q^A_i \sum_{B \in 2^A \backslash 2^{A \backslash i} : |B| > 1} \mu^w(B) + \sum_{B' \in 2^{A \backslash i}} \mu^w(B') \right) =$$

$$= \left(p^A_i - q^A_i \right) \sum_{B \in 2^A \backslash 2^{A \backslash i} : |B| > 1} \mu^w(B).$$

Now assume that $\mathbf{q}$ is *not* a local maximizer, i.e. $W(\mathbf{p}) - W(q_i | \mathbf{p}_{-i}) < 0$. Since $p^A_i - q^A_i > 0$ (because $p^A_i = 1$ and $q_i \in \Delta_i$ is a deviation from p_i), then

$$\sum_{B \in 2^A \backslash 2^{A \backslash i} : |B| > 1} \mu^w(B) = w(A) - w(A \backslash i) - w(\{i\}) < 0.$$

Hence $\mathbf{q}$ cannot be the output of *Second Loop*. $\square$

In local search methods, the chosen initial candidate solution determines what neighborhoods shall be visited. The range of the objective function in a neighborhood is a set of real values. In a neighborhood $\mathcal{N}(\mathbf{p})$ of a $\mathbf{p} \in \Delta_N$ or partition P only those $\sum_{A \in P : |A| > 1} |A|$ elements $i \in A$ in non-sigleton blocks $A \in P, |A| > 1$ can modify global worth by reallocating their membership. In view of (the proof of) proposition 13, the only admissible variations obtain by deviating from $p^A_i = 1$ with an alternative membership distribution $q^A_i \in [0, 1)$, with $W(q_i | \mathbf{p}_{-i}) - W(\mathbf{p})$ equal to $(q^A_i - 1) \sum_{B \in 2^A \backslash 2^{A \backslash i}, |B| > 1} \mu^w(B) + (1 - q^A_i) w(\{i\})$. Hence, choosing partitions as initial candidate solutions of LOCALSEARCH is evidently poor. A sensible choice should conversely allow the search to explore different neighborhoods where the objective function may range widely. A simplest example of such an initial candidate solution is $q^A_i = 2^{1-n}$ for all $A \in 2^N_i$ and all $i \in N$, i.e. the uniform distribution. On the other hand, the input of local search algorithms is commonly desired to be close to a global optimum, i.e. a maximizer in the present setting.

This translates here into the idea of defining the input by means of set function w. In this view, consider $q_i^A = w(A)/\sum_{B\in 2_i^N} w(B)$, yielding $\frac{q_i^A}{q_i^B} = \frac{w(A)}{w(B)} = \frac{q_j^A}{q_j^B}$ for all $A, B \in 2_i^N \cap 2_j^N$ and all $i, j \in N$ (see lines (c), (d)).

With a suitable initial candidate solution, the search may be restricted to explore only a maximum number of fuzzy partitions, thereby containing the computational burden. In particular, if $\mathbf{q}(0)$ is the finest partition $\{\{1\},\ldots,\{n\}\}$ or $q_i^{\{i\}}(0) = 1$ for all $i \in N$, then the search does not explore any neighborhood at all, and such an input coincides with the output. More reasonably, let $\mathbb{A}_{\mathbf{q}}^{\max} = \{A_1,\ldots,A_k\}$ denote the collection of $\supseteq$-maximal subsets where input memberships are strictly positive. That is, $q_i^{A_{k'}} > 0$ for all $i \in A_{k'}, 1 \leq k' \leq k$ as well as $q_j^B = 0$ for all $B \in 2^N \backslash \left(2^{A_1} \cup \cdots \cup 2^{A_k}\right)$ and all $j \in B$. Then, the output shall be a partition P each of whose blocks $A \in P$ satisfies $A \subseteq A_{k'}$ for some $1 \leq k' \leq k$. Hence, by suitably choosing the input $\mathbf{q}$, LOCALSEARCH outputs a partition with no less than some maximum desired number $k(\mathbf{q})$ blocks.

3 LOWER-DIMENSIONAL CASE

If $\mathcal{F} \subset 2^N$, then $2_i^N \cap \mathcal{F} \neq \emptyset$ for every $i \in N$, otherwise the problem reduces to packing the proper subset $N \backslash \{i : \mathcal{F} \cap 2_i^N = \emptyset\}$ of elements contained in at least one feasible subset. As outlined in section 1, without additional notation simply let $\{\emptyset\} \in \mathcal{F} \ni \{i\}$ for all $i \in N$ with null weigths $w(\emptyset) = 0 = w(\{i\})$ if $\{i\} \notin \mathcal{F}$. Thus $(\mathcal{F}, \supseteq)$ is a poset (partially ordered set) with bottom element $\emptyset$, and weight function $w : \mathcal{F} \to \mathbb{R}_+$ has well-defined Möbius inversion $\mu^w : \mathcal{F} \to \mathbb{R}$ (Rota, 1964b). Memberships q_i distribute over $\mathcal{F}_i = 2_i^N \cap \mathcal{F} = \{A_1,\ldots,A_{|\mathcal{F}_i|}\}$, with lower($|\mathcal{F}_i|$)-dimensional unit simplices

$$\bar{\Delta}_i = \left\{ \left(q_i^{A_1},\ldots,q_i^{A_{|\mathcal{F}_i|}}\right) \in \mathbb{R}_+^{|\mathcal{F}_i|} : \sum_{1\leq k\leq A_{|\mathcal{F}_i|}} q_i^{A_k} = 1 \right\}$$

and corresponding fuzzy covers $\mathbf{q} \in \bar{\Delta}_N = \times_{1\leq i\leq n} \bar{\Delta}_i$. Note that a fuzzy cover now may maximally consist of $|\mathcal{F}| - 1$ points in the unit n-dimensional hypercube $[0,1]^n$. Accordingly, hypercube $[0,1]^n$ is replaced with $\mathcal{C}(\mathcal{F}) = co(\{\chi_A : A \in \mathcal{F}\}) \subseteq [0,1]^n$, i.e. the convex hull of feasible characteristic functions, regarded as n-vectors (Grünbaum, 2001). Recursively (with $w(\emptyset) = 0$), Möbius inversion $\mu^w : \mathcal{F} \to \mathbb{R}$ is

$$\mu^w(A) = w(A) - \sum_{B\in\mathcal{F}:B\subset A} \mu^w(B),$$

while the MLE $f^w : \mathcal{C}(\mathcal{F}) \to \mathbb{R}$ of w is

$$f^w(q^A) = \sum_{B\in\mathcal{F}\cap 2^A} \left(\prod_{i\in B} q_i^A\right) \mu^w(B).$$

Therefore, every fuzzy cover $\mathbf{q} \in \bar{\Delta}_N$ has global worth

$$W(\mathbf{q}) = \sum_{A\in\mathcal{F}} \sum_{B\in\mathcal{F}\cap 2^A} \left(\prod_{i\in B} q_i^A\right) \mu^w(B).$$

For all $i \in N, q_i \in \bar{\Delta}_i$, and $\mathbf{q}_{-i} \in \bar{\Delta}_{N\backslash i} = \underset{j\in N\backslash i}{\times} \bar{\Delta}_j$

$$W_{-i}(\mathbf{q}_{-i}) = \sum_{A\in\mathcal{F}_i} \left[\sum_{B\in\mathcal{F}\cap 2^{A\backslash i}} \left(\prod_{j\in B} q_j^A\right) \mu^w(B) \right] +$$

$$+ \sum_{A'\in\mathcal{F}\backslash\mathcal{F}_i} \left[\sum_{B'\in\mathcal{F}\cap 2^{A'}} \left(\prod_{j'\in B'} q_{j'}^{A'}\right) \mu^w(B') \right],$$

$$W_i(q_i|\mathbf{q}_{-i}) = \sum_{A\in\mathcal{F}_i} q_i^A \left[\sum_{B\in\mathcal{F}_i\cap 2^A} \left(\prod_{j\in B\backslash i} q_j^A\right) \mu^w(B) \right],$$

yielding again

$$W(\mathbf{q}) = W_i(q_i|\mathbf{q}_{-i}) + W_{-i}(\mathbf{q}_{-i}). \tag{6}$$

From (4) above, $w_{\mathbf{q}_{-i}} : \mathcal{F}_i \to \mathbb{R}$ now is

$$w_{\mathbf{q}_{-i}}(A) = \sum_{B\in\mathcal{F}_i\cap 2^A} \left(\prod_{j\in B\backslash i} q_j^A\right) \mu^w(B) \tag{7}$$

for all $i \in N$, all $A \in \mathcal{F}_i$ and all $\mathbf{q}_{-i} \in \bar{\Delta}_{N\backslash i}$.

For each $i \in N$, denote by $ex(\bar{\Delta}_i)$ the set of $|\mathcal{F}_i|$ extreme points of simplex $\bar{\Delta}_i$. Like in the full-dimensional case, at any fuzzy cover $\hat{\mathbf{q}} \in \bar{\Delta}_N$ every $i \in N$ such that $\hat{q}_i \notin ex(\bar{\Delta}_i)$ may deviate by concentrating its whole membership on some $A \in \mathcal{F}_i$ such that $w_{\hat{\mathbf{q}}_{-i}}(A) \geq w_{\hat{\mathbf{q}}_{-i}}(B)$ for all $B \in \mathcal{F}_i$. This yields a non-decreasing variation $W(q_i|\hat{\mathbf{q}}_{-i}) \geq W(\hat{\mathbf{q}})$ in global worth, with $q_i \in ex(\bar{\Delta}_i)$. When all n elements do so, one after the other while updating $w_{\mathbf{q}_{-i(t)}}$ as in ROUNDUP above, i.e. $t = 0,1,\ldots$, then the final $\mathbf{q} = (q_1,\ldots,q_n)$ satisfies $\mathbf{q} \in \underset{i\in N}{\times} ex(\bar{\Delta}_i)$. Yet cases $\mathcal{F} \subset 2^N$ and $\mathcal{F} = 2^N$ are different in terms of exactness. Specifically, consider any $\emptyset \neq A \in \mathcal{F}$ such that $|\{i : q_i^A = 1\}| \notin \{0, |A|\}$ or $A_{\mathbf{q}}^+ = \{i : q_i^A = 1\} \subset A$, with $f^w(q^A) = \sum_{B\in\mathcal{F}\cap 2^{A_{\mathbf{q}}^+}} \mu^w(B)$. Then, $\mathcal{F} \cap 2^{A_{\mathbf{q}}^+}$ is likely to admit no shrinking (see above) yielding an exact fuzzy cover with same global worth as (non-exact) $\mathbf{q}$.

Proprosition 14. *The values taken on exact fuzzy covers do not saturate the range of* $W : \bar{\Delta}_N \to \mathbb{R}_+$.

Proof. Consider this example: $N = \{1,2,3,4\}$ and $\mathcal{F} = \{N,\{4\},\{1,2\},\{1,3\},\{2,3\}\}$, with worth $w(N) = 3, w(\{4\}) = 2, w(\{i,j\}) = 1$ for $1 \leq i < j \leq 3$. Define $\mathbf{q} = (q_1,\ldots,q_4)$ by $q_4^{\{4\}} = 1 = q_i^N, i = 1,2,3$, with non-exactness $|\{i : q_i^N > 0\}| = 3 < 4 = |N|$. As

$$W(\mathbf{q}) = w(\{4\}) + \sum_{1\leq i<j\leq 3} w(\{i,j\}) = 2+1+1+1$$

and $A^+_{\mathbf{q}} = \{1,2,3\}$, for all distributions $\hat{q}_1,\hat{q}_2,\hat{q}_3$ placing membership only over feasible $B \in \mathcal{F} \cap 2^{A^+_{\mathbf{q}}}$ global worth is $W(\hat{q}_1,\hat{q}_2,\hat{q}_3,q_4) < W(\mathbf{q})$. □

This observation simply indicates that an arbitrary search for optimal fuzzy covers may yield a maximizer (global or local) which is not reducible to any feasible solution of the original set packing problem. On the other hand, such feasible solutions are partitions P all of whose blocks are feasible, and where singleton blocks with worth 0 are not included in the packing. In fact, similarly to the full-dimensional case, fairly simple conditions may be shown to be sufficient for a partition to be a local maximizer.

Definition 15. *Any $\hat{q}_i|\hat{\boldsymbol{q}}_{-i} = \hat{\boldsymbol{q}} \in \bar{\Delta}_N$ is a local maximizer of $W : \bar{\Delta}_N \to \mathbb{R}_+$ if $W_i(\hat{q}_i|\hat{\boldsymbol{q}}_{-i}) \geq W_i(q_i|\hat{\boldsymbol{q}}_{-i})$ for all $i \in N$ and all $q_i \in \bar{\Delta}_i$ (see (6) above).*

The neighborhood $\mathcal{N}(\mathbf{q}) \subset \bar{\Delta}_N$ of $\mathbf{q} \in \bar{\Delta}_N$ thus is

$$\mathcal{N}(\hat{\mathbf{q}}) = \bigcup_{i\in N}\left\{\mathbf{q} : \mathbf{q} = q_i|\hat{\mathbf{q}}_{-i}, q_i \in \bar{\Delta}_i\right\}.$$

Any partition P with $A \in \mathcal{F}$ for each block $A \in P$ clearly has associated $\mathbf{p}$ such that $\mathbf{p} \in \underset{i\in N}{\times} ex(\bar{\Delta}_i) \subset \bar{\Delta}_N$.

Proprosition 16. *Any partition P with associated $\boldsymbol{p}$ such that $\boldsymbol{p} \in \bar{\Delta}_N$ is a local maximizer if for all $A \in P$*

$$w(A) \geq w(\{i\}) + \sum_{\hat{B}\in\mathcal{F}\cap 2^{A\backslash i}} \mu^w(\hat{B}).$$

Proof. Firstly note that for all blocks $A \in P$, if any, such that $|A| = 1$ there is nothing to prove, as the summation reduces to $w(\emptyset) = 0$, and thus there only remains $w(\{i\}) \geq w(\{i\})$. Accordingly, let $A \in P$ and $|A| > 1$. For every $i \in A$, any membership reallocation $q_i \in \bar{\Delta}_i$ deviating from p_i (i.e. $p_i^A = 1$), given memberships $\mathbf{p}_{-i}$ of other elements $j \in N\backslash i$ (i.e. $\bar{\Delta}_j \ni p_j^{A'} = 1$ for all $A' \in P$ and all $j \in A'$), yields $\mathbf{q} = (q_i|\mathbf{p}_{-i}) \in \bar{\Delta}_N$ which is best analyzed by distinguishing between $\mathcal{F}_i\backslash A$ and A. In particular,

$$w(A) = w(\{i\}) + \sum_{\substack{B\in\mathcal{F}_i\cap 2^A\\|B|>1}} \mu^w(B) + \sum_{\hat{B}\in\mathcal{F}\cap 2^{A\backslash i}} \mu^w(\hat{B}).$$

All membership mass $\sum_{B\in\mathcal{F}_i\backslash A} q_i^B > 0$ collapses on singleton $\{i\}$, because $\prod_{i'\in B\backslash i} p_{i'}^B = 0$ for all $B \in \mathcal{F}_i\backslash A$ by the definition of $\mathbf{p}_{-i}$ (see example 2 above). Thus,

$$W(\mathbf{p}) - W(q_i|\mathbf{p}_{-i}) = w(A) - w(\{i\}) +$$

$$-\left(q_i^A \sum_{\substack{B\in\mathcal{F}_i\cap 2^A\\|B|>1}} \mu^w(B) + \sum_{\hat{B}\in\mathcal{F}\cap 2^{A\backslash i}} \mu^w(\hat{B})\right) =$$

$$= \left(p_i^A - q_i^A\right) \sum_{\substack{B\in\mathcal{F}_i\cap 2^A\\|B|>1}} \mu^w(B).$$

Now assume that $\mathbf{p}$ is *not* a local maximizer, i.e. $W(\mathbf{p}) - W(q_i|\mathbf{p}_{-i}) < 0$. Since $p_i^A - q_i^A > 0$ because $p_i^A = 1$ and $q_i \in \bar{\Delta}_i$ is a deviation from p_i, then

$$\sum_{\substack{B\in\mathcal{F}_i\cap 2^A\\|B|>1}} \mu^w(B) = w(A) - w(\{i\}) - \sum_{\hat{B}\in\mathcal{F}\cap 2^{A\backslash i}} \mu^w(\hat{B}) < 0$$

must hold. This contradicts precisely the premise $w(A) \geq w(\{i\}) + \sum_{\hat{B}\in\mathcal{F}\cap 2^{A\backslash i}} \mu^w(\hat{B})$ for all $A \in P$ and $i \in A$, thus completing the proof. □

4 LOCAL SEARCH WITH COST

In order to design a gradient-based local search for this lower-dimensional case, the only tool still missing is the derivative, which clearly shall reproduce definition 11 above with $\mathcal{F}_i$ in place of 2_i^N. Before that, as outlined in section 1, let $c : \mathcal{F} \to \mathbb{N}$ count the number $c(A) = |\{B : B \in \mathcal{F}, B \cap A \neq \emptyset\}|$ of feasible subsets with which each $A \in \mathcal{F}$ has non-empty intersection, itself included, i.e. $c(A) \in \{1,\ldots,|\mathcal{F}|\}$ is the cost of including A in the packing. Accordingly, the underlying poset function $\hat{w} : \mathcal{F} \to \mathbb{R}_+$ now used (still taking positive values only) incorporates both weights $w(A), A \in \mathcal{F}$ (used thus far) and costs by means of ratio $\hat{w}(A) = \frac{w(A)}{c(A)}$. The result is quasi-objective function $\hat{W} : \bar{\Delta}_N \to \mathbb{R}_+$, obtained via MLE $f^{\hat{w}} : \mathcal{C}(\mathcal{F}) \to \mathbb{R}_+$ of $\hat{w}$, i.e. $\hat{W}(\mathbf{q}) = \sum_{A\in\mathcal{F}} f^{\hat{w}}(q^A)$, and all of the above applies invariately simply replacing W with $\hat{W}$.

Definition 17. *The (i,A)-derivative of $\hat{W}$ at $\boldsymbol{q} \in \bar{\Delta}_N$,*

$$A \in \mathcal{F}_i, \text{ is } \partial\hat{W}(\boldsymbol{q})/\partial q_i^A = \hat{W}(\overline{\boldsymbol{q}}(i,A)) - \hat{W}(\underline{\boldsymbol{q}}(i,A)) =$$

$$= \hat{W}_i\left(\overline{q}_i(i,A)|\overline{\boldsymbol{q}}_{-i}(i,A)\right) - \hat{W}_i\left(\underline{q}_i(i,A)|\underline{\boldsymbol{q}}_{-i}(i,A)\right),$$

with $\overline{\boldsymbol{q}}(i,A) = \left(\overline{q}_1(i,A),\ldots,\overline{q}_n(i,A)\right)$ given by

$$\overline{q}_j^B(i,A) = \begin{cases} q_j^B \text{ for all } j \in N\backslash i, B \in \mathcal{F}_j \\ 1 \text{ for } j = i, B = A \\ 0 \text{ for } j = i, B \neq A \end{cases},$$

and $\underline{\boldsymbol{q}}(i,A) = \left(\underline{q}_1(i,A),\ldots,\underline{q}_n(i,A)\right)$ given by

$$\underline{q}_j^B(i,A) = \begin{cases} q_j^B \text{ for all } j \in N\backslash i, B \in \mathcal{F}_j \\ 0 \text{ for } j = i \text{ and all } B \in \mathcal{F}_i \end{cases}.$$

The (full) gradient of $\hat{W}$ at $\boldsymbol{q} \in \bar{\Delta}_N$ is

$$\nabla \hat{W}(\boldsymbol{q}) = \{\partial \hat{W}(\boldsymbol{q})/\partial q_i^A : i \in N, A \in \mathcal{F}_i\} \in \mathbb{R}^{\sum_{i \in N}|\mathcal{F}_i|}$$

as well as the i-gradient $\nabla_i \hat{W}(\boldsymbol{q}) \in \mathbb{R}^{|\mathcal{F}_i|}$ of $\hat{W}$ at $\boldsymbol{q} = (q_i|\boldsymbol{q}_{-i}) \in \bar{\Delta}_N$ is poset function $\nabla_i \hat{W}(\boldsymbol{q}) : \mathcal{F}_i \to \mathbb{R}$ defined by $\nabla_i \hat{W}(\boldsymbol{q})(A) = \hat{w}_{\boldsymbol{q}_{-i}}(A)$ for all $A \in \mathcal{F}_i$, where $\hat{w}_{\boldsymbol{q}_{-i}}$ is given by (7) with $\hat{w}$ in place of w. Again, membership distribution $\underline{q}_i(i,A)$ is the null one: its $|\mathcal{F}_i|$ entries are all 0, hence $\underline{q}_i(i,A) \notin \bar{\Delta}_i$.

LS-WITHCOST$(\hat{w}, \mathbf{q})$

Initialize: Set $t = 0$ and $\mathbf{q}(0) = \mathbf{q}$, with requirement $|\{i : q_i^A > 0\}| \in \{0, |A|\}$ for all $A \in \mathcal{F}, \hat{w}(A) > 0$.

Loop 1: While $0 < \sum_{i \in A} q_i^A(t) < |A|$ for a $A \in \mathcal{F}$, set $t = t + 1$ and:

(a) select a $A^*(t) \in \mathcal{F}$ such that

$$\min_{i \in A^*(t)} \hat{w}_{\mathbf{q}_{-i}(t-1)}(A^*(t)) \geq \min_{j \in B} \hat{w}_{\mathbf{q}_{-j}(t-1)}(B)$$

for all $B \in 2^N$ such that $0 < \sum_{i \in B} q_j^B(t) < |B|$,

(b) for $i \in A^*(t)$ and $A \in \mathcal{F}_i$, define

$$q_i^A(t) = \begin{cases} 1 \text{ if } A = A^*(t), \\ 0 \text{ if } A \neq A^*(t), \end{cases}$$

(c) for $j \in N \backslash A^*(t)$ and $A \in \mathcal{F}_j$ with $A \cap A^*(t) = \emptyset$, define $q_j^A(t) = q_j^A(t-1) +$

$$+ \left(\hat{w}(A) \sum_{\substack{B \in \mathcal{F}_j \\ B \cap A^*(t) \neq \emptyset}} q_j^B(t-1) \right) \left(\sum_{\substack{B' \in \mathcal{F}_j \\ B' \cap A^*(t) = \emptyset}} \hat{w}(B') \right)^{-1}$$

(d) for $j \in N \backslash A^*(t)$ and $A \in \mathcal{F}_j$ with $A \cap A^*(t) \neq \emptyset$, define

$$q_j^A(t) = 0.$$

(e) for $A \in \mathcal{F}$ with $A \cap A^*(t) = \emptyset$, update cost function by $c(A) = |\{B : B \in \mathcal{F}, B \cap A \neq \emptyset = B \cap A^*(t)\}|$ and plug it into $\hat{w}$.

Loop 2: While $q_i^A(t) = 1, |A| > 1$ for a $i \in N$ and

$$w(A) < w(\{i\}) + \sum_{\hat{B} \in \mathcal{F} \cap 2^{A \backslash i}} \mu^w(\hat{B}),$$

set $t = t + 1$ and define:

$$q_i^{\hat{A}}(t) = \begin{cases} 1 \text{ if } |\hat{A}| = 1 \\ 0 \text{ otherwise} \end{cases} \text{ for all } \hat{A} \in \mathcal{F}_i,$$

$$q_j^B(t) = \begin{cases} 1 \text{ if } B = A \backslash i \\ 0 \text{ otherwise} \end{cases} \text{ for all } j \in A \backslash i, B \in \mathcal{F}_j,$$

$$q_{j'}^{\hat{B}}(t) = q_{j'}^{\hat{B}}(t-1) \text{ for all } j' \in A^c, \hat{B} \in \mathcal{F}_{j'}.$$

Output: Set $\mathbf{q}^* = \mathbf{q}(t)$.

Both LOCALSEARCH and LS-WITHCOST generate in $|P| \leq n$ iterations of *Loop 1* a partition $\{A^*(1), \ldots, A^*(|P|)\} = P$. Now selected blocks $A^*(t) \in \mathcal{F}$, $1 \leq t \leq |P|$ are any of those feasible subsets where the minimum over elements $i \in A^*(t)$ of $(i, A^*(t))$-derivatives $\partial \hat{W}(\mathbf{q}(t-1))/\partial q_i^{A^*(t)}(t-1)$ is maximal. The following *Loop 2* again checks whether the partition generated by *Loop 1* may be improved by exctracting some elements from existing blocks and putting them in singleton blocks of the final output, which thus allows for the following.

Proprosition 18. LS-WITHCOST$(W, \boldsymbol{q})$ *outputs a local maximizer $\boldsymbol{q}^*$.*

Proof. Follows from proposition 16 since *Loop 2* deals with w, not with $\hat{w}$. □

Concerning input $\mathbf{q} = \mathbf{q}(0)$, consider again setting

$$q_i^A = \frac{\hat{w}(A)}{\sum_{B \in \mathcal{F}_i} \hat{w}(B)} \text{ for all } A \in \mathcal{F}_i, i \in N, \text{ which entails}$$

$$\frac{q_i^A}{q_i^B} = \frac{w(A)c(B)}{w(B)c(A)} = \frac{q_j^A}{q_j^B} \text{ for all } A, B \in \mathcal{F}_i \cap \mathcal{F}_j, i, j \in N.$$

Evidently, *Loop 1* may take exactly the same form as in LOCALSEARCH, that is with selected blocks $A^*(t) \in \mathcal{F}, t = 1, \ldots, |P|$ of the generated partition P being any of those feasible subsets where the sum, rather than the minimum, over elements $i \in A^*(t)$ of $(i, A^*(t))$-derivatives $\partial \hat{W}(\mathbf{q}(t-1))/\partial q_i^{A^*(t)}(t-1)$ is maximal. This possibility may be useful in those settings where set packing appears in its weighted version, while using the minimum in place of the sum may be interesting for k-uniform set packing problems (see section 1). In fact, for the k-uniform case Möbius inversion is $\mu^{\hat{w}}(A) = \frac{1}{c(A)}$ if $|A| = k$ and $\mu^{\hat{w}}(A) = 0$ if $|A| \in \{0, 1\}$ for all $A \in \mathcal{F}$ (recall the convention $\{\emptyset\} \in \mathcal{F} \ni \{i\}$ for all $i \in N$), with the cost function iteratively updated in line (e). It is also plain that in k-uniform set packing *Loop 2* is ineffective.

5 NEAR-BOOLEAN FUNCTIONS

Boolean functions (Crama and Hammer, 2011) provide key analytical tools and methods with a variety of important applications. Beyond set packing problems that here constitute the main benchmark, this section further develops the full-dimensional case detailed in section 2 with the aim to indicate additional opportunities obtained from expanding the standard framework where pseudo-Boolean models are traditionally exploited. Recall that Boolean functions of n variables have form $f : \{0,1\}^n \to \{0,1\}$, and constitute a subclass of pseudo-Boolean functions $f:\{0,1\}^n\rightarrow\mathbb{R}$, which in turn admit the unique MLE $\hat{f}:[0,1]^n\rightarrow\mathbb{R}$ over the whole n-dimensional unit hypercube extensively employed thus far. The n variables thus range each in the unit interval $[0,1]$. Such a setting is here expanded by letting each variable $i=1,\ldots,n$ range in a $2^{n-1}-1$-dimensional simplex Δ_i, with the goal to evaluate collections of fuzzy subsets of a n-set through the MLE given by (1) and (2).

Definition 19. *Near-Boolean functions of n variables have form* $F:\underset{1\leq i\leq n}{\times} ex(\Delta_i)\rightarrow\mathbb{R}$.

Following (Hammer and Holzman, 1992, p. 4), denote by $N=\{1,\ldots,n\}$ the set of indices of variables (i.e., the ground set in previous sections). As already observed, any pseudo-Boolean function has a unique expression as a multilinear polynomial $f(x_1,\ldots,x_n)=\sum_{A\subseteq N}\left(\alpha_A\prod_{i\in A}x_i\right)$ in n variables (Boros and Hammer, 2002, p. 162), since $\alpha_A, A\in 2^N$ is in fact the Möbius inversion (Rota, 1964b) of a unique set function $w:2^N\rightarrow\mathbb{R}$ such that $w(A)=f(\chi_A)$, where χ_A is the characteristic function defined in section 2.

Definition 20. *The MLE $\hat{F}$ of near-Boolean functions F has polynomial form $\hat{F}:\underset{1\leq i\leq n}{\times}\Delta_i\rightarrow\mathbb{R}$ given by expression (2) in section 2, that is*

$$\hat{F}(\boldsymbol{q})=\sum_{A\in 2^N}\left[\sum_{B\supseteq A}\left(\prod_{i\in A}q_i^B\right)\right]\mu^w(A),\text{ with (see above)}$$

$\boldsymbol{q}=(q_1,\ldots,q_n)$ *and* $q_i=(q_i^{A_1},\ldots,q_i^{A_{2^{n-1}}})\in\Delta_i$.

5.1 Bounded-degree LS Approximations

In line with (Hammer and Holzman, 1992), the issue of approximating a given near-Boolean function F by means of the least squares LS criterion amounts to determine a near-Boolean function F_k such that

$$\sum_{\mathbf{q}\in\underset{i\in N}{\times}ex(\Delta_i)}[F(\mathbf{q})-F_k(\mathbf{q})]^2 \tag{8}$$

attains its minimum over all near-Boolean functions F_k with polynomial MLE $\hat{F}_k$ of degree k, that is

$$\hat{F}_k(\mathbf{q})=\sum_{\substack{A\in 2^N\\|A|\leq k}}\left[\sum_{B\supseteq A}\left(\prod_{i\in A}q_i^B\right)\right]\mu^w(A),$$

or, equivalently stated in terms of the underlying set function w, such that $\mu^w(A)=0$ if $|A|>k$.

Near-Boolean functions F take their values on n-product $\underset{i\in N}{\times} ex(\Delta_i)$, and $|ex(\Delta_i)|=2^{n-1}$ for each $i\in N$. They might thus be regarded as points $F\in\mathbb{R}^{2^{n(n-1)}}$ in a $2^{n(n-1)}$-dimensional vector space. In view of proposition 4 formalizing exactness, this seems conceptually incorrect and with useless enumerative demand. Specifically, for every partition $P\in\mathcal{P}^N$ with associated $\mathbf{p}\in\underset{i\in N}{\times} ex(\Delta_i)$, there clearly exist many non-exact $\mathbf{q}\in\underset{i\in N}{\times} ex(\Delta_i)$ such that $F(\mathbf{q})=F(\mathbf{p})$ (see corollary 8 above). Counting these redundant extreme points of simplices appears wothless. Hence k-degree approximation is to be dealt with by replacing expression (8) with the following, applying to partitions $\mathbf{p}\leftrightarrow P$ only:

$$\sum_{\substack{\mathbf{p}\in\underset{i\in N}{\times}ex(\Delta_i)\\ \text{with }\mathbf{p}\leftrightarrow P\in\mathcal{P}^N}}[F(\mathbf{p})-F_k(\mathbf{p})]^2. \tag{9}$$

The number $|\mathcal{P}^N|$ of partitions of a n-set is given by *Bell number* $\mathcal{B}_n$ (Rota, 1964a; Aigner, 1997). Accordingly, near-Boolean functions might be regarded as points $F\in\mathbb{R}^{\mathcal{B}_n}$ in a $\mathcal{B}_n$-dimensional vector space. Still, this also is far too large, as points in such a vector space correspond in fact to generic partition functions, i.e. with Möbius inversion free to live on every partition $P\in\mathcal{P}^N$. Conversely, near-Boolean functions factually involve only partition functions $h:\mathcal{P}^N\rightarrow\mathbb{R}$ such that $h(P)=h_w(P)=\sum_{A\in P}w(A)$ for some set function $w:2^N\rightarrow\mathbb{R}$. The Möbius inversion of these partition functions lives only on the 2^n-n modular elements (Stanley, 1971) of lattice $(\mathcal{P}^N,\wedge,\vee)$, namely on those partitions with a number of non-sigleton blocks ≤ 1. When regarded as points in a vector space (i.e. expressed as a linear combination of a basis, see above) these functions may be seen as $h_w\in\mathbb{R}^{2^n-n}$. This is shown below via recursion through the Möbius inversion of *additively separable* partition functions (Gilboa and Lehrer, 1990; Gilboa and Lehrer, 1991).

It seems crucial emphasizing that while pseudo-Boolean functions admit a unique set function providing their best k-degree approximation, $0\leq k\leq n$ (Hammer and Holzman, 1992), every near-Boolean function admits a continuum of set functions w determining their unique best k-degree approximation. In particular, consider first the linear (i.e. $k=1$) case: the issue is to find a best (least squares) approximation F_1 of any given F. That is, the set function w determining F_1 has to satisfy $w(A)=\sum_{i\in A}w(\{i\})$ for all $A\in 2^N$. Then,

$$h_w(P)=\sum_{A\in P}w(A)=\sum_{A\in P}\sum_{i\in A}w(\{i\})=w(N)$$

for all $P\in\mathcal{P}^N$. Thus h_w is a constant partition function, or a valuation (Aigner, 1997) of partition lattice $(\mathcal{P}^N,\wedge,\vee)$. Also, any further linear $v:2^N\rightarrow\mathbb{R}$

such that $v(N) = w(N)$ also satisfies $h_v(P) = h_w(P)$ for all $P \in \mathcal{P}^N$. In other terms, there is a continuum of equivalent linear $v \neq w$ such that $h_w = h_v$, obtained each by distributing arbitrarily the whole of $w(N)$ over the n singletons $\{i\} \in 2^N$. Cases $k > 1$ maintain this same feature: consider a set function w such that $\mu^w(A) \neq 0$ for one or more (possibly all $\binom{n}{k}$) subsets $A \in 2^N$ such that $|A| = k$. Now fix arbitrarily n values $v(\{i\}), i \in N$ with $\sum_{i\in N} w(\{i\}) = \sum_{i\in N} v(\{i\})$. For all $A \in 2^N, |A| > 1$ Möbius inversion $\mu^v : 2^N \to \mathbb{R}$ can always be determined uniquely via recursion by

$$v(A) + \sum_{i\in A^c} v(\{i\}) = \sum_{B\subseteq A} \mu^v(B) + \sum_{i\in A^c} v(\{i\}) =$$

$$= w(A) + \sum_{i\in A^c} w(\{i\}) = \sum_{B\subseteq A} \mu^w(B) + \sum_{i\in A^c} w(\{i\}).$$

If set function w additively separates partition function h, i.e. $h = h_w$, and $v' = w - v$ is a linear set function, then $v + v'$ also additively separates h, i.e. $h_w = h_{v+v'}$. Hence, there is a continuum of equivalent set functions w and $v + v'$ available for the sought k-degree approximation F_k, but still the $\mathcal{B}_n$ values taken by F_k (more precisely, the $2^n - n$ values taken by F_k on the modular elements of partition lattice $(\mathcal{P}^N, \wedge, \vee)$) are unique and independent from the chosen set function in the continuum of set functions $w, v + v'$ available, each determining an equivalent MLE $\hat{F}$ of F.

5.2 A Continuum of Polynomials

In this work, two lattices play a central role, namely the Boolean lattice $(2^N, \cap, \cup)$ of subsets of N ordered by inclusion $\supseteq$, and the geometric lattice $(\mathcal{P}^N, \wedge, \vee)$ of partitions of N ordered by coarsening $\geqslant$ (Aigner, 1997; Stern, 1999). Both, of course, are posets (partially ordered sets), and Möbius inversion applies to any (locally finite) poset, provided a bottom element exists (Rota, 1964b). The bottom subset is $\emptyset$, while the bottom partition $P_\perp = \{\{1\}, \ldots, \{n\}\}$ is the finest one. For a lattice $(L, \wedge, \vee)$ ordered by $\geqslant$ and with generic elements $x, y, z \in L$, any lattice function $f : L \to \mathbb{R}$ has Möbius inversion $\mu^f : L \to \mathbb{R}$ given by $\mu^f(x) = \sum_{x_\perp \leqslant y \leqslant x} \mu_L(y, x) f(y)$, where $x_\perp$ is the bottom element and μ_L is the Möbius function, defined recursively on ordered pairs $(y, x) \in L \times L$ by $\mu_L(y, x) = -\sum_{y \leqslant z < x} \mu_L(z, x)$ if $y < x$ (i.e. $y \leqslant x$ and $y \neq x$) as well as $\mu_L(y, x) = 1$ if $y = x$, while $\mu_L(y, x) = 0$ if $y \not\leqslant x$. The Möbius function of the subset lattice implicitly appears since the beginning of this work, and is $\mu_{2^N}(B, A) = (-1)^{|A\setminus B|}$, with $B \subseteq A$. Concerning the Möbius function of $\mathcal{P}^N$, given any two partitions $P, Q \in \mathcal{P}^N$, if $Q < P = \{A_1, \ldots, A_{|P|}\}$, then for every block $A \in P$ there are blocks $B_1, \ldots, B_{k_A} \in Q$ such that $A = B_1 \cup \cdots \cup B_{k_A}$, with $k_A > 1$ for at least one $A \in P$. Segment $[Q, P] = \{P' : Q \leqslant P' \leqslant P\}$ is thus isomorphic to product $\times_{A\in P} \mathcal{P}(k_A)$, where $\mathcal{P}(k)$ denotes the lattice of partitions of a k-set. Accordingly, let $m_k = |\{A : k_A = k\}|$ for $k = 1, \ldots, n$. Then (Rota, 1964b, pp. 359-360),

$$\mu_{\mathcal{P}^N}(Q, P) = (-1)^{-n+\sum_{1\leq k\leq n} m_k} \prod_{1<k<n} (k!)^{m_{k+1}}.$$

If a partition function $h : \mathcal{P}^N \to \mathbb{R}$ admits a set function $v : 2^N \to \mathbb{R}$ to satisfy $h(P) = \sum_{A\in P} v(A)$ for all $P \in \mathcal{P}^N$, then it may be said to be additively separable (Gilboa and Lehrer, 1990; Gilboa and Lehrer, 1991), with the notation $h = h_v$. As already outlined, any such an additively separable partition function $h = h_v$ has Möbius inversion μ^{h_v} that lives only on the modular elements of the partition lattice, i.e. where only one block, at most, has cardinality > 1. That is to say, together with the bottom $P_\perp$ and top $P^\top = \{N\}$, all other modular elements are those partitions of the form $\{A\} \cup P_\perp^{A^c}$ for $A \in 2^N$ such that $1 < |A| < n$, where $P_\perp^{A^c}$ is the finest partition of A^c (Aigner, 1997, Ex. 13, p. 71). The total number of these modular partitions thus is $2^n - n$. The Möbius inversion of an additively separable partition function h_v is detailed hereafter; see (Gilboa and Lehrer, 1990, Prop. 4.4, p. 138 and Appendix, p. 144) and (Gilboa and Lehrer, 1991, Prop. 3.3, p. 452).

Proprosition 21. *If $h = h_v$, then $h = h_w$ for a continuum of set functions $w : 2^N \to \mathbb{R}, w \neq v$.*

Proof. Firstly, by direct substitution, for all $P \in \mathcal{P}^N$,

$$\mu^{h_v}(P) = \sum_{A\in P} \sum_{B\subseteq A} v(B) \sum_{Q\leqslant P: B\in Q} \mu_{\mathcal{P}^N}(Q, P).$$

Secondly, if $P \neq \{B\} \cup P_\perp^{B^c}$, then the recursive definition of Möbius function $\mu_{\mathcal{P}^N}$ yields

$$\sum_{Q\leqslant P: B\in Q} \mu_{\mathcal{P}^N}(Q, P) = 0.$$

Möbius inversion μ^{h_v} thus takes non-zero values only on modular elements, where it obtains recursively by $\mu^{h_v}(P_\perp) = \sum_{i\in N} v(\{i\})$, $\mu^{h_v}(P^\top) = \mu^v(N)$ as well as $\mu^{h_v}(\{A\} \cup P_\perp^{A^c}) = \mu^v(A)$ for $1 < |A| < n$. In other terms, any $w \neq v$ satisfying $\sum_{i\in N} v(\{i\}) = \sum_{i\in N} w(\{i\})$ and $\mu^v(A) = \mu^w(A)$ for all $A \in 2^N, |A| > 1$ also additively separates h, i.e. $h_v = h_w$. □

In view of corollary 8, the setting considered in this work deals precisely with additively separable partition functions, and thus the polynomial expression (2) in section 2 is not unique. Specifically, recall that the degree of a polynomial is the highest degree of its terms. Hence in (2), for any chosen set function w additively separating partiton function $h = h_w$, the degree is $\max\{|A| : \mu^w(A) \neq 0\}$, while every non-zero value of Möbius inversion $\mu^w : 2^N \to \mathbb{R}$ is a coefficient of the polynomial. The only degree k such that there is a unique set function available for polynomial expression (2) is $k = 0$, in which case the unique additively separating set function w is trivial: $w(A) = 0$ for all $A \in 2^N$. On the other hand, for any degree $k, 0 < k \leq n$ there exists a continuum of set functions available for additive separability and such that $\max\{|A| : \mu^w(A) \neq 0\} = k$, each defining alternative but equivalent coefficients of the polynomial.

6 NEAR-BOOLEAN GAMES

In view of the above definition of local maximizers relying on equilibrium conditions for strategic n-player games, and having mentioned additive separablity of partition functions or global games (Gilboa and Lehrer, 1990), it seems now useful to regard variables as players in near-Boolean coalition formation games (for pseudo-Boolean functions and coalitional games see (Hammer and Holzman, 1992, section 3)).

Definition 22. *A near-Boolean n-player game is a triple (N, F, π) such that $N = \{1, \ldots, n\}$ is the player set and F is a near-Boolean function taking real values on profiles $\boldsymbol{q} = (q_1, \ldots, q_n) \in \underset{i \in N}{\times} ex(\Delta_i)$ of strategies, while $\pi : \underset{i \in N}{\times} ex(\Delta_i) \to \mathbb{R}^n$ efficiently assigns payoffs $\pi(\boldsymbol{q}) = (\pi_1(\boldsymbol{q}), \ldots, \pi_n(\boldsymbol{q}))$ to players, i.e. $\sum_{i \in N} \pi_i(\boldsymbol{q}) = F(\boldsymbol{q})$ at all $\boldsymbol{q} \in \underset{i \in N}{\times} ex(\Delta_i)$.*

Definition 23. *A fuzzy near-Boolean n-player game is a triple $(N, \hat{F}, \pi)$ such that $N = \{1, \ldots, n\}$ is the player set and $\hat{F}$ is the MLE of a near-Boolean function taking values on profiles $\boldsymbol{q} = (q_1, \ldots, q_n) \in \underset{i \in N}{\times} \Delta_i$, while $\pi : \underset{i \in N}{\times} \Delta_i \to \mathbb{R}^n$ efficiently assigns payoffs to players, i.e. $\sum_{i \in N} \pi_i(\boldsymbol{q}) = \hat{F}(\boldsymbol{q})$ at all $\boldsymbol{q} \in \underset{i \in N}{\times} \Delta_i$.*

In both near-Boolean games and fuzzy ones the player set is finite. Given this, a main distinction is between games where players have either finite or else infinite sets of strategies, with near-Boolean games in the former class and fuzzy ones in the latter. In addition, players may play either deterministic (i.e. pure) or else random (i.e. mixed) strategies. In the latter case equilibrium conditions are stated in terms of expected payoffs, and by means of fixed point arguments for upper hemicontinuous correspondences such conditions are commonly fulfilled (Mas-Colell et al., 1995, p. 260). The sets of deterministic strategies in fuzzy near-Boolean games are precisely the sets of random strategies in near-Boolean games. Nevertheless, the payoffs for the fuzzy setting are *not*, in general, expectations.

The main framework where these games seem interesting is coalition formation, which combines both strategic and cooperative games. A generic strategy profile $\mathbf{q} \in \underset{i \in N}{\times} ex(\Delta_i)$ of near-Boolean (non-fuzzy) games may well fail to be exact (see proposition 4), but it seems plain that there is a unique partition P of N with associated $\mathbf{p} \in \underset{i \in N}{\times} ex(\Delta_i)$ obtained through shrinkings of $\mathbf{q}$ and such that $F(\mathbf{p}) = F(\mathbf{q})$. Let $\mathbf{p}(\mathbf{q})$ be such a unique $\mathbf{p}$. In view of the above discussion, it is also evident that for every $\mathbf{p}$ there are many (non-exact) $\mathbf{q}$ such that $\mathbf{p}=\mathbf{p}(\mathbf{q})$. In these terms, near-Boolean games model stategic coalition formation in a very handy manner, in that they totally bypass the need to define a mechanism mapping strategy profiles into partitions of players or coalition structures (Slikker, 2001). More precisely, a mechanism is a mapping $M : \underset{i \in N}{\times} ex(\Delta_i) \to \mathcal{P}^N$ such that when each player $i \in N$ specifies a coalition $A_i \in 2^N_i$, then $M(A_1, \ldots, A_n) = P$ is a resulting coalition structure. If the n specified coalitions $A_i, i \in N$ are such that for some partition P it holds $A_i = A$ for all $i \in A$ and all $A \in P$, then $M(A_1, \ldots, A_n) = P$. Otherwise, the generated partition $P' = M(A_1, \ldots, A_n)$ depends on what mechanism is chosen, and generally may be a rather fine one, i.e. possibly consisting of many small blocks. Conversely, near-Boolean games do not need any mechanism, in that even if players' strategies $(q_1, \ldots, q_n) = \mathbf{q}$ are such that $\mathbf{q}$ does not correspond to a partition, still the global worth $F(\mathbf{q})$ is that attained at the partition P with corresponding $\mathbf{p}(\mathbf{q})$, i.e. whose blocks $A \in P$ each include maximal subsets of players choosing the same superset $A' \supseteq A$.

Given coalitional game $v : 2^N \to \mathbb{R}_+, v(\emptyset) = 0$, with $F(\mathbf{p}) = \sum_{A \in P} v(A)$ for all partitions $P \leftrightarrow \mathbf{p}$, let payoffs be defined, for $i \in A$ and $A \in P$ and $\mathbf{p} = \mathbf{p}(\mathbf{q})$,

$$\text{by } \pi_i(\mathbf{q}) = \sum_{B \in 2^A \setminus 2^{A \setminus i}} \frac{\mu^w(B)}{|B|} \text{ for all } \mathbf{q} \in \underset{j \in N}{\times} ex(\Delta_j),$$

where P thus is the partition with associated $\mathbf{p}(\mathbf{q})$. Apart from the absence of any coalition structure generation mechanism (see above), this is in fact a well-known coalition formation game (Slikker, 2001), with payoffs given by the *Shapley value* (Roth, 1988).

Definition 24. *A local maximizer of near-Boolean function F is any $\boldsymbol{q} \in \underset{i \in N}{\times} ex(\Delta_i)$ such that for all $i \in N$ and all $q'_i \in ex(\Delta_i)$ inequality $F(\boldsymbol{q}) \geq F(q'_i|\boldsymbol{q}_{-i})$ holds.*

Remark 25. *If payoffs are given by $\pi_i(\boldsymbol{q}) = \frac{\omega_i F(\boldsymbol{q})}{\sum_{j \in N} \omega_j}$ for all $i \in N$, with $\omega_1, \ldots, \omega_n > 0$, then near-Boolean games are (pure) common interest potential games (Monderer and Shapley, 1996; Bowles, 2004). That is to say, the set of equilibria of (N, F, π) coincides* *with the set of local maximizers of F, and players' preferences all agree on the set of strategy profiles.*

In artificial intelligence, these games may model coalition formation in multiagent systems (Rahwan and Jenning, 2007; Conitzer and Sandholm, 2006).

7 CONCLUSIONS AND FUTURE WORK

Via polynomial MLE, near-Boolean functions of n variables take values on the n-product of high-dimensional unit simplices, thus enabling to approach discrete optimization problems, namely set packing/partitioning, through an objective function defined over a continuous domain, and with feasible solutions found at extreme points of the simplices. Least squares approximations with polynomials of bounded degree are discussed in terms of additive separability of partition functions, while near-Boolean n-player games flexibly model strategic coalition formation.

Apart from settings in artificial intelligence such as combinatorial auctions and coalition structure generation in multiagent systems discussed above, near-Boolean optimization seems generally interesting for objective function-based clustering (Rossi, 2015), and for graph clustering in particular (Schaeffer, 2007). Specifically, for a given weighted graph, edge weights and vertices' weighted degrees may be used (in alternative ways) to obtain a quadratic MLE in expression (2), i.e. a polynomial with degree 2. Then, optimization with respect to such an objective function provides a method for partitioning the vertex set. This deserves separate investigation in a forthcoming work.

Along an alternative route, through the core concept in cooperative game theory, it may be interesting to consider the following problem: for given generic (i.e. *non*-additively separable) global game or partition function $h \in \mathbb{R}^{\mathcal{B}_n}$, determine a quadratic MLE in (2), i.e. a set function w such that $\mu^w(A) = 0$ for all $A \in 2^N, |A| > 2$, satisfying $h_w(P) \geq h(P)$ for all partitions $P \in \mathcal{P}^N$ of players, where $h_w(P) = \sum_{A \in P} w(A)$, and with $h_w(P^\top) = h(P^\top)$ for the coarsest partition. The idea behind this problem comes from a novel approach to the solution of global games h (Gilboa and Lehrer, 1990), and shall be explored in future work.

REFERENCES

Aigner, M. (1997). *Combinatorial Theory*. Springer. Reprint of the 1979 Edition.

Alidaee, B., Kochenberger, G., Lewis, K., Lewis, M., and Wang, H. (2008). A new approach for modeling and solving set packing problems. *European Journal of Operational Research*, 186(2):504 – 512.

Boros, E. and Hammer, P. (2002). Pseudo-Boolean optimization. *Discrete App. Math.*, 123:155–225.

Bowles, S. (2004). *Microeconomics: Behavior, Institutions, and Evolution*. Princeton University Press.

Chandra, B. and Halldorsson, M. M. (2001). Greedy local improvement and weighted set packing. *Journal of Algorithms*, (39):223–240.

Conitzer, V. and Sandholm, T. (2006). Complexity of constructing solutions in the core based on synergies among coalitions. *Artificial Intel.*, 170:607–619.

Crama, Y. and Hammer, P. L. (2011). *Boolean Functions: Theory, Algorithms, and Applications*. Cambridge University Press.

Gilboa, I. and Lehrer, E. (1990). Global games. *International Journal of Game Theory*, (20):120–147.

Gilboa, I. and Lehrer, E. (1991). The value of information - an axiomatic approach. *Journal of Mathematical Economics*, 20(5):443–459.

Grünbaum, B. (2001). *Convex Polytopes 2nd ed.* Springer.

Hammer, P. and Holzman, R. (1992). Approximations of pseudo-Boolean functions; applications to game theory. *Math. Methods of Op. Res. - ZOR*, 36(1):3–21.

Hazan, E., Safra, S., and Schwartz, O. (2006). On the complexity of approximating k-set packing. *Computational Complexity*, (15):20–39.

Korte, B. and Vygen, J. (2002). *Combinatorial Optimization. Theory and Algorithms*. Springer.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press.

Milgrom, P. (2004). *Putting Auction Theory to Work*. Cambridge University Press.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14(1):124–143.

Papadimitriou, C. (1994). *Computational Complexity*. Addison Wesley.

Rahwan, T. and Jenning, N. (2007). An algorithm for distributing coalitional value calculations among cooperating agents. *Artificial Intelligence*, 171:535–567.

Rossi, G. (2015). Multilinear objective function-based clustering. In *Proc. 7th Int. J. Conf. on Computational Intelligence*, volume 2 (FCTA), pages 141–149.

Rota, G.-C. (1964a). The number of partitions of a set. *American Mathematical Monthly*, 71:499–504.

Rota, G.-C. (1964b). On the foundations of combinatorial theory I: theory of Möbius functions. *Z. Wahrscheinlichkeitsrechnung u. verw. Geb.*, 2:340–368.

Roth, A. (1988). *The Shapley value*. Cambridge Univ. Press.

Sandholm, T. (2002). Algorithm for optimal winner determination in combinatorial auctions. *Artificial Intelligence*, (135):1–54.

Schaeffer, S. E. (2007). Graph clustering. *Computer Science Review*, 1(1):27–64.

Slikker, M. (2001). Coalition formation and potential games. *Games and Econ. Behavior*, 37(2):436 – 448.

Stanley, R. (1971). Modular elements of geometric lattices. *Algebra Universalis*, (1):214–217.

Stern, M. (1999). *Semimodular Lattices. Theory and Applications.* Cambridge University Press.

Trevisan, L. (2001). Non-approximability results for optimization problems on bounded degree instances. In *Proc. 33rd ACM Symp. on Theory of Computing*, pages 453–461.