

Exact bounds of the Möbius inverse of monotone set functions

Michel GRABISCH* Pedro MIRANDA†

March 31, 2015

Abstract

We give the exact upper and lower bounds of the Möbius inverse of monotone and normalized set functions (a.k.a. normalized capacities) on a finite set of n elements. We find that the absolute value of the bounds tend to $\frac{4^{n/2}}{\sqrt{\pi n/2}}$ when n is large. We establish also the exact bounds of the interaction transform and Banzhaf interaction transform, as well as the exact bounds of the Möbius inverse for the subfamilies of k -additive normalized capacities and p -symmetric normalized capacities.

Keywords: Möbius inverse, monotone set function, interaction

AMS Classification: 05, 06, 91

1 Introduction

The Möbius function is a well-known tool in combinatorics and partially ordered sets (see, e.g., [?, ?, ?]). In the field of decision theory, the Möbius inverse of a monotone set function (called a capacity) is a fundamental concept permitting to derive simple expressions of nonadditive integrals and to analyze the core of capacities (set of probability measures dominating a capacity) [?]. Set functions can also be seen as pseudo-Boolean functions, and it is well known that the Möbius inverse corresponds to the coefficients of the polynomial representation of a pseudo-Boolean function. In particular, monotone and normalized pseudo-Boolean functions correspond to semicoherent structure functions in reliability theory (see, e.g., Marichal and Mathonet [?], Marichal [?]).

Consider $N = \{1, \dots, n\}$ and a monotone set function $\mu : 2^N \rightarrow [0, 1]$ with the property $\mu(\emptyset) = 0$ and $\mu(N) = 1$ (normalized capacity). In optimization problems involving capacities or monotone pseudo-Boolean functions (as in reliability) it is often useful to know the bounds of the Möbius inverse to use algorithmic methods (see Crama and Hammer [?], Chapter 13). This is the case for example when dealing with k -additive

*Corresponding author. Paris School of Economics, University of Paris I, 106-112, Bd. de l'Hôpital, 75013 Paris, France. Tel. (33) 144-07-82-85, Fax (33)-144-07-83-01. Email: michel.grabisch@univ-paris1.fr

†Complutense University of Madrid. Plaza de Ciencias, 3, 28040 Madrid, Spain Email: pmiranda@mat.ucm.es

measures, which are best represented through their Möbius inverse (see below); then, when solving optimization problems like model fitting, algorithms usually need to fix an interval where the searched values lay, and the upper and lower bounds are the natural limits of these intervals. Surprisingly, although μ takes values in $[0, 1]$, the exact bounds of its Möbius inverse grow rapidly with n , approximately in $\frac{4^{n/2}}{\sqrt{\pi n/2}}$ when n is large.

The aim of the paper is to establish this result, correcting wrong bounds obtained in a previous paper by the authors [?], and providing a complete proof of the result. We extend this result to the interaction transform, another useful linear invertible transform of set functions, and we consider also specific subclasses of capacities, like k -additive and p -symmetric capacities.

2 Preliminaries

Let $N = \{1, \dots, n\}$. A *capacity* on N is a set function $\mu : 2^N \rightarrow \mathbb{R}$ satisfying $\mu(\emptyset) = 0$ and monotonicity: $A \subseteq B \subseteq N$ implies $\mu(A) \leq \mu(B)$. A capacity is *normalized* if in addition $\mu(N) = 1$. We denote respectively by $\mathcal{C}(N)$ and $\mathcal{NC}(N)$ the set of capacities and normalized capacities on N . The set $\mathcal{NC}(N)$ is a convex closed polytope, whose extreme points are all $\{0, 1\}$ -valued normalized capacities (as the polytope of normalized capacities is an order polytope, this result has been shown by Stanley [?]. For a direct proof, see [?]). We denote by $\mathcal{NC}_{0,1}(N)$ the set of all $\{0, 1\}$ -valued normalized capacities.

Consider a set function ξ on N such that $\xi(\emptyset) = 0$. The *monotonic cover* of ξ is the smallest capacity μ such that $\mu \geq \xi$. We denote it by $\widehat{\xi}$, and it is given by

$$\widehat{\xi}(A) = \max_{B \subseteq A} \xi(B) \quad (A \subseteq N). \quad (1)$$

Consider now a set function $\xi : 2^N \rightarrow \mathbb{R}$. The linear system

$$\xi(A) = \sum_{B \subseteq A} m(B) \quad (A \in 2^N) \quad (2)$$

has always a unique solution, known as the *Möbius inverse* [?], and is given by

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \xi(B) \quad (A \in 2^N). \quad (3)$$

Since m is also a set function, we view now the Möbius inverse as a transform on the set of set functions:

$$m : \mathbb{R}^{(2^N)} \rightarrow \mathbb{R}^{(2^N)}; \xi \mapsto m^\xi \text{ given by (3).}$$

We call m the *Möbius transform* of ξ . Remark that it is a linear invertible transform.

We introduce another linear invertible transform, which is useful in decision making, called the *interaction transform*. To this end we introduce the derivative of a set function ξ . Let $i \in N$ and $A \subseteq N \setminus \{i\}$. The *derivative* of ξ w.r.t. i at A is defined by $\Delta_i \xi(A) = \xi(A \cup \{i\}) - \xi(A)$. Derivatives w.r.t. sets are defined recursively by

$$\Delta_K \xi(A) = \Delta_{K \setminus \{i\}}(\Delta_{\{i\}} \xi(A)) \quad (|K| \geq 1)$$

with $i \in K$, $\Delta_{\{i\}}\xi = \Delta_i\xi$, and $\Delta_{\emptyset}\xi = \xi$. For $A \subseteq N \setminus K$, we obtain

$$\Delta_K\xi(A) = \sum_{L \subseteq K} (-1)^{|K \setminus L|} \xi(A \cup L).$$

Also, observe that

$$m^\xi(A) = \Delta_A\xi(\emptyset) \quad (A \in 2^N). \quad (4)$$

The interaction transform $I : \mathbb{R}^{(2^N)} \rightarrow \mathbb{R}^{(2^N)}$ computes a weighted average of the derivatives:

$$I^\xi(A) = \sum_{B \subseteq N \setminus A} \frac{(n - b - a)!b!}{(n - a + 1)!} \Delta_A\xi(B) \quad (A \in 2^N), \quad (5)$$

where $a = |A|$, $b = |B|$. Its expression through the Möbius transform is much simpler:

$$I^\xi(A) = \sum_{B \supseteq A} \frac{1}{b - a + 1} m^\xi(B), \quad (6)$$

while the inverse relation uses the Bernoulli numbers B_k :

$$m^\xi(A) = \sum_{B \supseteq A} B_{a-b} I^\xi(B).$$

(see [?, ?] for details). Another related transform is the *Banzhaf interaction transform* I_B , which is the (unweighted) average of the derivatives:

$$I_B^\xi(A) = \frac{1}{2^{n-a}} \sum_{B \subseteq N \setminus A} \Delta_A\xi(B) \quad (A \in 2^N). \quad (7)$$

Lastly, we introduce two specific families of normalized capacities. A normalized capacity μ is said to be *at most k -additive* ($1 \leq k \leq n$) if $m^\mu(A) = 0$ for every set $A \in 2^N$ such that $|A| > k$ [?]. 1-additive capacities are ordinary additive capacities, i.e., satisfying $\mu(A \cup B) = \mu(A) + \mu(B)$ for disjoint sets A, B . Note that by (6), m^μ can be replaced by I^μ in the above definition.

We denote by $\mathcal{NC}^{\leq k}(N)$ the set of at most k -additive capacities on N . It is a convex closed polytope (see [?] for a study of its properties).

Another family of interest is the family of p -symmetric capacities [?]. A capacity μ is *symmetric* if $\mu(A) = \mu(B)$ whenever $|A| = |B|$. We denote by $\mathcal{SNC}(N)$ the set of symmetric normalized capacities. This notion can be generalized as follows. A nonempty subset $A \subseteq N$ is a *subset of indifference* for μ if for all $B_1, B_2 \subset A$ with $|B_1| = |B_2|$, we have $\mu(C \cup B_1) = \mu(C \cup B_2)$ for every $C \subseteq N \setminus A$. The *basis* of the capacity is the coarsest partition of N into subsets of indifference. It always exists and is unique [?]. Now, μ is *p -symmetric* with respect to the partition $\{A_1, \dots, A_p\}$ if this partition is its basis. Symmetric games are therefore 1-symmetric games (with respect to the basis $\{N\}$). We denote by $\mathcal{SNC}^{\leq p}(A_1, \dots, A_p)$ the set of normalized capacities such that A_1, \dots, A_p are subsets of indifference. It is a convex closed polytope (again, see [?] for a study of its properties).

Lastly, we mention a combinatorial result on the binomial coefficients:

$$\sum_{\ell=0}^k (-1)^\ell \binom{n}{\ell} = (-1)^k \binom{n-1}{k} \quad (k < n), \quad (8)$$

for any positive integer n .

3 Exact bounds of the Möbius inverse

We present in this section the main result of the paper.

Theorem 1. For any normalized capacity μ , its Möbius transform satisfies for any $A \subseteq N$, $|A| > 1$:

$$-\binom{|A|-1}{l'_{|A|}} \leq m^\mu(A) \leq \binom{|A|-1}{l_{|A|}},$$

with

$$l_{|A|} = 2 \left\lfloor \frac{|A|}{4} \right\rfloor, \quad l'_{|A|} = 2 \left\lfloor \frac{|A|-1}{4} \right\rfloor + 1 \quad (9)$$

and for $|A| = 1 < n$:

$$0 \leq m^\mu(A) \leq 1,$$

and $m^\mu(A) = 1$ if $|A| = n = 1$. These upper and lower bounds are attained by the normalized capacities μ_A^*, μ_{A^*} , respectively:

$$\mu_A^*(B) = \begin{cases} 1, & \text{if } |A| - l_{|A|} \leq |B \cap A| \leq |A| \\ 0, & \text{otherwise} \end{cases}, \quad \mu_{A^*}(B) = \begin{cases} 1, & \text{if } |A| - l'_{|A|} \leq |B \cap A| \leq |A| \\ 0, & \text{otherwise} \end{cases}$$

for any $B \subseteq N$.

We give in Table 1 the first values of the bounds. Using the well-known Stirling's

$ A $	1	2	3	4	5	6	7	8	9	10	11	12
u.b. of $m^\mu(A)$	1	1	1	3	6	10	15	35	70	126	210	462
l.b. of $m^\mu(A)$	1(0)	-1	-2	-3	-4	-10	-20	-35	-56	-126	-252	-462

Table 1: Lower and upper bounds for the Möbius transform of a normalized capacity

approximation $\binom{2n}{n} \simeq \frac{4^n}{\sqrt{\pi n}}$ for $n \rightarrow \infty$, we deduce that

$$-\frac{4^{\frac{n}{2}}}{\sqrt{\frac{\pi n}{2}}} \leq m^\mu(N) \leq \frac{4^{\frac{n}{2}}}{\sqrt{\frac{\pi n}{2}}}$$

when n tends to infinity.

Proof. Let us prove the result for the upper bound when $A = N$. We consider the group S_n of permutations on N . For any $\sigma \in S_n$ and any capacity $\mu \in \mathcal{NC}(N)$, we define the capacity $\sigma(\mu) \in \mathcal{NC}(N)$ by $\sigma(\mu)(B) = \mu(\sigma^{-1}(B))$ for any $B \subseteq N$.

We observe that the target function $m^\mu(N)$ is invariant under permutation. Indeed,

$$\begin{aligned} m^{\sigma(\mu)}(N) &= \sum_{B \subseteq N} (-1)^{n-|B|} \mu(\sigma^{-1}(B)) \\ &= \sum_{B' \subseteq N} (-1)^{n-|B'|} \mu(B') \quad (\text{letting } B' = \sigma^{-1}(B)) \\ &= m^\mu(N). \end{aligned}$$

For every set function μ on N , define its symmetric part $\mu^s = \frac{1}{n!} \sum_{\sigma \in S_n} \sigma(\mu)$, which is a symmetric function. By convexity of $\mathcal{NC}(N)$, if $\mu \in \mathcal{NC}(N)$, then so is μ^s , and by linearity of the Möbius inverse, we have

$$m^{\mu^s}(N) = \frac{1}{n!} \sum_{\sigma \in S_n} m^{\sigma(\mu)}(N) = \frac{1}{n!} \sum_{\sigma \in S_n} m^\mu(N) = m^\mu(N).$$

It is therefore sufficient to maximize $m^\mu(N)$ on the set of symmetric normalized capacities $\mathcal{SNC}(N)$. But this set is also a convex polytope, whose extreme points are the following $\{0, 1\}$ -valued capacities μ_k defined by

$$\mu_k(B) = 1 \text{ iff } |B| \geq n - k \quad (k = 0, \dots, n - 1).$$

Indeed, if μ is symmetric, it can be written as a convex combination of these capacities:

$$\mu = \mu(\{1\})\mu_{n-1} + \sum_{k=2}^n (\mu(\{1, \dots, k\}) - \mu(\{1, \dots, k-1\}))\mu_{n-k}$$

It follows that the maximum of $m^\mu(N)$ is attained on one of these capacities, say μ_k . We compute

$$\begin{aligned} m^{\mu_k}(N) &= \sum_{B \subseteq N} (-1)^{|N \setminus B|} \mu(B) = \sum_{i=n-k}^n (-1)^{n-i} \binom{n}{i} \\ &= \sum_{i'=0}^k (-1)^{i'} \binom{n}{n-i'} = (-1)^k \binom{n-1}{k}, \end{aligned} \quad (10)$$

where the third equality is obtained by letting $i' = n - i$ and the last one follows from (8). Therefore k must be even. If $n - 1$ is even, the maximum of $\binom{n-1}{k}$ for k even is attained for $k = \frac{n-1}{2}$ if this is an even number, otherwise $k = \frac{n-3}{2}$. If $n - 1$ is odd, the maximum of $\binom{n-1}{k}$ is reached for $k = \lceil \frac{n-1}{2} \rceil$ and $k - 1 = \lfloor \frac{n-1}{2} \rfloor$, among which the even one must be chosen. As it can be checked (see Table 2 below), this amounts to taking

$$k = 2 \left\lfloor \frac{n}{4} \right\rfloor$$

that is, $k = l_n$ as defined in (9), and we have defined the capacity

$$\mu^*(B) = 1 \text{ if } n - l_n \leq |B| \leq n,$$

which is μ_N^* as defined in the theorem.

For establishing the upper bound of $m^\mu(A)$ for any $A \subset N$, remark that the value of $m^\mu(A)$ depends only on the subsets of A . It follows that applying the above result to the sublattice 2^A , the set function ξ_A^* defined on 2^N by

$$\xi_A^*(B) = 1 \text{ if } B \subseteq A \text{ and } |A| - l_{|A|} \leq |B| \leq |A|, \text{ and } 0 \text{ otherwise,}$$

yields an optimal value for $m^\mu(A)$. It remains to turn this set function into a capacity on N , without destroying optimality. This can be done since ξ_A^* is monotone on 2^A , so that taking the monotonic cover of ξ_A^* by (1) yields an optimal capacity, given by

$$\widehat{\xi}_A^*(B) = \max_{C \subseteq B} \xi_A^*(C) = 1 \text{ if } |A| - l_{|A|} \leq |B \cap A| \leq |A|, \text{ and } 0 \text{ otherwise,}$$

which is exactly μ_A^* as desired. Note however that this is not the only optimal solution in general, since values of the capacity on the sublattice $2^{N \setminus A}$ are irrelevant.

One can proceed in a similar way for the lower bound. In this case however, as it can be checked, the capacity must be equal to 1 on the $l'_n + 1$ first lines of the lattice 2^N , with $l'_n = 2 \lfloor \frac{n-1}{4} \rfloor + 1$ (see Table 2). \square

n/k	0	1	2	3	4	5	6	7	8	9	10	11
$n = 1$	1											
$n = 2$	1	-1										
$n = 3$	1	-2	1									
$n = 4$	1	-3	3	-1								
$n = 5$	1	-4	6	-4	1							
$n = 6$	1	-5	10	-10	5	-1						
$n = 7$	1	-6	15	-20	15	-6	1					
$n = 8$	1	-7	21	-35	35	-21	7	-1				
$n = 9$	1	-8	28	-56	70	-56	28	-8	1			
$n = 10$	1	-9	36	-84	126	-126	84	-36	9	-1		
$n = 11$	1	-10	45	-120	210	-252	210	-120	45	-10	1	
$n = 12$	1	-11	55	-105	330	-462	462	-330	165	-55	11	-1

Table 2: Computation of the upper (red) and lower (blue) bounds. The value of the capacity μ is 1 for the $k + 1$ first lines of the lattice 2^N . Each entry (n, k) equals $m^\mu(N)$, as given by (10).

4 Exact bounds of the interaction transforms

We begin by establishing a technical lemma which will permit to get the results easily from Theorem 1.

Lemma 1. Let $A, B \subset N$, $A \neq \emptyset$, be disjoint sets. Then

$$\max_{\mu \in \mathcal{NC}(N)} \sum_{C \subseteq A} (-1)^{a-c} \mu(B \cup C) = \max_{\mu \in \mathcal{NC}(N)} m^\mu(A), \quad (11)$$

and the maximum is attained for $\mu = \mu_A^*$.

Proof. The function we have to maximize is simply the derivative $\Delta_A \mu(B)$. As this is a linear function in μ and $\mathcal{NC}(N)$ is a polytope, its maximum is attained on a vertex, i.e. a $\{0, 1\}$ -valued capacity. If $\mu(B \cup A) = \mu(B)$, then by monotonicity of μ we get $\Delta_A \mu(B) = 0$. Since this is clearly not the maximum of the derivative, we can discard such capacities μ from the analysis. Assuming then $\mu(B \cup A) > \mu(B)$, we define a capacity $\mu_B \in \mathcal{C}(A)$ by

$$\mu_B(C) = \mu(B \cup C) - \mu(B) \quad (C \subseteq A). \quad (12)$$

Observe that if μ is $\{0, 1\}$ -valued, then necessarily $\mu(B \cup A) = 1$ and $\mu(B) = 0$, hence (12) collapses to $\mu_B(C) = \mu(B \cup C)$, for any $C \subseteq A$, and μ_B is $\{0, 1\}$ -valued and normalized

too. Moreover, any $\{0, 1\}$ -valued normalized capacity on A can be obtained from a $\{0, 1\}$ -valued normalized capacity on N by the latter equality. On the other hand, remark that for any $\mu \in \mathcal{NC}(N)$

$$m^{\mu_B}(A) = \sum_{C \subseteq A} (-1)^{a-c} \mu_B(C) = \sum_{C \subseteq A} (-1)^{a-c} \mu(B \cup C)$$

since $\sum_{C \subseteq A} (-1)^{a-c} = 0$. In summary, we have

$$\max_{\mu \in \mathcal{NC}(N)} \Delta_A \mu(B) = \max_{\mu \in \mathcal{NC}_{0,1}(N)} \Delta_A \mu(B) = \max_{\mu \in \mathcal{NC}_{0,1}(A)} m^\mu(A) = \max_{\mu \in \mathcal{NC}(A)} m^\mu(A) = \max_{\mu \in \mathcal{NC}(N)} m^\mu(A),$$

the last equality coming from Theorem 1. Hence (11) is established, the value of the maximum is given by Theorem 1, as well as the capacity attaining the maximum. \square

A similar result can be established for the lower bound.

Corollary 1. Consider $A \subseteq N$. The upper and lower bounds for the interaction transform $I(A)$ are the same as for $m(A)$, and they are obtained for the capacities μ_A^* and μ_{A^*} of Theorem 1.

Proof. We will obtain the upper bound, the proof for the lower bound being similar. From Lemma 1, we see that the maximum of $\Delta_A \mu(B)$ does not depend on B . Thus, from (5), letting $m^*(A) = \max_{\mu \in \mathcal{NC}(N)} m^\mu(A)$ we obtain

$$\max_{\mu \in \mathcal{NC}(N)} I^\mu(A) = \sum_{B \subseteq N \setminus A} \frac{(n-a-b)!b!}{(n-a+1)!} m^*(A) = m^*(A) \sum_{b=0}^{n-a} \frac{(n-a-b)!b!}{(n-a+1)!} \binom{n-a}{b} = m^*(A).$$

\square

Similarly, we obtain the exact bounds for the Banzhaf interaction index.

Corollary 2. Consider $A \subseteq N$. The upper and lower bounds for $I_B(A)$ are the same as for $m(A)$. These upper and lower bounds are obtained for the capacities μ_A^* and μ_{A^*} of Theorem 1.

Proof. Proceeding as for Corollary 1, the result follows from the identity $\sum_{b=0}^{n-a} \binom{n-a}{b} = 2^{n-a}$. \square

5 Exact bounds for k -additive and p -symmetric capacities

We show in this section that the results established for the bounds of the Möbius and interaction transforms on the set of normalized capacities are still valid when one restricts to k -additive capacities and p -symmetric capacities.

Proposition 1. For any nonempty $A \subseteq N$, the normalized capacities μ_A^*, μ_{A^*} given in Theorem 1 are at most k -additive for any $|A| \leq k \leq n$. Therefore, the upper and lower bounds for the Möbius transform, the interaction transform and the Banzhaf interaction transform, are valid:

$$\max_{\mu \in \mathcal{NC}(N)} m^\mu(A) = \max_{\mu \in \mathcal{NC}^{\leq k}(N)} m^\mu(A), \quad \min_{\mu \in \mathcal{NC}(N)} m^\mu(A) = \min_{\mu \in \mathcal{NC}^{\leq k}(N)} m^\mu(A),$$

for $|A| \leq k \leq n, \emptyset \neq A \subseteq N$, and similarly for $I^\mu(A), I_B^\mu(A)$.

Proof. Given a nonempty $A \subseteq N$, it suffices to show that μ_A^*, μ_{A^*} are at most k -additive for $k = |A|$. Take $B \subseteq N$ such that $k < |B| \leq n$. Then, $B \setminus A \neq \emptyset$. On the other hand, observe that for any $i \notin A$,

$$\Delta_i \mu_A^*(K) = \mu_A^*(K \cup i) - \mu_A^*(K) = 0$$

for any $K \not\ni i$. It follows that $\Delta_B \mu_A^*(K) = 0$ for any K as soon as $B \setminus A \neq \emptyset$. Taking $K = \emptyset$, by (4), we conclude that $m^{\mu_A^*}(B) = 0$ if $k < |B| \leq n$, as desired. \square

Remark 1. Proposition 1 tells us what is the maximum achieved by $m^\mu(A)$ for the set of k -additive capacities when $|A| \leq k \leq n$, but says nothing when $k < |A|$. The question appears to be very complex, because in general μ_A^* will not be k -additive, and the vertices of the polytope of k -additive capacities are not known, except for $k = 1$ and 2. In particular, it is known that many vertices are *not* $\{0, 1\}$ -valued as soon as $k > 2$ (see [?]).

Proposition 2. For any $1 \leq p \leq n$ and any partition $\{A_1, \dots, A_p\}$ of N ,

$$\begin{aligned} \max_{\mu \in \mathcal{NC}(N)} m^\mu(A) &= \max_{\mu \in \mathcal{SNC}^{\leq p}(A_1, \dots, A_p)} m^\mu(A), & (\emptyset \neq A \subseteq N), \\ \min_{\mu \in \mathcal{NC}(N)} m^\mu(A) &= \min_{\mu \in \mathcal{SNC}^{\leq p}(A_1, \dots, A_p)} m^\mu(A), & (\emptyset \neq A \subseteq N), \end{aligned}$$

and similarly for $I^\mu(A), I_B^\mu(A)$.

Proof. Consider the capacities defined by

$$\mu_A^{**}(B) := \begin{cases} 1 & \text{if } |B| \geq l_{|A|} + 1 \\ 0 & \text{otherwise} \end{cases}, \quad \mu_{A^{**}}(B) := \begin{cases} 1 & \text{if } |B| \geq l_{|A|} \\ 0 & \text{otherwise} \end{cases}$$

Observe that $\mu_A^{**}(C) = \mu_A^*(C)$, $\mu_{A^{**}}(C) = \mu_{A^*}(C)$ for any $C \subseteq A$.

Therefore $m^{\mu_A^{**}}(A) = m^{\mu_A^*}(A)$, $m^{\mu_{A^{**}}}(A) = m^{\mu_{A^*}}(A)$. On the other hand, μ_A^{**} and $\mu_{A^{**}}$ are symmetric capacities, whence they are p -symmetric for any p and any partition of indifference. \square

6 Acknowledgements

We address all our thanks to the anonymous referees, for their constructive comments, which have permitted to correct and improve the paper, in particular the proof of Theorem 1 which is now much shorter and more clear. This work was partially supported by Grant MTM2012-33740.

References