Theory and Decision
https://doi.org/10.1007/s11238-025-10047-x

A characterization of delay averse Choquet integrals for intertemporal analysis

José Carlos R. Alcantud[1]

Accepted: 26 April 2025

Abstract
In an intertemporal framework with a finite horizon, we pose the problem of characterizing which discrete Choquet integrals satisfy strong aversion to delayed rewards. This property is arguably a minimal condition for a reasonable evaluation of vectors of rewards received at successive periods of time. We relate this property to other distributional axioms discussed for this framework. Particularly, we show that for monotonic evaluations, delay-aversion is equivalent to consistency with cumulative domination. Then, we define a property (temporal anti-buoyancy) and we prove that it is necessary and sufficient for a capacity to define a strongly delay-averse Choquet integral.

Keywords Intertemporal analysis · Delay aversion · Choquet integral · Capacity

1 Introduction

The general aim of this paper is to explore the ability of the Choquet integral to capture intertemporal assessments with appropriate axiomatic properties. Prior to stating our specific goals, let us review these two main components of our article.

1.1 Intertemporal analysis

The intertemporal analysis has its own characteristic features. One reason is that anomalies, also called paradoxes, abound in intertemporal problems. Their origins are not always obvious, for example, Van der Wal et al. (2013) found evidence that nature exposure reduces future discounting. These authors reported on other environmental effects of temporal discounting. Below we recall that the intertemporal analysis has implications in many other different fields.

✉ José Carlos R. Alcantud
jcr@usal.es

1 BORDA Research Group and IME, University of Salamanca, 37007 Salamanca, Spain

Published online: 23 May 2025

One specific topic that sparked our interest is the mounting evidence for time discounting, e.g., in financial choices, albeit it may be subject to the aforementioned anomalies. It is easy to test that people frequently prefer immediate profits at the expense of future gains. This paradoxical behavior is not exclusive of financial decisions. For example, Sparkman et al. (2021) conducted experiments to disentangle the relationship between psychological distance and environmental policy support. Recently, Ruggeri et al. (2022) conducted a test across sixty-one countries to explore the reasons why temporal discounting feeds into rising global inequality. They considered five anomalies of intertemporal choices. In life cycle assessment, Yuan et al. (2015) reviewed theoretical frameworks for temporal discounting. It is no surprise that to determine the present value of future rewards (such as cash flows in mathematical finance), time discounting provides an influential methodology. In other areas, the nature of the temporal discount must be well known too, e.g., for the purposes of public health and environmental policy-making (Ruggeri et al., 2022).

The question that stems is the computation of the "present value" of forthcoming cash flows. Although the "discount rate" is at the root of several solutions, the exact form of the evaluation is subject to intense debate, both from experimental and theoretical viewpoints. Exponential discounting defines the standard Discounted Utility model introduced by authors such as Fisher (1930) and Samuelson (1937). After experimental evidence proved that this model is unable to explain the time-inconsistent patterns observed in actual intertemporal choices (Loewenstein & Thaler, 1989; Thaler, 1981; Willis-Moore et al., 2024), more faithful accounts of people's evaluations of future endowments have been proposed, notably the hyperbolic discounting model (Green et al., 1997; Kirby, 1997). All these formulas and others proposed in the literature (e.g., quasi-hyperbolic discounting) boil down to weighted sums. But it may be that additivity lies at the heart of why of these models for time discounting are inadequate, a fact that has been long argued since Read (2001). This author asserted that time discounting must only be subadditive.

Aggregation operators offer a natural escape to this stalemate. Many aggregation operators alleviate the additivity property, and one remarkable example is the Choquet integral defined by Choquet (1954), who in particular, identified which Choquet integrals are subadditive. At any rate, in a (possibly discrete and finite) temporal modelization, basic constraints should be forcefully observed. In this regard, the indisputable lesson we draw from the complex scenario described above is that individuals consistently display aversion to delayed rewards. We will therefore be concerned with Choquet integrals that satisfy this property.

1.2 Choquet integrals

Although we explained that the Choquet integral was developed in 1954 (Choquet, 1954), Struk (2006) states that it was earlier proposed by Vitali (1925) who also anticipated the set functions later known as fuzzy measures (Sugeno, 1974). Other authors had referred to set functions that are monotone and vanish on the empty set under various names (Mesiar & Mesiarová, 2008). A celebrated characterization of the Choquet integral was given by Schmeidler (1986), who pioneered its utilization in decision making under uncertainty. Grabisch et al. (2009) and Grabisch and Labreuche (2010) established the foundation of its applicability in multi-attribute decisions. Nowadays its applications have extended to a large number of specialities (Alcantud & de Andrés Calle, 2017; Beliakov & James, 2022; Even & Lehrer, 2014; Li et al., 2013; Ok & Zhou, 2000; Torra & Narukawa, 2008, 2016). Concerning computability, Ontkovičová and Torra (2024) gives an updated overview and contribution to the calculation of Choquet integrals of arbitrary continuous functions. In this article the problem that we investigate only requires the utilization of the discrete version of the Choquet integral. Characterizations of the discrete Choquet integral are available as well (Couceiro & Marichal, 2011, 2013; Marichal, 2000). The literature attests to the existence of many generalizations of the Choquet integral (Dimuro et al., 2020; Horanská & Siposová, 2018).

A fundamental component of the construction of the Choquet integral is the significant class of non-additive set functions known as capacities, or fuzzy measures. Capacities form the basis of a generalization of measure theory with important applications in decision making (Grabisch, 2016). Wakker (1990) gave a behavioral foundation for capacities comparable to that of Savage (1954) for subjective probabilities. While Savage's target was the optimization of expected utility, Wakker opened the door to verification/falsification of capacities through observations. Grabisch (1995) inaugurated the problem of identification of capacities, a pivotal aspect in the aggregation of interacting inputs using the Choquet integral and its extensions. However this task is particularly demanding due to the high complexity of the polytope of capacities. Motivated by this problem, Grabisch (1997) introduced the important class of k-additive capacities. Their computational complexity is considerably smaller than in the general case, but k-additive capacities allow for sufficient flexibility even for small values of k. Another important topic approached for example by Beliakov and Divakov (2020) is the computational aspects of learning capacities, a subject reviewed by (Beliakov et al. 2020, Chapter 8) from the perspective of data fitting. Currently, a Python module is available that uses empirical data to learn capacities (Türkarslan & Torra, 2022). Yet another area of inspection is motivated by the fact that simulations and probabilistic optimization algorithms often require the random generation of capacities with specific properties. Studies that investigate this problem include Beliakov (2022), who focused on the case where it is required that the capacities are supermodular, resp., submodular.

Many contributions provide evidence of interaction with other topics. Narukawa et al. (2023) used the Choquet integral, as well as other fuzzy integrals, for building a type of probabilistic metric spaces that are based on capacities. In a multi-period context with uncertainty, Gilboa (1989) proves a representation theorem and an extension to an infinite horizon. As a by-product, he also proved that the Choquet integral is continuous with respect to monotonic convergence. Beliakov and James (2021) approached optimisation in the case that the objective functions are Choquet integrals. A simplified version optimizes problems with a single constraint based on OWA (Ordered Weighted Averaging) evaluations. OWAs operators (Yager, 1988) are discrete Choquet integrals defined from symmetric capacities. Aristondo et al. (2013) showed that OWA operators defined from antibuoyant weighting vectors perform well as welfare functions. These authors presented semantic justifications of this property based on links with the Pigou–Dalton progressive transfers principle, a sound principle from welfare economics. Beliakov and James (2022) used an extension of the antibuoyancy property to capacities in order to identify the Choquet integrals that satisfy the Pigou–Dalton principle. Alcantud and de Andrés Calle (2017) used the Choquet integral in the context of solving impossibilities in the group identification problem with the help of fuzzy interpretations. And Martino et al. (2024) considered aggregation by a discrete Choquet integral in their inspection of the role of (components of) financial literacy and gender in anomalies in intertemporal choice.

Finally, as an antecedent to the present study, we mention that the use of Sugeno's fuzzy measures led Alcantud (2024) to define subadditive discrete Choquet integrals that evaluate each time moment exactly like the additive hyperbolic (or exponential) discounting expressions do. That exercise aligns fuzzy integrals with Read (2001)'s requirement that time discounting should be subadditive rather than explicitly additive and hyperbolic.

1.3 Goals and outline of this article

It has been established that the comparison of temporal gains is a controversial task. Many axiomatic properties attempt to give an organized perspective. Evaluation functions (of the vectors) must be tested on these properties. We first aim at giving an overview of characteristics that are important for discussing reasonable evaluations of temporal rewards. Then we focus on delay aversion, which can arguably be a necessary condition for acceptable evaluations of the vectors in temporal problems. Indeed, the most popular formulas for intertemporal evaluations (i.e., those based on exponential, hyperbolic, and quasi-hyperbolic discounting) satisfy the axiomatization of delay aversion that we consider in this article. We show that this axiom is equivalent to another natural property, namely, the respect for cumulative domination (or c-monotonicity, for short), in the case of monotonic evaluations (a restriction that does not discard any rational evaluation, specifically, the formulas mentioned above). The same is true for respective strong versions of these axioms. In particular, monotonicity is a property of the Choquet integrals. The next question emerges: what Choquet integrals for intertemporal evaluations satisfy these equivalent properties? We identify an axiom that we call temporal-antibuoyancy, which gives a full solution to this problem. In passing, we argue that Choquet integrals that satisfy the Pigou–Dalton principle are basically incompatible with the intertemporal analysis. We also establish that when a capacity is buoyant, the Choquet integral that it defines can be neither strongly delay-averse nor strongly c-monotonic. And we prove results that confirm that c-monotonicity and symmetry are rarely combined in a numerical evaluation. An important conclusion is that among all OWA operators, only the simple arithmetic mean is c-monotonic.

The outline of this work is as follows. Section 2 gives preliminaries concerning both the 'fair' comparison/evaluation of vectors representing temporal rewards, and the discrete Choquet integral, inclusive of special cases that deserve our attention (OWAs and weighted arithmetic means). Section 3 explores how narrowly we can implement

c-monotonicity in the presence of symmetry. Section 4 produces our main results. First we prove the impossibility of a delay-averse function that satisfies the Pigou–Dalton progressive principle. Then we derive consequences of the strong version of delay-aversion of the Choquet integral on the capacity that produces it. Specifically, Choquet integrals computed from buoyant capacities cannot be strongly delay-averse. We give a simple proof that submodular capacities that satisfy a simple additional condition produce strongly delay-averse Choquet integrals. Finally, we axiomatically characterize strongly delay-averse Choquet integrals in terms of temporal-antibuoyant capacities. Section 5 gives final remarks, plus lines for future inspection.

The proofs of various lemmas and main results are postponed to Appendices. Specifically, Appendix A, Appendix B and Appendix C respectively include material from Sects. 2, 3 and 4.

2 Preliminaries

Henceforth, X^k denotes the Cartesian product of k copies of a non-empty set X, and $\mathscr{P}(X)$ denotes the set of all subsets of X; whereas $\mathscr{P}^*(X)$ stands for the set of all non-empty subsets of X.

Let $N = \{1, \dots, T\}$. Depending on the intended application, the set N may represent either T distinct properties (in multi-criteria decision analysis) or experts (in group decision making), or the results of an event with T possible outcomes, or a finite number of periods of time. The later will be the framework of this study, where time is divisible into discrete periods.

The elements of the canonical basis of $\mathbb{R}^T$ are denoted by $\mathbf{e}_i$, with $i = 1, \dots, T$.

For each $\mathbf{x} = (x_1, \dots, x_T) \in \mathbb{R}^T$, by $\mathbf{x}_{\nearrow} = (x_{(1)}, \dots, x_{(T)})$ we will denote a non-decreasing permutation of $\mathbf{x}$, i.e., a permutation of the entries of $\mathbf{x}$ such that $x_{(1)} \leqslant x_{(2)} \leqslant \dots \leqslant x_{(T)}$. An alternative presentation uses a permutation of the indices, i.e., a bijective mapping $\sigma : N \mapsto N$, in such way that $\mathbf{x}_{\nearrow} = \mathbf{x}_{\sigma} = (x_{\sigma(1)}, \dots, x_{\sigma(T)})$. And $\mathbf{x}_{\searrow} = (x_{[1]}, \dots, x_{[T]})$ will denote a non-increasing rearrangement of $\mathbf{x}$, i.e., the entries of $\mathbf{x}$ are such that $x_{[1]} \geqslant x_{[2]} \geqslant \dots \geqslant x_{[T]}$. It is also possible to reformulate this concept with permutations of the indices. A permutation that exchanges two elements, all other elements remaining unchanged, is a transposition.

The fundamental definitions and facts that we need to state our contribution concern evaluations that do not contradict basic assumptions in intertemporal comparisons. First, in subsect. 2.1 we overview several notions for the comparison of vectors, both in the case of binary relations and of numerical evaluations. Some useful relationships are established. Secondly, in subsect. 2.2 we brief the reader about the discrete Choquet integral and its fundamental properties. Then we describe two specific cases in subsect. 2.3, namely, OWA operators and weighted arithmetic means.

2.1 Comparing vectors: a few basic notions

For each $\mathbf{x} = (x_1, \dots, x_T) \in \mathbb{R}^T$ and $\mathbf{y} = (y_1, \dots, y_T) \in \mathbb{R}^T$, suppose for simplicity that they represent income allocations to T individuals. An evident situation of improvement happens when $x_i \geqslant y_i$ for each $i = 1, \ldots, T$. We represent this case by $\mathbf{x} \geqslant \mathbf{y}$, and say that $\mathbf{x}$ monotonically improves $\mathbf{y}$, or that $\mathbf{x}$ is monotonically better than $\mathbf{y}$. In different contexts, other types of transformations produce "better", "more desirable", or "more evenly distributed" vectors of allocations. Particularly, we say that:

a) $\mathbf{x}$ is obtained from $\mathbf{y}$ by a Pigou–Dalton *progressive* transfer if there is $\varepsilon > 0$ with $x_j = y_j - \varepsilon \geqslant x_k = y_k + \varepsilon$ for some $j, k \in N$, and $y_t = x_t$ when $j \neq t \neq k$. This is a transfer of $\varepsilon > 0$ from j to k, in such way that *their relative positions are not reversed*. So after the transfer, individual j still has a larger endowment than individual k.
b) $\mathbf{x}$ is obtained from $\mathbf{y}$ by a Pigou–Dalton transfer (Marshall et al., 2011, Chapter 1) if there is $y_j - y_k > \varepsilon > 0$ with $x_j = y_j - \varepsilon$, $x_k = y_k + \varepsilon$ for some $j, k \in N$, and $y_t = x_t$ when $j \neq t \neq k$. This is a transfer of $\varepsilon > 0$ from j (whose income is larger under $\mathbf{y}$) to k, that is no larger than the difference between j's and k's incomes. Note that now *the relative positions of the individuals can be reversed*: after the transfer, individual j may end up with less income than individual k. In other words: this shift of income may not be a Pigou–Dalton *progressive* transfer. Arnold (2012) labels this operation a *Robin Hood transfer*.

An h-delay (from component i to component $i+1$) of $\mathbf{x}$ is $\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1}$, for some $i = 1, \ldots, T-1$ and $h > 0$. Unit delays are the case $h = 1$.

Obviously, in an intertemporal setting where N represents periods, an h-delay defers h units from period i to period $i+1$.

We can formalize various ideas that implement the production of more 'fairly' distributed entries, or other concepts such as 'delay of a shift in income'.

a) $\mathbf{x}$ Lorenz dominates $\mathbf{y}$ (or $\mathbf{y}$ is Lorenz majorized by $\mathbf{x}$), represented $\mathbf{x} \succcurlyeq_L \mathbf{y}$, if $\sum_{i=1}^{T} x_i = \sum_{i=1}^{T} y_i$ and $\sum_{i=1}^{k} x_{(i)} \leqslant \sum_{i=1}^{k} y_{(i)}$ for each $k = 1, \ldots, T-1$. Note the following equivalence in terms of non-increasing rearrangements (Marshall et al., 2011, A.1): $\mathbf{x}$ Lorenz dominates $\mathbf{y}$ if and only if $\sum_{i=1}^{T} x_i = \sum_{i=1}^{T} y_i$ and $\sum_{i=1}^{k} x_{[i]} \geqslant \sum_{i=1}^{k} y_{[i]}$ for each $k = 1, \ldots, T-1$. The reason is that due to $\sum_{i=1}^{T} x_i = \sum_{i=1}^{T} y_i$, for each $k = 1, \ldots, T-1$, the inequality $\sum_{i=1}^{k} x_{[i]} \geqslant \sum_{i=1}^{k} y_{[i]}$ is equivalent to $\sum_{i=k+1}^{T} x_{[i]} \leqslant \sum_{i=k+1}^{T} y_{[i]}$, which is equivalent to $\sum_{i=1}^{T-k} x_{(i)} \leqslant \sum_{i=1}^{T-k} y_{(i)}$.
b) $\mathbf{x}$ cumulative dominates $\mathbf{y}$, represented $\mathbf{x} \succcurlyeq_c \mathbf{y}$, if $\sum_{i=1}^{k} x_i \geqslant \sum_{i=1}^{k} y_i$ for each $k = 1, \ldots, T$.

Of course, $\mathbf{x} \geqslant \mathbf{y}$ implies $\mathbf{x} \succcurlyeq_c \mathbf{y}$. Both $\geqslant$, $\succcurlyeq_L$, and $\succcurlyeq_c$ are reflexive and transitive.

To see why these definitions matter, suppose for simplicity that the vectors distribute income or wealth. Intuitively, when $\mathbf{x}$ Lorenz dominates $\mathbf{y}$ / $\mathbf{y}$ is Lorenz majorized by $\mathbf{x}$, the vector of endowments $\mathbf{y}$ is more evenly distributed than $\mathbf{x}$. Indeed, the same quantity of wealth is distributed, but also the richest person in $\mathbf{x}$ has no less than the richest person in $\mathbf{y}$ / the richest person in $\mathbf{y}$ has no more than the richest person in $\mathbf{x}$, the richest and second richest persons in $\mathbf{x}$ have no less than the richest and second richest persons in $\mathbf{y}$ / the richest person in $\mathbf{y}$ has no more than the richest person in $\mathbf{x}$, and so forth.

Clearly, if $\mathbf{x}$ is obtained from $\mathbf{y}$ by a Pigou–Dalton transfer then $\mathbf{x}$ Lorenz dominates $\mathbf{y}$. Conversely, Lorenz dominating vectors can always be obtained by a finite sequence of Pigou–Dalton transfers (which may not be progressive), and also by a finite sequence of Pigou–Dalton progressive transfers if permutations are allowed as well. This sequence is empty when $\mathbf{x} = \mathbf{y}$. To summarize, one has:

Lemma 2.1 (Marshall et al., 2011; Banerjee, 2014; Aristondo et al., 2013) For each $\mathbf{x} = (x_1, \ldots, x_T) \in \mathbb{R}^T$ and $\mathbf{y} = (y_1, \ldots, y_T) \in \mathbb{R}^T$, the following statements are equivalent:

1. $\mathbf{x}$ Lorenz dominates $\mathbf{y}$
2. $\mathbf{y}$ is obtained from $\mathbf{x}$ by a finite sequence of successive Pigou–Dalton transfers.
3. $\mathbf{y}$ is obtained from $\mathbf{x}$ by a finite sequence of successive Pigou–Dalton progressive transfers and/or permutations.

Example 2.1 With $T = 3$, let $\mathbf{x} = (0, 5, 4) \in \mathbb{R}^3$ and $\mathbf{y} = (4, 2, 3) \in \mathbb{R}^3$, hence $\mathbf{x}_{\nearrow} = (0, 4, 5)$ and $\mathbf{y}_{\nearrow} = (2, 3, 4)$. Then $\mathbf{x}$ Lorenz dominates $\mathbf{y}$ because $0 + 4 + 5 = 2 + 3 + 4$, $0 \leqslant 2$ and $0 + 4 \leqslant 2 + 3$.

We can obtain $\mathbf{y}$ from $\mathbf{x}$ doing two Pigou–Dalton transfers, namely, a transfer of 4 units from the second to the first individual, which produces (4, 1, 4) from $\mathbf{x} = (0, 5, 4)$, and then a transfer of 1 unit from the third to the second individual, which yields $\mathbf{y} = (4, 2, 3)$ from (4, 1, 4). Note that the first Pigou–Dalton transfer is not progressive.

Alternatively, we can obtain $\mathbf{y}$ from $\mathbf{x}$ doing one transposition and two successive Pigou–Dalton progressive transfers. First we obtain (5, 0, 4) from $\mathbf{x} = (0, 5, 4)$ by a transposition of the endowments of the first and second individuals. Then a transfer of 1 unit from the first to the second individual produces (4, 1, 4) from (5, 0, 4). Now a transfer of 1 unit from the third to the second individual produces $\mathbf{y} = (4, 2, 3)$ from (4, 1, 4).

In each framework (e.g., welfare economics or intertemporal analysis), an appropriate evaluation of vectors must be done by functions that respect reasonable properties. To define some apposite ideas, fix $X \subseteq \mathbb{R}$ and a function $\mathbf{f} : X^T \to \mathbb{R}$. Then we say that:

1. The function $\mathbf{f}$ is monotonic if $\mathbf{x} \geqslant \mathbf{y}$ implies $\mathbf{f}(\mathbf{x}) \geqslant \mathbf{f}(\mathbf{y})$, for each $\mathbf{x}, \mathbf{y} \in X^T$.
2. The function $\mathbf{f}$ satisfies the **Pigou–Dalton progressive principle** if when $\mathbf{x}, \mathbf{y} \in X^T$ and $\mathbf{x}$ is obtained from $\mathbf{y}$ by a Pigou–Dalton progressive transfer, then $\mathbf{f}(\mathbf{x}) > \mathbf{f}(\mathbf{y})$. Beliakov and James (2021, Definition 10) call this property the Pigou–Dalton (or the progressive transfers) principle.
3. The function $\mathbf{f}$ satisfies the **Pigou–Dalton principle** if when $\mathbf{x}, \mathbf{y} \in X^T$ and $\mathbf{x}$ is obtained from $\mathbf{y}$ by a Pigou–Dalton transfer, then $\mathbf{f}(\mathbf{x}) > \mathbf{f}(\mathbf{y})$. Note that this

property implies the Pigou–Dalton progressive principle (because progressive Pigou–Dalton transfers are Pigou–Dalton transfers).

4. The function $\mathbf{f}$ is Schur-concave or **S-concave** (Marshall et al., 2011; Beliakov & James, 2021) if for each $\mathbf{x}, \mathbf{y} \in X^T$, $\mathbf{x} \succcurlyeq_L \mathbf{y}$ implies $\mathbf{f}(\mathbf{x}) \geqslant \mathbf{f}(\mathbf{y})$, i.e., if $\mathbf{f}(\mathbf{x}) \geqslant \mathbf{f}(\mathbf{y})$ whenever $\mathbf{x}$ Lorenz dominates $\mathbf{y}$. Strict S-concavity holds when in addition to S-concavity, $\mathbf{x} \neq \mathbf{y}$ always implies $\mathbf{f}(\mathbf{x}) > \mathbf{f}(\mathbf{y})$. Note that when $\mathbf{f}$ is S-concave then it is **symmetric**, i.e., $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_{\nearrow})$ for all $\mathbf{x} \in X^T$.
5. The function $\mathbf{f}$ is consistent with cumulative domination if for each $\mathbf{x}, \mathbf{y} \in X^T$, $\mathbf{x} \succcurlyeq_c \mathbf{y}$ implies $\mathbf{f}(\mathbf{x}) \geqslant \mathbf{f}(\mathbf{y})$, i.e., if $\mathbf{f}(\mathbf{x}) \geqslant \mathbf{f}(\mathbf{y})$ whenever $\mathbf{x}$ cumulative dominates $\mathbf{y}$. In this case we simply say that $\mathbf{f}$ is ***c*-monotonic**. Note that this property implies that $\mathbf{f}$ is monotonic. We also say that $\mathbf{f}$ is strongly consistent with cumulative domination if $\mathbf{f}(\mathbf{x}) > \mathbf{f}(\mathbf{y})$ whenever $\mathbf{x}$ cumulative dominates $\mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$. Then we simply say that $\mathbf{f}$ is **strongly *c*-monotonic**. This property implies that $\mathbf{f}$ is c-monotonic thus monotonic.
6. The function $\mathbf{f}$ is consistent with Axiom Z if $\epsilon \in X^T$, $\epsilon \succcurlyeq_c 0$ implies $\mathbf{f}(\epsilon) \geqslant 0$.
7. The function $\mathbf{f}$ is delay-averse if $\mathbf{f}(\mathbf{x}) \geqslant \mathbf{f}(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1})$ when $i \in \{1, \ldots, T-1\}$, $h > 0$, and $\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1} \in X^T$. It is strongly delay-averse if we require $\mathbf{f}(\mathbf{x}) > \mathbf{f}(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1})$ in the conclusion. Obviously, strongly delay-averse implies delay-averse.

In cases such as $X = \mathbb{R}^T_+$, $\mathbf{f}$ is delay-averse if and only if $\mathbf{f}(\mathbf{x}) \geqslant \mathbf{f}(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_j)$ when $i, j \in N$, $j > i$, $h > 0$, and $\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_j \in X^T$. We only need to apply a simple recursive argument to prove this equivalence.

A comparative tool that is weaker than an evaluative function is a binary relation. A preorder on X^T is a reflexive and transitive binary relation R on X^T. It is complete when for all $\mathbf{x}, \mathbf{y} \in X^T$, either $\mathbf{x}R\mathbf{y}$ or $\mathbf{y}R\mathbf{x}$ is true. The concepts above can be transferred *ceteris paribus* to this language. For example, a binary relation R on X^T is monotonic if $\mathbf{x} \geqslant \mathbf{y}$ implies $\mathbf{x}R\mathbf{y}$, for each $\mathbf{x}, \mathbf{y} \in X^T$. Let P denote its asymmetric part,[1] Then R satisfies the Pigou–Dalton progressive principle if when $\mathbf{x}, \mathbf{y} \in X^T$ and $\mathbf{x}$ is obtained from $\mathbf{y}$ by a Pigou–Dalton progressive transfer, then $\mathbf{x}P\mathbf{y}$. And R satisfies Axiom Z if $\epsilon \in X^T$, $\epsilon \succcurlyeq_c 0$ implies $\epsilon R 0$. Below we establish some properties that can be exported to the case of evaluative functions (v., Appendix A for a proof):

Proposition 2.1 Let R be a transitive binary relation on X^T, and suppose $X \subseteq \mathbb{R}$ is closed under sums. Then:

1. R is strict S-concave if and only if R satisfies the Pigou–Dalton principle. This case implies that R also satisfies the Pigou–Dalton progressive principle.
2. R is c-monotonic (resp., strongly c-monotonic) if and only if for each $\epsilon \succcurlyeq_c 0$ (resp., with $\epsilon \neq 0$) and $\epsilon, x \in X^T$, one has $(\mathbf{x} + \epsilon)R\mathbf{x}$ (resp., $(\mathbf{x} + \epsilon)P\mathbf{x}$). In particular, if R is c-monotonic then R satisfies Axiom Z and it is monotonic.

[1] The asymmetric part P of R is defined by $\mathbf{x}P\mathbf{y}$ if and only if $\mathbf{x}R\mathbf{y}$ but not $\mathbf{y}R\mathbf{x}$, for all $\mathbf{x}\mathbf{y} \in X^T$.

3. If R is reflexive and additive (i.e., $\mathbf{x}R\mathbf{y}$, $\mathbf{x}'R\mathbf{y}'$ implies $(\mathbf{x}+\mathbf{x}')R(\mathbf{y}+\mathbf{y}')$ for all $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in X^T$) and it satisfies Axiom Z then R is c-monotonic.
4. If R is c-monotonic then it is delay-averse. The reverse is not true. However, if R is a monotonic and delay-averse preorder, then it is c-monotonic.
5. If R is strongly c-monotonic then it is strongly delay-averse. The reverse is not true. However, if R is a monotonic and strongly delay-averse preorder, then it is strongly c-monotonic.

Despite claim 4 in Proposition 2.1, delay-aversion and c-monotonicity are not in general equivalent (v., Example A.1 in Appendix A).

2.2 The discrete Choquet integral

In this article we will work with non-negative real inputs only. In this context, aggregation operators are functions of T arguments mapping T-dimensional vectors with non-negative real components onto non-negative real numbers, that are monotonic and satisfy a boundary condition:

Definition 2.1 An **aggregation operator** on $\mathbb{R}_+^T$ is a mapping $f : \mathbb{R}_+^T \longrightarrow \mathbb{R}_+$ which is monotonic (i.e., $f(\mathbf{x}) \leqslant f(\mathbf{y})$ whenever $\mathbf{x} \leqslant \mathbf{y}$) and satisfies $f(0, \dots, 0) = 0$.

The standard version of aggregation operators (Beliakov et al., 2007, Definition 1.5) alludes to mappings $f : [0,1]^T \longrightarrow [0,1]$, for which in addition to the restriction established above, it is also required that $f(1, \dots, 1) = 1$.

The discrete Choquet integral provides a remarkable class of aggregation operators that are constructed from a discrete capacity defined as follows:

Definition 2.2 A discrete capacity (or a **fuzzy measure**) on N is a set function $\mu : 2^N \longrightarrow [0,1]$ which is monotonic (i.e., $\mu(A) \leqslant \mu(A')$ whenever $A \subseteq A' \subseteq N$) and satisfies $\mu(\varnothing) = 0$, $\mu(N) = 1$.

If $\mu(A) = \mu(B)$ when $A, B \subseteq N$ are such that $|A| = |B|$, then the capacity μ is **symmetric**. Also, **additivity** holds if it is the case that when $A, B \subseteq N$ are disjoint then $\mu(A \cup B) = \mu(A) + \mu(B)$. Additive set functions are uniquely determined by $\{\mu(1), \dots, \mu(T)\}$. An additive normalized (i.e., with the property $\mu(N) = 1$, that is sometimes dropped) capacity is a probability measure.

Capacities for which $\mu(A \cup B) \geqslant \mu(A) + \mu(B)$, respectively, $\mu(A \cup B) \leqslant \mu(A) + \mu(B)$, whenever A and B are disjoint non-empty subsets of N, are called **superadditive**, respectively, **subadditive**.

Furthermore, the capacity μ is **supermodular** when

$$\mu(A \cup B) + \mu(A \cap B) \geqslant \mu(A) + \mu(B) \text{ for all } A, B \subseteq N \tag{1}$$

and it is **submodular** when

$$\mu(A \cup B) + \mu(A \cap B) \leqslant \mu(A) + \mu(B) \text{ for all } A, B \subseteq N. \tag{2}$$

Supermodular capacities are superadditive, and submodular capacities are subadditive. For this reason, another term for supermodular, resp., submodular, is “strongly superadditive”, resp., “strongly subadditive” (e.g., Anger (1977)). In a wider framework, supermodularity boils down to 2-monotonicity and submodularity to the 2-alternating property (Grabisch, 2016, Remark 2.19).

It is worth mentioning that de Campos Ibáñez and Bolaños Carmona (1989, Propositions 3.4, 4.2) characterize submodularity and supermodularity of a capacity.

Since we only work with non-negative real inputs in this article, the discrete Choquet integral can be defined in the following way:

Definition 2.3 The **discrete Choquet integral** $C^\mu : \mathbb{R}_+^T \longrightarrow \mathbb{R}_+$ with respect to a discrete capacity μ is defined as: for each $\mathbf{x} = (x_1, \dots, x_T) \in \mathbb{R}_+^T$,

$$C^\mu(x_1, \dots, x_T) = \sum_{i=1}^{T} \left(x_{(i)} - x_{(i-1)}\right) \mu(H_i), \tag{3}$$

provided that we set $x_{(0)} = 0$ by convention, and $H_i = \{(i), \dots, (T)\}$ (thus H_i captures the set of indices of the largest $T - i + 1$ components of $\mathbf{x}$).

An alternative definition uses the concept of discrete derivative of the capacity μ at A with respect to i, which is $\Delta_i \mu(A \cup \{i\}) = \mu(A \cup \{i\}) - \mu(A)$, for each $A \subseteq N \backslash \{i\}$ and $i \in N$:

Definition 2.4 In the conditions of Definition 2.3,

$$C^\mu(x_1, \dots, x_T) = \sum_{i=1}^{T} \Delta_{(i)} \mu(H_i) x_{(i)}, \text{ for each } \mathbf{x} = (x_1, \dots, x_T) \in \mathbb{R}_+^T. \tag{4}$$

Either symmetric or asymmetric Choquet integral must be used when working with positive and negative integrands (Grabisch et al., 2009, Section 5.4). The extension of the formulas presented above to this general case produces the asymmetric Choquet integral, which in fact is the most widely acknowledged definition of the discrete Choquet integral for real integrands (Grabisch et al., 2009, Remark 5.24). Note that this is no longer true in the continuous version of the formula.

The Choquet integral is continuous with respect to the Euclidean topology. It is neither additive nor symmetric, but it has a number of other interesting properties such as the following list.

- **Compensativeness**: $\forall \mathbf{x} = (x_1, \dots, x_T) \in \mathbb{R}_+^T$,

  $$\min\{x_1, \dots, x_T\} \leqslant C^\mu(x_1, \dots, x_T) \leqslant \max\{x_1, \dots, x_T\}.$$

  Alternatively, we say that the Choquet integral is **averaging** (Beliakov et al., 2007, Definition 1.7). In particular, the Choquet integral is **idempotent**, i.e., $C^\mu(x, \dots, x) = x$ when $x \in \mathbb{R}_+$.
- **Monotonicity**: $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}_+^T$, $\mathbf{x} \leqslant \mathbf{y} \Rightarrow C^\mu(\mathbf{x}) \leqslant C^\mu(\mathbf{y})$.

- **Positive homogeneity** (of degree 1): $\forall \mathbf{x} \in \mathbb{R}_+^T$, $\forall \lambda \in \mathbb{R}_+$, $C^\mu(\lambda \mathbf{x}) = \lambda\, C^\mu(\mathbf{x})$.
- **Translation invariance** (also, shift-invariance): $\forall \mathbf{x} \in \mathbb{R}_+^T$, $\forall x \in \mathbb{R}_+$, $C^\mu(\mathbf{x} + (x, \dots, x)) = C^\mu(\mathbf{x}) + x$.
- **Comonotone additivity**:$C^\mu(\mathbf{x} + \mathbf{y}) = C^\mu(\mathbf{x}) + C^\mu(\mathbf{y})$ when $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^T$ are comonotone. Recall that $\mathbf{x}, \mathbf{y}$ are comonotone when a common permutation $\{(1), \dots, (T)\}$ of N exists, such that $\mathbf{x}_\nearrow = (x_{(1)}, \dots, x_{(T)})$ and $\mathbf{y}_\nearrow = (y_{(1)}, \dots, y_{(T)})$ are respective non-decreasing permutations of $\mathbf{x}$ and $\mathbf{y}$. This is equivalent to the property that $(x_j - x_i)(y_j - y_i) \geqslant 0$ for each $i, j \in N$ (Grabisch, 2016, Lemma 4.27 *iv*)). A comparable test holds true for comonotonicity of functions (Grabisch, 2016, Definition 4.16). Note that being comonotonic is not a transitive property (Grabisch, 2016, p. 215).

Note that translation invariance implies that $C^\mu(\mathbf{x} - (x, \dots, x)) = C^\mu(\mathbf{x}) - x$, for all $\mathbf{x} \in \mathbb{R}_+^T$ and $x \in \mathbb{R}_+$ with $\mathbf{x} - (x, \dots, x) \in \mathbb{R}_+$.

Positive homogeneity permits to prove the next equivalence in a simple manner:

Lemma 2.2 The Choquet integral is convex (resp., concave) if and only if it is subadditive (resp., superadditive).

Recall that **convexity** means

$$C^\mu(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leqslant \lambda C^\mu(\mathbf{x}) + (1 - \lambda) C^\mu(\mathbf{x})$$

for all $\mathbf{x} = (x_1, \dots, x_T), \mathbf{y} = (y_1, \dots, y_T) \in \mathbb{R}_+^T$, and for all $\lambda \in [0, 1]$. And **subadditivity** means in this context

$$C^\mu(\mathbf{x} + \mathbf{y}) \leqslant C^\mu(\mathbf{x}) + C^\mu(\mathbf{y})$$

for all $\mathbf{x} = (x_1, \dots, x_T), \mathbf{y} = (y_1, \dots, y_T) \in \mathbb{R}_+^T$. Concavity and superadditivity are dual properties, which are stated by reversing the respective inequalities.

A property that is much more difficult to prove (although many proofs are available) relates the convexity or subadditivity of the Choquet integral with submodularity of the capacity that defines it. And the same holds for concavity or superadditivity of the Choquet integral with supermodularity of the capacity.

Proposition 2.2 (Alfonsi, 2015; Buja, 1984; Choquet, 1954; Denneberg, 2013) The discrete Choquet integral $C^\mu : \mathbb{R}_+^T \longrightarrow \mathbb{R}_+$ with respect to a discrete capacity μ is convex and subadditive (resp., concave and superadditive) if and only if μ is submodular (resp., supermodular).

Huber (2004) directly characterizes subadditivity (resp., superadditivity) of C^μ by submodularity (resp., supermodularity) of μ: v., de Campos Ibáñez and Bolaños Carmona (1989, Proposition 4.3).

Another class of capacities with important implications for aggregation by fuzzy integrals are buoyant/antibuoyant capacities (Beliakov &

James, 2021, Definition 8). We say that the capacity μ is **buoyant** when $\Delta_i\mu(A \cup \{i,j\}) \leqslant \Delta_j\mu(A \cup \{j\})$ for all $i \neq j \in N$ and $A \subseteq N \backslash \{i,j\}$. **Antibuoyancy** is defined by reversing the inequality in the definition of buoyancy. The next Proposition summarizes evidence of the significance of antibuoyancy:

Proposition 2.3 (Beliakov & James, 2021) Let μ be a capacity on N.

1. If μ is antibuoyant (resp., buoyant) then μ must be supermodular (resp., submodular), but not conversely.
2. If μ is antibuoyant (resp., buoyant) then C^μ must be concave (resp., convex).
3. μ is antibuoyant if and only if C^μ satisfies the Pigou–Dalton progressive principle.

Beliakov and James (2021, Proposition 3) proved the first assertion. Now using Proposition 2.2, one deduces from this fact that Choquet integrals with respect to antibuoyant capacities must be concave. Example 4.5 below demonstrates that submodular discrete capacities exist that are not buoyant. Beliakov and James (2021, Proposition 8) prove that a capacity μ is antibuoyant if and only if C^μ is consistent with the Pigou–Dalton progressive principle.

Considering that the Pigou–Dalton principle is logically stronger than the Pigou–Dalton progressive principle, one should conclude that antibuoyant capacities do not necessarily produce Choquet integrals that satisfy the Pigou–Dalton principle. The next example is a testament to this fact. Note that henceforth, when T is small we will represent discrete capacities on N by arrangement in an array that replicates the Hasse diagram of the subsethood relation defined on the parts of N (cf., Figure 1).

Example 2.2 With $T = 3$, consider the following antibuoyant discrete capacity:

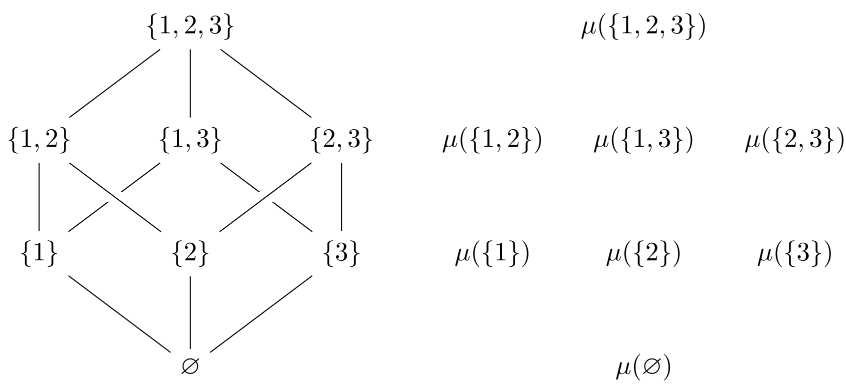

Fig. 1 A convention for the presentation of discrete capacities when $T = 3$ (right). It is inspired by the Hasse diagram of the subsethood relation defined on $\mathscr{P}(\{1,2,3\})$ (left)

$$\mu(\{1\}) = \mu(\{2\}) = \mu(\{1,2\}) = 0,$$
$$\mu(\{3\}) = 0.1,\ \mu(\{1,3\}) = \mu(\{2,3\}) = 0.2,\ \mu(\{1,2,3\}) = 1$$

that we summarize by

$$\begin{array}{ccc} & 1 & \\ 0 & 0.2 & 0.2 \\ 0 & 0 & 0.1 \\ & 0 & \end{array}$$

Then C^μ defines a discrete Choquet integral $C^\mu : \mathbb{R}^3_+ \longrightarrow \mathbb{R}_+$ that is not consistent with the Pigou–Dalton principle because $C^\mu(5,0,10) > C^\mu(5,10,0)$.

As far as we know, the characterization of capacities that produce C^μ with the Pigou–Dalton principle (i.e., that are strict S-concave) is an open question.

We underline that Beliakov and James (2021, Proposition 7) characterize the λ -fuzzy measures, also called Sugeno measures, that are antibuoyant. Let us recall this important concept:

Definition 2.5 A discrete capacity μ on N is a λ-fuzzy measure, where $\lambda \in (-1, +\infty)$ is fixed, when $\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\, \mu(A)\, \mu(B)$ for all disjoint $A, B \subseteq N$.

For a λ-fuzzy measure μ, both λ and the collection of values $\mu(\{i\})$ with $i \in N$ suffice to define the whole μ. The value of λ is related to the collection of values $\{\mu(\{i\})\}_{i\in N}$ by a normalization condition proven by Leszczyński et al. (1985) and described e.g., in Grabisch (2016, Sect. 2.8.6) or Wang and Klir (1992, Theorem 4.7), namely,

$$\lambda + 1 = \prod_{j=1}^{T} (1 + \lambda\mu(\{j\})). \tag{5}$$

In addition, one can observe $\Delta_i \mu(A \cup \{i\}) = \mu(A)\,(1 + \lambda\, \mu(\{i\})$, for each $A \subseteq N \backslash \{i\}$ and $i \in N$. Besides, λ-fuzzy measures are supermodular therefore superadditive when $\lambda \geqslant 0$, otherwise they are submodular (Narukawa et al., 2023). In fact, λ-fuzzy measures are belief measures (therefore totally monotone) when $\lambda \geqslant 0$, and they are plausibility measures (therefore totally alternating) when $\lambda \in (-1, 0]$ (Grabisch, 2016, Theorem 2.28).

Regrettably, Beliakov and James (2021) neither give any explicit example of buoyant or antibuoyant λ-fuzzy measures, nor characterize buoyant λ-fuzzy measures. We next show an example with a buoyant λ-fuzzy measure. The characterization of buoyant λ-fuzzy measures can be achieved by a duality argument coupled with the characterization of antibuoyancy.

Example 2.3 With $T = 3$, consider the λ-discrete capacity that stems from (we drop brackets for simplicity):

$$\mu(1) = 0.9,\ \mu(2) = 0.8,\ \mu(3) = 0.7.$$

We use (5) to compute that $\lambda = -0.993373$ produces the λ-discrete capacity

$$\begin{array}{ccc} & 1 & \\ 0.984771 & 0.974175 & 0.943711 \\ 0.9 & 0.8 & 0.7 \\ & 0 & \end{array}$$

This is a buoyant discrete capacity.

Henceforth we will use the following notation: when $A \subseteq N = \{1, \ldots, T\}$, $\chi_A \in \{0, 1\}^T$ is the binary T-dimensional vector whose components are 1 exactly for the indices defined by A, i.e.,

$$\chi_A = (\chi_1, \ldots, \chi_T) = (\chi_A(1), \ldots, \chi_A(T)) \text{ such that } \begin{cases} \chi_i = \chi_A(i) = 1 & \text{if } i \in A, \\ \chi_i = \chi_A(i) = 0 & \text{if } i \notin A. \end{cases}$$

The proof of the next Lemma is straightforward (Grabisch, 2016, states its continuous version):

Lemma 2.3 Let μ be a discrete capacity on N. If $A_1 \supseteq \ldots \supseteq A_k$ is any chain of (possibly repeated) subsets of N, then $C^\mu(\sum_{j=1}^k \chi_{A_j}) = \sum_{j=1}^k \mu(A_j)$.

2.3 Two particular cases of the Choquet integral

When μ is symmetric, Definition 2.3 yields an OWA operator on the domain (Yager, 1988). To define OWA, assume that $\mathbf{w} = (w_1, \ldots, w_T) \in [0, 1]^T$ is a weighting vector with $\sum_{i=1}^T w_i = 1$. The OWA operator associated with these weights is $F^{\mathbf{w}} : \mathbb{R}^T \longrightarrow \mathbb{R}$ defined as

$$F^{\mathbf{w}}(x_1, \ldots, x_T) = \sum_{i=1}^{T} w_i x_{[i]} \text{ for each } (x_1, \ldots, x_T) \in \mathbb{R}^T, \tag{6}$$

where $x_{[i]}$ is the i-th largest element in the collection of (possibly repeated) values $\{x_1, \ldots, x_T\}$.

In the case where C^μ is the Choquet integral associated with a symmetric capacity, defining $w_i = \mu(H_{T-i+1}) - \mu(H_{T-i})$, one has $C^\mu = F^{\mathbf{w}}$ on the restricted domain of non-negative entries. If the capacity μ is additive, then Definition 2.3 produces the weighted arithmetic mean associated with weights $w_i = \mu(\{i\})$, namely, for $\mathbf{w} = (\mu(\{1\}), \ldots, \mu(\{T\})) \in [0, 1]^T$,

$$C^\mu(x_1, \ldots, x_T) = M_{\mathbf{w}}(x_1, \ldots, x_T) = \sum_{i=1}^{T} w_i x_i \text{ for each } (x_1, \ldots, x_T) \in \mathbb{R}^T_+. \tag{7}$$

So symmetric and additive capacities define a Choquet integral that is the arithmetic mean, defined as follows:

$$M(x_1, \dots, x_T) = \frac{\sum_{i=1}^{T} x_i}{T} \text{ for each } (x_1, \dots, x_T) \in \mathbb{R}_+^T. \tag{8}$$

3 Some bounds on *c*-monotonic evaluations

Statement 5 in Proposition 2.1 helps to produce examples of strongly delay-averse and c-monotonic complete preorders that are representable by a utility. Let us recall this concept:

Definition 3.1 A complete preorder $\succcurlyeq$ on a set X is representable by a utility when there is $f : X \longrightarrow \mathbb{R}$ such that for all $x, y \in X$, $x \succcurlyeq y$ if and only if $f(x) \geqslant f(y)$.

The next examples link the intertemporal models of evaluation of future gains with strong c-monotonicity and delay-aversion. Fix $X = \mathbb{R}_+$.

1. When $\succcurlyeq$ is a complete preorder on X^N representable by an exponential discounting $\mathbf{f}_\delta : X^T \to \mathbb{R}$ such that $\mathbf{f}_\delta(\mathbf{x}) = \sum_{t=1}^{T} \delta^t x_t$ for a fixed $\delta \in (0, 1)$, then $\succcurlyeq$ is strongly c-monotonic because it is obviously monotonic and strongly delay-averse.
2. When $\succcurlyeq$ is a complete preorder on X^N representable by a hyperbolic discounting $\mathbf{h}_k : X^T \to \mathbb{R}$ such that $\mathbf{h}_k(\mathbf{x}) = \sum_{t=1}^{T} \frac{1}{1+kt} x_t$ (for a fixed $k > 0$) then $\succcurlyeq$ is strongly c-monotonic, also because it is monotonic and strongly delay-averse.
3. When $\succcurlyeq$ is a complete preorder on X^N representable by a quasi-hyperbolic discounting $\mathbf{q}_{\beta,\delta} : X^T \to \mathbb{R}$ such that $\mathbf{q}_{\beta,\delta}(\mathbf{x}) = \beta \sum_{t=1}^{T} \delta^t x_t$ for fixed $\beta, \delta \in (0, 1)$, then $\succcurlyeq$ is strongly c-monotonic, also because it is monotonic and strongly delay-averse.

However, finding operators that in addition satisfy symmetry is a tougher problem. In fact, there is a tension between c-monotonicity and symmetry, that can only be resolved by the standard arithmetic mean among all idempotent aggregation operators. This issue can be captured by the following result, whose proof is given in Appendix B:

Proposition 3.1 Suppose that the mapping $f : \mathbb{R}_+^T \longrightarrow \mathbb{R}$ is symmetric, idempotent and c-monotonic. Then f is the arithmetic mean.

Thus the disappointing message is that among all OWA operators, only the arithmetic mean is c-monotonic (see Remark 3.1 below for an independent proof). This result renders moot the entire class of OWA operators, except for the arithmetic mean, in the implementation of c-monotonic aggregation operators.

The geometric mean proves that we cannot relax c-monotonicity to monotonicity in Proposition 3.1.

Remark 3.1 A direct proof that only the arithmetic mean is a c-monotonic OWA operator proceeds as follows. We consider the case when $T = 2$ for simplicity.

Suppose first that the weighting vector $\mathbf{w} = (w_1, w_2) \in [0,1]^2$ satisfies $w_1 > w_2$, hence $w_1 > \frac{1}{2}$. Fix $K \in (\frac{1}{2}, w_1)$. Then $(K, K) \succ^c (0, 1)$, but $F^{\mathbf{w}}(K, K) = K < w_1 = F^{\mathbf{w}}(0, 1)$, contradicting c-monotonicity.

In case $w_1 < w_2$, since $w_1 < \frac{1}{2}$ we can mimic the argument with $(1, 0) \succ^c (K, K)$ for any fixed $K \in (w_1, \frac{1}{2})$. Now $F^{\mathbf{w}}(1, 0) = w_1 < K = F^{\mathbf{w}}(K, K)$, contradicting c-monotonicity.

We conclude $w_1 = w_2$.

4 Characterization of strongly delay-averse Choquet integrals

The main goal of this section is to identify the capacities that produce strongly delay-averse (or equivalently, c-monotonic) Choquet integrals. After some preliminary considerations, we prove technical implications of this type of Choquet integrals in Sect. 4.1. In view of the importance of submodular discrete capacities in the literature, in Sect. 4.2 we give a simple proof that the combination of submodularity with a necessary condition for strong delay-aversion guarantees that the Choquet integral is strongly delay-averse. Then we solve the main problem in Sect. 4.3.

To begin with, we reveal a general incompatibility that is easy to justify and will shed light on our problem. We state it for mappings. One can immediately produce a version for binary relations.

Proposition 4.1 Suppose that $[0,1]^T \subseteq X$. Then there is no mapping $f : X \longrightarrow \mathbb{R}$ that is delay-averse and satisfies the Pigou–Dalton progressive principle.

Proof By the Pigou–Dalton progressive principle, $f(1, 0, \dots, 0) < f(\frac{2}{3}, \frac{1}{3}, 0, \dots, 0)$. But $f(1, 0, \dots, 0) \geqslant f(\frac{2}{3}, \frac{1}{3}, 0, \dots, 0)$ because f is delay-averse. □

Now the message is quite transparent: both strict S-concavity and the Pigou–Dalton principle, in addition to the (logically weaker) Pigou–Dalton progressive principle, are unsuitable for basic intertemporal analyses (cf., Statement 1 in Proposition 2.1). These three properties are incompatible with c-monotonicity, owing to Statement 4 in Proposition 2.1.

Therefore in our investigation of Choquet integrals in the intertemporal framework, we must turn our attention to other admissible properties. The very interesting Proposition 2.3 is of no use in exploring intertemporal choices although it sets an excellent precedent.

4.1 Implications of strongly *c*-monotonic Choquet integrals

Choquet integrals are c-monotonic (resp., strongly c-monotonic) if and only if they are delay-averse (resp., strongly delay-averse), due to Statement 4 (resp., Statement 5) in Proposition 2.1. This section explores the properties of the capacities that produce strongly delay-averse, or strongly c-monotonic, Choquet integrals. Henceforth in this section 4.1, μ will denote a discrete capacity on N, with $|N| > 1$, such that $C^\mu : \mathbb{R}_+^T \longrightarrow \mathbb{R}_+$ is both strongly delay-averse and strongly c-monotonic, unless otherwise stated.

Note that when $A, A' \subseteq N = \{1, \dots, T\}$ have the same number of elements (i.e., when $|A| = |A'|$), and $\chi_A \succcurlyeq_c \chi_{A'}$, it is possible to obtain A' from A by a finite number of successive 1-delays. Using Lemma 2.3, we conclude:

Proposition 4.2 Let μ be a discrete capacity on N such that C^μ is strongly delay-averse. Then for each $A, A' \subseteq N$ such that $A \neq A'$, $\chi_A \succcurlyeq_c \chi_{A'}$, and $|A| = |A'|$, one has $\mu(A) > \mu(A')$.

In particular, μ in the conditions of Proposition 4.2 must satisfy

$$\text{for all } i, j \in N,\ j > i \text{ implies } \mu(\{i\}) > \mu(\{j\}). \tag{9}$$

It is trivial to prove that in addition, such μ satisfies

$$\mu(A \setminus \{i\}) < \mu(A \setminus \{j\}) \text{ for all } i, j \in N,\ j > i, \ \text{ and } i, j \in A \subseteq N. \tag{10}$$

One reason is that $C^\mu(\chi_{A\setminus\{j\}}) > C^\mu(\chi_{A\setminus\{i\}})$ by strong delay-aversion, which boils down to $\mu(A \setminus \{j\}) > \mu(A \setminus \{i\})$. Subtracting from $\mu(A)$ each side of the inequality (10), we deduce the next consequence of strong delay-aversion:

Proposition 4.3 Let μ be a discrete capacity on N such that C^μ is strongly delay-averse. Then

$$\Delta_i \mu(A) > \Delta_j \mu(A) \text{ for all } i, j \in A \subseteq N \ \text{ with } j > i. \tag{11}$$

Example 4.1 With $T = 3$, Proposition 4.2 imposes the necessary conditions $\mu(\{1,2\}) > \mu(\{1,3\}) > \mu(\{2,3\})$, in addition to (9) which means $\mu(\{1\}) > \mu(\{2\}) > \mu(\{3\})$.

The simple restriction (9) permits to discard that specific discrete capacities define strongly delay-averse Choquet integrals by inspection:

Example 4.2 With $T = 3$, consider the antibuoyant discrete capacity in Example 2.2. Then C^μ defines a discrete Choquet integral $C^\mu : \mathbb{R}_+^3 \longrightarrow \mathbb{R}_+$ that is not delay-averse because $C^\mu(5, 0, 10) > C^\mu(5, 10, 0)$. We can anticipate this problem because the necessary condition $\mu(\{2\}) > \mu(\{3\})$ is false.

The aforementioned necessary conditions (9) can be generalized to the set of restrictions described in the following result whose proof is in Appendix C:

Proposition 4.4 Let μ be a discrete capacity on N such that C^μ is strongly delay-averse. Then for each $\emptyset \neq A_1 \subsetneq A_2 \subsetneq \dots \subsetneq A_k \subseteq N$, letting $c = \sum_{i=1}^{k} |A_i|$ one has:

(a) If $i \leqslant i'$ for all $i' \in A_k$, $A_k \neq \{i\}$, then $\mu(i) > \frac{\sum_{j=1}^{k} \mu(A_j)}{c}$.

(b) If $i \geqslant i'$ for all $i' \in A_k$, $A_k \neq \{i\}$, then $\mu(i) < \frac{\sum_{j=1}^{k} \mu(A_j)}{c}$.

Note that the application of claim (a) to the case $k = 1$ and $A_k = \{j\}$ produces the necessary condition (9).

Proposition 4.4 gives an *ex-ante* set of restrictions on the values of μ conducive to strongly delay-averse Choquet integral. They implicitly impose constraints such as

$$i_1 < i_2 < \ldots < i_t \, (t > 1) \text{ imply } \mu(\{i_1, i_2, \ldots, i_t\}) \in \Big(t \cdot \mu(\{i_1\}), t \cdot \mu(\{i_t\})\Big) \tag{12}$$

which holds by application to $k = 1$, $A_k = \{i_1, i_2, \ldots, i_t\}$, and $i = i_1$ with statement (a), plus $i = i_t$ with statement (b). And also

$$\mu(\{i + t\}) < \frac{\mu(\{i\}) + \sum_{j=1}^{t} \mu(\{i, i+1, \ldots, i+j\})}{\frac{(t+1)(t+2)}{2}} < \mu(\{i\}) \tag{13}$$

when $i \in N$, $i + t \in \{i + 1, \ldots, T\}$, which follows from Proposition 4.4 applied to the nested sequence $\{i\} \subsetneq \{i, i+1\} \subsetneq \ldots \subsetneq \{i, i+1, \ldots, i+t\} \subseteq N$ (here $k = t + 1$, $c = \frac{1+(t+1)}{2}(t + 1)$).

We now proceed to establish another implication of strong delay-aversion in the vein of Proposition 4.3.

Proposition 4.5 Let μ be a discrete capacity on N such that C^μ is strongly delay-averse. Then

$$\Delta_i \mu(A) > \Delta_j \mu(B) \text{ for all } B \subsetneq A \subseteq N \text{ with } i \in A \setminus B, j \in B. \tag{14}$$

Proof In this case $C^\mu(\chi_A + \chi_{B\setminus\{j\}}) > C^\mu(\chi_{A\setminus\{i\}} + \chi_B)$ by strong delay-aversion, and because $B \setminus \{j\} \subseteq A$ and $B \subseteq A \setminus \{i\}$ we can use Lemma 2.3 as follows:

$$\begin{aligned} C^\mu(\chi_A + \chi_{B\setminus\{j\}}) =& \mu(A) + \mu(B \setminus \{j\}) > \\ \mu(A \setminus \{i\}) + \mu(B) =& C^\mu(\chi_{A\setminus\{i\}} + \chi_B). \end{aligned}$$

This inequality is equivalent to $\Delta_i \mu(A) > \Delta_j \mu(B)$. □

Another implication of the property that a discrete capacity μ on N generates a strongly delay-averse C^μ is that μ cannot satisfy buoyancy (which in turn means that C^μ cannot be convex). This conclusion is a consequence of another simple to test necessary condition. Both facts are described in the next result:

Proposition 4.6 Let μ be a discrete capacity on N such that C^μ is strongly delay-averse. Then

(a) For each $i, j \in N$ with $i < j$, $\mu(N\setminus\{i\}) + \mu(\{j\}) < 1$.

(b) μ is not buoyant, and C^{μ} is not convex.

Proof To prove claim (a), we observe that strong delay-aversion in particular implies

$$1 = \mu(N) = C^{\mu}(1,\dots,1) > C^{\mu}((1,\dots,1) - \mathbf{e}_i + \mathbf{e}_j) = \mu(N \setminus \{i\}) + \mu(\{j\}).$$

Now claim (b) follows from the fact that that the latter inequality can be written as $\mu(N) - \mu(N \backslash \{i\}) > \mu(\{j\}) - \mu(\emptyset)$, i.e., $\Delta_i \mu(N) > \Delta_j \mu(\{j\})$, which contradicts buoyancy. Proposition 2.3 assures that C^{μ} is not convex. $\square$

In many cases, these conditions allow us to dismiss the idea that a discrete capacity generates a strongly delay-averse Choquet integral by inspection, in the same way as in Example 4.2 but with a much larger set of restrictions. Now we consider a simple situation to illustrate this possibility:

Example 4.3 With $T = 3$, Proposition 4.4 (a) imposes conditions such as (we drop brackets for simplicity) $\mu(1) > \frac{\mu(1)+\mu(1,2)+\mu(1,2,3)}{6} = \frac{\mu(1)+\mu(1,2)+1}{6}$, $\mu(1) > \frac{\mu(1)+\mu(1,3)+\mu(1,2,3)}{6}$, $\mu(1) > \frac{\mu(1,3)+\mu(1,2,3)}{5}$ and $\mu(1) > \frac{\mu(1)+\mu(1,2,3)}{4} = \frac{\mu(1)+1}{4} \geqslant 0.25$.

Conversely, the following restrictions apply too due to Proposition 4.4 (b): $\mu(3) < \frac{\mu(1)+\mu(2,3)+\mu(1,2,3)}{6}$, $\mu(3) < \frac{\mu(2,3)+\mu(1,2,3)}{5}$, $\mu(2) < \frac{\mu(2)+\mu(1,2)}{3}$.

Let the discrete capacity μ on $N = \{1, 2, 3\}$ be defined as follows (we drop brackets for simplicity):

$$\mu(1) = 0.5,\ \mu(2) = 0.4,\ \mu(3) = 0.3,\ \mu(1,2) = 0.9,\ \mu(1,3) = 0.8,\ \mu(2,3) = 0.7.$$

that we summarize by

$$\begin{array}{ccc} & 1 & \\ 0.9 & 0.8 & 0.7 \\ 0.5 & 0.4 & 0.3 \\ & 0 & \end{array}$$

By inspection, one can check that $\mu(\{1\}) > \mu(\{2\}) > \mu(\{3\})$. In addition, μ is submodular.

However the property and $\mu(N \setminus \{1\}) + \mu(\{2\}) < 1$ does not hold true. We conclude that μ defines a discrete Choquet integral $C^{\mu} : \mathbb{R}^3_+ \longrightarrow \mathbb{R}_+$ that is not strongly delay-averse. We can check this failure because $1 = C^{\mu}(1, 1, 1) < C^{\mu}(0, 2, 1) = 1.1$, a contradiction with delay-aversion.

4.2 The production of strongly delay-averse Choquet integrals with submodular capacities

Strongly delay-averse Choquet integrals can never be produced from buoyant capacities, by Proposition 4.6. It is known that buoyancy implies submodularity (v., Sect. 2.2). Therefore it cannot be true that submodularity of a capacity suffices to produce strongly delay-averse Choquet integrals, which is further shown by Example 4.3. Our next theorem is remarkable because it establishes that submodularity in conjunction with a necessary condition for strong delay-aversion obtained in Proposition 4.6, guarantee that the Choquet integral is strongly delay-averse. We prove it in Appendix C. Then we put forward an equivalent statement for this theorem (cf., Corollary 4.1).

Theorem 4.1 Suppose that the discrete capacity μ on N is submodular and $\mu(N \setminus \{i\}) + \mu(\{j\}) < 1$, for each $i, j \in N$ with $i < j$.

Then the Choquet integral $C^\mu : \mathbb{R}_+^T \longrightarrow \mathbb{R}_+$ is both strongly delay-averse and strongly c-monotonic.

We conclude that the condition $\mu(N\setminus\{i\}) + \mu(\{j\}) < 1$, for each $i, j \in N$ with $i < j$ must be incompatible with buoyancy of μ. Note that in fact, $\mu(N\setminus\{i\}) + \mu(\{j\}) < 1$ is equivalent to $\Delta_j\mu(\{j\}) < \Delta_i\mu(N)$ with $i < j$, which contradicts buoyancy.

Furthermore, we can express the assumptions of Theorem 4.1 differently. The reason is that when a discrete capacity μ on N is submodular, the condition $\mu(N\setminus\{i\}) + \mu(\{j\}) < 1$ when $i, j \in N$ and $i < j$ is equivalent to the combination of $\mu(\{i\}) > \mu(\{j\})$ when $i, j \in N$ and $i < j$, plus $\mu(N\setminus\{i\}) + \mu(\{i+1\}) < 1$ for each $i = 1, \ldots, T-1$. It is obvious that the latter two properties imply $\mu(N\setminus\{i\}) + \mu(\{j\}) < 1$, for each $i, j \in N$ with $i < j$. Conversely, suppose μ is submodular, and fix $i, j \in N$ with $i < j$. Assuming $\mu(N \setminus \{i\}) + \mu(\{j\}) < 1$,

$$\begin{aligned}\mu(N \setminus \{i\}) + \mu(\{j\}) < 1 =& \mu((N \setminus \{i\}) \cup \{i\}) + \mu((N \setminus \{i\}) \cap \{i\}) \\ \leqslant& \mu(N \setminus \{i\}) + \mu(\{i\})\end{aligned}$$

due to submodularity. This inequality forcefully yields $\mu(\{j\}) < \mu(\{i\})$.

Therefore the statement of Theorem 4.1 can be replaced by the following equivalent form:

Corollary 4.1 Suppose that the discrete capacity μ on N is submodular, $\mu(\{i\}) > \mu(\{j\})$ for each $i, j \in N$ with $i < j$, and $\mu(N\setminus\{i\}) + \mu(\{i+1\}) < 1$ for each $i = 1, \ldots, T-1$. Then the Choquet integral $C^\mu : \mathbb{R}_+^T \longrightarrow \mathbb{R}_+$ is both strongly delay-averse and strongly c-monotonic.

The next example shows that even in the case of λ-fuzzy measures, the statement of Corollary 4.1 cannot dispense with the necessary condition prescribed by Proposition 4.6. Nevertheless, Example 4.5 below proves that λ-fuzzy measures exist that satisfy the conditions in Theorem 4.1 or Corollary 4.1.

Example 4.4 Consider the discrete capacity μ on $N = \{1, 2, 3\}$ defined as follows:

$$\begin{array}{ccc} & 1 & \\ 0.9285 & 0.8523 & 0.7877 \\ 0.7 & 0.6 & 0.4 \\ & 0 & \end{array}$$

This is a λ-fuzzy measure with $\lambda = -0.8845$ (Beliakov et al., 2020, Example 2.9) hence it is submodular (v. Sect. 2.2).

Although $\mu(\{1\}) > \mu(\{2\}) > \mu(\{3\})$ and μ is submodular, the properties $\mu(N\backslash\{1\}) + \mu(\{2\}) < 1$ and $\mu(N\backslash\{2\}) + \mu(\{3\}) < 1$ do not hold true. Therefore μ defines a discrete Choquet integral $C^\mu : \mathbb{R}^3_+ \longrightarrow \mathbb{R}_+$ that cannot be strongly delay-averse, by Proposition 4.6. For example, a contradiction with this property is $C^\mu(4, 1, 6) = 4.3569 < 4.6092 = C^\mu(4, 0, 7)$.

Example 4.5 Consider the λ-fuzzy measure on $N = \{1, 2, 3\}$ defined by

$$\mu(\{1\}) = 0.5,\ \mu(\{2\}) = 0.4,\ \mu(\{3\}) = 0.2.$$

We use (5) to compute the value of λ that produces a normalized capacity, and we find $\lambda = -0.270882$, which means that the discrete capacity produced from these data is submodular (v. Sect. 2.2). Therefore

$$\mu(\{1, 2\}) = 0.845824,\ \mu(\{1, 3\}) = 0.672912,\ \mu(\{2, 3\}) = 0.578329.$$

So to summarize, μ is the discrete capacity represented by

$$\begin{array}{ccc} & 1 & \\ 0.845824 & 0.672912 & 0.578329 \\ 0.5 & 0.4 & 0.2 \\ & 0 & \end{array}$$

From inspection, one can check that $\mu(\{1\}) > \mu(\{2\}) > \mu(\{3\})$. The properties $\mu(N \setminus \{1\}) + \mu(\{2\}) < 1$ and $\mu(N \setminus \{2\}) + \mu(\{3\}) < 1$ hold true too.

Then μ defines a discrete Choquet integral $C^\mu : \mathbb{R}^3_+ \longrightarrow \mathbb{R}_+$ that is strongly delay-averse, by Corollary 4.1. However μ is not buoyant because $\mu(\{2\}) < \mu(\{1, 2\}) - \mu(\{2\})$, a fact prescribed by Proposition 4.6.

4.3 Characterization of strongly delay-averse Choquet integrals

It could be argued that the approach in Sect. 4.2 is based on rejecting convexity-type properties of the Choquet integral, due to Proposition 4.6. Now we take a different path that will lead us to the solution of our main problem.

Let us define temporal-antibuoyant capacities as follows:

Definition 4.1 The capacity μ on N is **temporal-antibuoyant** when for all $i, j \in N$ with $j > i$:

(T.1) $\Delta_i\mu(A) > \Delta_j\mu(A)$ for all $A \subseteq N$ with $i, j \in A$, and
(T.2) $\Delta_i\mu(A \cup \{i, j\}) > \Delta_j\mu(B \cup \{j\})$ for all $B \subseteq A \subseteq N$ with $i, j \notin B$.

We can express this axiom in other different manners because:

(T.1) can be written as $\mu(A\backslash\{i\}) < \mu(A\backslash\{j\})$ for all $A \subseteq N$ with $i,j \in A$, $j > i$ (cf., Sect. 4.1).
(T.2) can be written as $\Delta_i \mu(A) > \Delta_j \mu(B)$ for all $B \subsetneq A \subseteq N$ with $i \in A\backslash B$, $j \in B$, $j > i$.

4.3.1 Semantic interpretation

Antibuoyant capacities are defined by possessing increasing discrete derivatives with respect to set cardinality. They are the exact class of capacities with respect to which the Choquet integral satisfies the Pigou–Dalton principle (Beliakov & James, 2021). More specifically, OWA operators defined from antibuoyant weighting vectors are justified by their connection with welfare economics through the Pigou–Dalton progressive transfers principle (Aristondo et al., 2013). Our main goal is to prove a comparable result for the evaluation of temporal rewards, namely, Theorem 4.2 below, with the help of Definition 4.1. The semantic description of this property is less transparent than that of antibuoyancy, but nevertheless, compatible with intuitions underlying the foundations of intertemporal discounting:

(T.1) means that at any set of periods (which in our description is $A \setminus \{i,j\}$), the derivative with respect to an earlier period i is strictly higher than the derivative with respect to a later period j. Or, adding one period i that comes earlier than another period j gives a higher evaluation than the evaluation at the initial set of periods.
(T.2) is a more purely "temporal" version of antibuoyancy: it states that discrete derivatives are strictly increasing with respect to set cardinality, *when the derivative at the larger set is taken with respect to an earlier moment*.

4.3.2 Example and main result

Below is an example with temporal-antibuoyant capacities.

Example 4.6 With $T = 3$, from inspection we can check that the following two discrete capacities are temporal-antibuoyant:

$$\begin{array}{ccc} & 1 & \\ 0.5 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0 \\ & 0 & \end{array} \qquad\qquad \begin{array}{ccc} & 1 & \\ \frac{3}{4} & \frac{1}{2} & \frac{13}{32} \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ & 0 & \end{array}$$

For further use, in the conditions of Definition 4.1 we associate with μ the following strictly positive figures: for all $i,j \in N$ with $j > i$,

$$\delta^{\mu}_{ij} = \min\{\Delta_i \mu(A) - \Delta_j \mu(A) \mid i,j \in A \subseteq N\} > 0, \tag{15a}$$

$$\epsilon_{ij}^{\mu} = \min\{\Delta_i \mu(A \cup \{i,j\}) - \Delta_j \mu(B \cup \{j\} \mid B \subseteq A \subseteq N, i,j \notin B\} > 0, \tag{15b}$$

and also

$$\delta^{\mu} = \min\{\delta_{ij}^{\mu} \mid j > i\} > 0, \tag{16a}$$

$$\epsilon^{\mu} = \min\{\epsilon_{ij}^{\mu} \mid j > i\} > 0, \text{ and} \tag{16b}$$

$$\rho^{\mu} = \min\{\delta^{\mu}, \epsilon^{\mu}\} > 0. \tag{16c}$$

Obviously, temporal-antibuoyancy does not imply antibuoyancy. Temporal-antibuoyant capacities satisfy Eq. (9), i.e., $\mu(\{i\}) > \mu(\{j\})$ when $j > i$, using part (T.1) with $A = \{i,j\}$. Note that the antibuoyant capacity defined in Example 2.2 is not temporal-antibuoyant because $\mu(\{2\}) > \mu(\{3\})$ is false.

The main goal of this article is to prove the following result:

Theorem 4.2 A capacity μ on N produces a Choquet integral $C^{\mu} : \mathbb{R}_+^T \longrightarrow \mathbb{R}_+$ that is both strongly delay-averse and strongly c-monotonic if and only if it is temporal-antibuoyant.

Theorem 4.2 (the proof of which is in Appendix C) identifies the exact class of capacities that generate Choquet aggregation operators with both strong delay-aversion and strong c-monotonicity, which are equivalent properties by Proposition 2.1.

5 Conclusions and future research lines

In a framework of intertemporal choice (which is here regarded as a finite and discrete process in time), we have posed and investigated the discrete Choquet integrals that satisfy strong aversion to delayed rewards. This property has been defined and related to other axioms. Robust arguments testify to its validity as a minimal requirement for reasonable evaluations in this framework. We have demonstrated that neither convex Choquet integrals, nor those that respect the Pigou–Dalton progressive transfers principle, produce strongly delay-averse evaluations of the non-negative vectors representing the successive rewards. Finally, we have identified a new property of a capacity that is equivalent to this axiom of delay aversion for the Choquet integral that it defines.

Several tasks remain for further work.

The discussion in Sect. 2.2 recalled that the characterization of the capacities yielding Choquet integrals that satisfy the Pigou–Dalton principle remains open. For inspiration, the logically weaker version called the Pigou–Dalton *progressive* principle is characterized by antibuoyancy. But we gave an example proving that antibuoyant capacities do not necessarily produce Choquet integrals that satisfy the Pigou–Dalton principle.

A problem that is more aligned with the contribution of this paper is the identification of the exact class of Sugeno capacities that produce strongly delay-averse Choquet integrals. An example in a setting with 3 periods has proven that this class is not empty. The same question can be posed for 2-additive, or in general, k-additive capacities (Grabisch, 1997).

Another line for future inspection is the random generation of temporal-antibuoyant capacities, a topic that aligns with the results of e.g., Beliakov (2022) for supermodular capacities, or Beliakov et al. (2022) for k-interactive capacities.

Learning is feasible as well. Note that Theorem 4.2 provides a set of linear constraints on the values of a capacity, which similarly to the case of buoyant/antibuoyant capacities, can be coded algorithmically. Existing techniques on random generation/learning suitable capacities can be adapted with relatively few modifications.

Appendix A Example and proof of result of Section 2.1

Example A.1 In relation with the assumptions of claim 4 in Proposition 2.1, we note that they are tight for the following reasons.

(a) There are non-transitive, monotonic and delay-averse relations on X^T that are not c-monotonic.
(b) There are non-monotonic, transitive and delay-averse relations on X^T that are not c-monotonic.

To prove (a), consider $X = \mathbb{R}_+$ and the relation R_a on X^T defined by $\mathbf{x}R_a\mathbf{y}$ if and only if the number of non-negative components of $\mathbf{x}-\mathbf{y}$ is at least as high as the number of its non-positive components. Then R_a is not transitive because in the case $T = 2$, $(1,1)R_a(0,3)R_a(2,2)$ but $(1,1)R_a(2,2)$ is false. This relation is obviously monotonic. It is delay-averse because if $i \in \{1,\dots,T-1\}$, $h > 0$, and $\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1} \in X^T$, then $\mathbf{x}R_a(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1})$ is true. The reason is that the vector $\mathbf{x} - (\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1}) = h\mathbf{e}_i - h\mathbf{e}_{i+1}$ has exactly one positive and one negative component. To see that it is not c-monotonic, note that when $T = 3$ one has $(3,0,0) \succcurlyeq_c (1,1,1)$ but $(3,0,0)R_a(1,1,1)$ is false.

To prove (b), consider $X = \mathbb{R}_+$ and the relation R_b on X^T defined by $\mathbf{x}R_b\mathbf{y}$ if and only if $\sum_{i=1}^{T-1} x_i \geqslant \sum_{i=1}^{T-1} y_i$ and $x_T \leqslant y_T$. It is not monotonic because $(1,7)R_b(1,1)$ is false. Transitivity is straightforward. To prove that R_b is delay-averse, we argue $\mathbf{x}R_b\mathbf{y} = \mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1}$ separately for $i < T-1$ and $i = T-1$. In the first case, because $\sum_{i=1}^{T-1} x_i = \sum_{i=1}^{T-1} y_i$ and $x_T = y_T$. In the second, because $\sum_{i=1}^{T-1} x_i > \sum_{i=1}^{T-1} y_i = \sum_{i=1}^{T-1} x_i - h$ and $x_T < y_T + h$. Finally, R_b is not c-monotonic because $(1,7) \succcurlyeq_c (1,1)$ but $(1,7)R_b(1,1)$ is false.

In relation with claim 5 in Proposition 2.1, we note that the following non-transitive, non-monotonic binary relation on X^T is strongly delay-averse but not c-monotonic: $\mathbf{x}R_c\mathbf{y}$ if and only if there is $j > i$ with $x_i > y_i$ and $x_j < y_j$. It is strongly delay-averse because if $i \in \{1,\dots,T-1\}$, $h > 0$, and $\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1} \in X^T$, then clearly $\mathbf{x}R_c(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1})$ is true but $(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1})R_c\mathbf{x}$ does not hold. To see that it is not c-monotonic, note that when $T = 2$ one has $(1, 1) \succcurlyeq_c (0, 0)$ but $(1, 1)R_c(0, 0)$ is false.

Proof of Proposition 2.1 Lemma 2.1 and transitivity guarantee the equivalence in claim 1. We explained that when R satisfies the Pigou–Dalton principle, it satisfies the Pigou–Dalton progressive principle as well (because progressive Pigou–Dalton transfers are Pigou–Dalton transfers).

To prove claim 2, suppose first that R is consistent with cumulative domination (also called c-monotonicity). Fix $\epsilon \succcurlyeq_c 0$, $\epsilon \subseteq X^T$. Then obviously $(\mathbf{x} + \epsilon) \succcurlyeq_c \mathbf{x}$ when $x \in X^T$, therefore $(\mathbf{x} + \epsilon)R\mathbf{x}$. Conversely, if $(\mathbf{x} + \epsilon)R\mathbf{x}$ whenever $\epsilon \succcurlyeq_c 0$ and $x \in X^T$, then to prove that R is c-monotonic we fix $\mathbf{x}, \mathbf{y} \in X^T$ such that $\mathbf{x} \succcurlyeq_c \mathbf{y}$. Observe that $\mathbf{x} = \mathbf{y} + (\mathbf{x} - \mathbf{y})$ and $\mathbf{x} - \mathbf{y} \succcurlyeq_c 0$ to deduce the conclusion.

A simple modification proves that R is strongly c-monotonic if and only if for each $\epsilon \succcurlyeq_c 0$ with $\epsilon \neq 0$, and each $x \in X^T$, one has $(\mathbf{x} + \epsilon)P\mathbf{x}$.

Furthermore, the case $\mathbf{x} = (0, \dots, 0)$ proves that when R is c-monotonic then R satisfies Axiom Z. And we stated earlier that c-monotonicity implies monotonicity.

Now claim 3 is straightforward: for each $\epsilon \subseteq X^T$ with $\epsilon \succcurlyeq_c 0$, Axiom Z guarantees $\epsilon R0$, thus $(\mathbf{x} + \epsilon)R\mathbf{x}$ because $\mathbf{x}R\mathbf{x}$ (by reflexivity) and R is additive. Claim 2 yields the conclusion.

To prove claim 4, we proceed in two steps.

First, note that $h\mathbf{e}_i - h\mathbf{e}_{i+1} \succcurlyeq_c 0$, and $\mathbf{x} = (\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1}) + (h\mathbf{e}_i - h\mathbf{e}_{i+1})$. Owing to claim 2, c-monotonicity yields the conclusion of the first statement. Example A.1 below proves that the reverse is not true.

Now suppose that R is reflexive, monotonic, and delay-averse. To prove c-monotonicity, fix $\mathbf{x}, \mathbf{y} \in X^T$ with $\mathbf{x} \succcurlyeq_c \mathbf{y}$. When $\mathbf{x} = \mathbf{y}$, we conclude by reflexivity. Otherwise we use a recursive argument.

Because $\mathbf{x} \succcurlyeq_c \mathbf{y}$ implies $x_1 - y_1 \geqslant 0$, we can assure $\mathbf{x}R(y_1, x_2 + x_1 - y_1, x_3, \dots, x_T)$, both if $x_1 = y_1$ or $x_1 > y_1$. The first case is trivial by reflexivity, the second is justified by delay-aversion because $(y_1, x_2 + x_1 - y_1, x_3, \dots, x_T)$ follows from a $(x_1 - y_1)$ -delay from the first to the second components.

Similarly, using that $\mathbf{x} \succcurlyeq_c \mathbf{y}$ implies $x_1 + x_2 - (y_1 + y_2) \geqslant 0$,

$$(y_1, x_2 + x_1 - y_1, x_3, \dots, x_T)R(y_1, y_2, x_3 + (x_1 + x_2 - (y_1 + y_2)), x_4, \dots, x_T)$$

because we obtain $(y_1, y_2, x_3 + (x_1 + x_2 - (y_1 + y_2)), x_4, \dots, x_T)$ from the vector $(y_1, x_2 + x_1 - y_1, x_3, \dots, x_T)$ shifting $x_1 + x_2 - (y_1 + y_2) \geqslant 0$ from the second to the third component. Again, we use either reflexivity or delay-aversion.

A recursive argument and transitivity of R produces

$$\mathbf{x}R\left(y_1, \dots, y_{T-1}, x_T + \left(\sum_{i=1}^{T-1}(x_i - y_i)\right)\right)$$

Using that $\mathbf{x} \succcurlyeq_c \mathbf{y}$ implies $\sum_{i=1}^{T}(x_i - y_i) \geqslant 0$, hence $x_T + \sum_{i=1}^{T-1}(x_i - y_i) \geqslant y_T$, and the fact that R is monotonic, we get

$$\mathbf{x}R\left(y_1, \ldots, y_{T-1}, x_T + \left(\sum_{i=1}^{T-1}(x_i - y_i)\right)\right) R\,\mathbf{y}.$$

The conclusion follows from transitivity.

We can prove claim 5 with arguments that replicate those of claim 4. □

Appendix B Proofs of results of Section 3

To prove Proposition 3.1, we will avail ourselves of the next auxiliary result:

Lemma B.1 Suppose $x_1 \leqslant \ldots \leqslant x_T$, $x_1, \ldots, x_T \in \mathbb{R}^T$. Then

$$(x_T, \ldots, x_1) \succcurlyeq_c (M(x_1, \ldots, x_T), \ldots, M(x_1, \ldots, x_T)) \succcurlyeq_c (x_1, \ldots, x_T).$$

Proof For each $k = 1, \ldots, T$, we can distribute

$$\begin{aligned} kM(x_1, \ldots, x_T) = & \frac{x_T + \ldots \overset{k}{} \ldots + x_T + x_{T-k} + x_{T-k-1} + \cdots + x_1}{T} \\ & + \frac{x_{T-1} + \ldots \overset{k}{} \ldots + x_{T-1} + x_{T-k} + x_{T-k-1} + \ldots + x_1}{T} + \cdots \cdots \\ & + \frac{x_{T-k+1} + \ldots \overset{k}{} \ldots + x_{T-k+1} + x_{T-k} + x_{T-k-1} + \ldots + x_1}{T} \end{aligned}$$

Therefore due to the assumption $x_1 \leqslant \ldots \leqslant x_T$,

$$kM(x_1, \ldots, x_T) \leqslant x_T + \ldots + x_{T-k+1}.$$

We conclude $\frac{x_T + \ldots + x_{T-k+1}}{k} \geqslant M(x_1, \ldots, x_T)$ for each $k = 1, \ldots, T$, which accounts for $(x_T, \ldots, x_1) \succcurlyeq_c (M(x_1, \ldots, x_T), \ldots, M(x_1, \ldots, x_T))$.

We can show $(M(x_1, \ldots, x_T), \ldots, M(x_1, \ldots, x_T)) \succcurlyeq_c (x_1, \ldots, x_T)$ in a similar way, using that for each $k = 1, \ldots, T$, we can distribute

$$\begin{aligned} kM(x_1, \ldots, x_T) = & \frac{x_1 + \ldots \overset{k}{} \ldots + x_1 + x_{k+1} + x_{k+2} + \ldots + x_T}{T} \\ & + \frac{x_2 + \ldots \overset{k}{} \ldots + x_2 + x_{k+1} + x_{k+2} + \ldots + x_T}{T} + \cdots \cdots \\ & + \frac{x_k + \ldots \overset{k}{} \ldots + x_k + x_{k+1} + x_{k+2} + \ldots + x_T}{T} \geqslant x_1 + \ldots + x_k \end{aligned}$$

□

Proof of Proposition 3.1 Suppose $\mathbf{x} = (x_1, \ldots, x_T) \in \mathbb{R}^T_+$. We do not lose generality if we assume $x_1 \leqslant \ldots \leqslant x_T$ in order to prove $f(\mathbf{x}) = M(\mathbf{x})$, because

$f(\mathbf{x}) = f(x_{(1)}, \dots, x_{(T)})$ by symmetry, and $M(\mathbf{x}) = M(x_{(1)}, \dots, x_{(T)})$. Lemma B.1 and symmetry prove

$$f(\mathbf{x}) = f(x_T, \dots, x_1) \geqslant f(M(x_1, \dots, x_T), \dots, M(x_1, \dots, x_T))$$

(as $(x_T, \dots, x_1) \succcurlyeq_c (M(x_1, \dots, x_T), \dots, M(x_1, \dots, x_T))$, and f is c-monotonic). Similarly, the facts $(M(x_1, \dots, x_T), \dots, M(x_1, \dots, x_T)) \succcurlyeq_c (x_1, \dots, x_T)$ and c-monotonicity prove

$$f(M(x_1, \dots, x_T), \dots, M(x_1, \dots, x_T)) \geqslant f(x_1, \dots, x_T) = f(\mathbf{x})$$

using that f is symmetric. This concludes the proof, since we have shown

$$f(\mathbf{x}) = f(M(x_1, \dots, x_T), \dots, M(x_1, \dots, x_T)) = M(x_1, \dots, x_T),$$

because f is idempotent. □

Appendix C Proofs of results of Section 4

To prove Proposition 4.4, we avail ourselves of the next auxiliary result:

Lemma C.1 In the conditions of statement (a), respectively, statement (b), of Proposition 4.4: $C^\mu(c \cdot \mathbf{e}_i) > C^\mu(\sum_{j=1}^k \chi_{A_j})$, respectively, $C^\mu(c \cdot \mathbf{e}_i) < C^\mu(\sum_{j=1}^k \chi_{A_j})$.

Proof Let $\mathbf{x} = (x_1, \dots, x_T) = \sum_{j=1}^k \chi_{A_j} \in \mathbb{R}_+^T$, then $x_t = |\{j \in N | t \in A_j\}|$. The assumption in (a) assures $x_t = 0$ when $t < i$. Also, $c > x_i$, even when $k = 1$ (because $A_k \neq \{i\}$). Therefore one can make a finite sequence of 1-delays on $c \cdot \mathbf{e}_i$ that produce $\mathbf{x}$ because $i \leqslant i'$ for all $i' \in A_k$. Using that C^μ is strongly delay-averse, the conclusion $C^\mu(c \cdot \mathbf{e}_i) > C^\mu(\sum_{j=1}^k \chi_{A_j})$ follows.

Statement (b) can be proven in a similar manner. □

Proof of Proposition 4.4 We only need to apply Lemmata C.1 and 2.3 to prove claim (a). □

Proof of Theorem 4.1 We use the fact that C^μ is convex (by Proposition 2.2) and subadditive (by Lemma 2.2).

To prove that C^μ is strongly delay-averse, fix $i \in \{1, \dots, T-1\}$ and $h > 0$ such that $\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1} \in \mathbb{R}_+^N$. Define $\mathbf{h} = (h, \dots, h)$, then $\mathbf{h} - h\mathbf{e}_i + h\mathbf{e}_{i+1} \in \mathbb{R}_+^N$ and $C^\mu(\mathbf{h} - h\mathbf{e}_i + h\mathbf{e}_{i+1}) = h \cdot \mu(N \setminus \{i\}) + h \cdot \mu(\{i+1\})$. Therefore $C^\mu(\mathbf{h} - h\mathbf{e}_i + h\mathbf{e}_{i+1}) < h$ by assumption.

Using translation-invariance and then additivity,

$$
\begin{aligned}
C^{\mu}(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_{i+1}) &= C^{\mu}(\mathbf{x} + \mathbf{h} - h\mathbf{e}_i + h\mathbf{e}_{i+1} - \mathbf{h}) \\
&= C^{\mu}(\mathbf{x} + \mathbf{h} - h\mathbf{e}_i + h\mathbf{e}_{i+1}) - h \\
&\leqslant C^{\mu}(\mathbf{x}) + C^{\mu}(\mathbf{h} - h\mathbf{e}_i + h\mathbf{e}_{i+1}) - h < C^{\mu}(\mathbf{x}).
\end{aligned}
$$

□

We need some preparatory results to prove Theorem 4.2.

Lemma C.2 Suppose $\mathbf{x} = (x_1, \dots, x_T) \in \mathbb{R}^T$ and $x_i = x_j + 1$. Let σ be a permutation of indices such that $\mathbf{x}_{\nearrow} = \mathbf{x}_{\sigma}$ satisfies $\sigma(t) = i$, $\sigma(t') = j$ (therefore $t > t'$). If τ is the transposition of i and j, then $\mathbf{x}'_{\nearrow} = \mathbf{x}'_{\tau\circ\sigma}$ is a non-decreasing ordering of $\mathbf{x}' = \mathbf{x} - \mathbf{e}_i + \mathbf{e}_j$. Also, $\mathbf{x}_{\nearrow} = \mathbf{x}'_{\nearrow}$.

Proof It is obvious that $\mathbf{x}_{\nearrow} = \mathbf{x}'_{\nearrow}$ in this case. Note that $x'_i = x_i - 1 = x_j$, $x'_j = x_j + 1 = x_i$. Therefore an ordering that permutes these indices after any permutation producing $\mathbf{x}_{\nearrow}$ guarantees a non-decreasing ordering of $\mathbf{x}'$. □

Lemma C.3 For each $\mathbf{x} = (x_1, \dots, x_T) \in \mathbb{N}^T$, the vectors $\mathbf{x}$ and $\mathbf{x}' = \mathbf{x} - \mathbf{e}_i + \mathbf{e}_j$ are comonotonic except when $x_i = x_j + 1$.

Proof We use the standard characterization of comonotonicity (Beliakov et al., 2020, page 94): $\mathbf{x}$ and $\mathbf{x}'$ are comonotonic if and only if $(x_p - x_q)(x'_p - x'_q) \geqslant 0$ for all $\{p,q\} \subseteq N$ with $p \neq q$.

When $\{p,q\} \cap \{i,j\} = \varnothing$, $x_p = x'_p$ and $x_q = x'_q$ thus $(x_p - x_q)(x'_p - x'_q) = 0$.

When $\{p,q\} \cap \{i,j\} = \{i\}$, without loss of generality $x_q = x'_q$ and $x_p = x_i$, $x'_p = x_i - 1$ thus $(x_p - x_q)(x'_p - x'_q) = (x_i - x_q)(x_i - 1 - x_q)$. If $x_i \leqslant x_q + 1$, then the inequality $(x_i - x_q)(x_i - 1 - x_q) \geqslant 0$ follows easily. Otherwise $x_i \geqslant x_q + 2$, then $x_i - x_q > 0$ and $x_i - 1 - x_q > 0$, proving the claim.

When $\{p,q\} \cap \{i,j\} = \{j\}$, without loss of generality $x_q = x'_q$ and $x_p = x_j$, $x'_p = x_j + 1$ thus $(x_p - x_q)(x'_p - x'_q) = (x_j - x_q)(x_j + 1 - x_q)$. If $x_j - x_q \geqslant 0$, then the inequality $(x_j - x_q)(x_j + 1 - x_q) \geqslant 0$ follows easily. Otherwise $x_j - x_q \leqslant -1$, which also yields $(x_j - x_q)(x_j + 1 - x_q) \geqslant 0$.

When $\{p,q\} \cap \{i,j\} = \{i,j\}$, without loss of generality $x_q = x_i$ and $x_p = x_j$, then $x'_q = x_i - 1$ and $x'_p = x_j + 1$ thus $(x_p - x_q)(x'_p - x'_q) = (x_j - x_i)(x_j + 1 - (x_i - 1))$. The inequality $(x_j - x_i)(x_j - x_i + 2)) \geqslant 0$ follows easily when $x_j \geqslant x_i$. Otherwise because $x_i = x_j + 1$ cannot happen by assumption, one has $x_i > x_j + 1$ therefore the conclusion $(x_j - x_i)(x_j - x_i + 2) \geqslant 0$ follows too. □

Note that the restriction $x_i = x_j + 1$ in Lemma C.3 applies to cases such as $\mathbf{x} = (1, 0)$, $\mathbf{x} - \mathbf{e}_1 + \mathbf{e}_2 = (0, 1)$, which are not comonotonic. Lemma C.2 considers this type of situations, even if the components of $\mathbf{x}$ are not integers.

Now we are ready to prove our main result.

Proof of Theorem 4.2 Propositions 4.3 and 4.5 prove necessity.

Let us now prove sufficiency. We fix μ that is temporal-antibuoyant. To prove that C^μ is strongly delay-averse, we proceed in several steps.

Initially, we demonstrate that the conclusion follows when $\mathbf{x} = (x_1, \ldots, x_T) \in \mathbb{N}^T$ and for unit delays only. Define $\mathbf{x}' = \mathbf{x} - \mathbf{e}_i + \mathbf{e}_j$ with $j > i$. We distinguish two cases.

Case 1) $x_i = x_j + 1$ is false.

Since $\mathbf{x} - \mathbf{e}_i + \mathbf{e}_j$ and $\mathbf{x}$ are comonotonic by Lemma C.3, we can order $\mathbf{x}_\nearrow = (x_{(1)}, \ldots, x_{(T)})$ in such way that $\mathbf{x}'_\nearrow = (x'_{(1)}, \ldots, x'_{(T)})$. If we set $x_{(0)} = x'_{(0)} = 0$ and $H_l = \{(l), \ldots, (T)\}$ for all l, Eq. (4) produces

$$\begin{aligned} C^\mu(\mathbf{x} - \mathbf{e}_i + \mathbf{e}_j) - C^\mu(\mathbf{x}) &= \sum_{l=1}^{T} \Delta_{(l)}\mu(H_l)x'_{(l)} - \sum_{l=1}^{T} \Delta_{(l)}\mu(H_l)x_{(l)} \\ &= \Delta_{(t)}\mu(H_t)(x'_{(t)} - x_{(t)}) + \Delta_{(t')}\mu(H_{t'})(x'_{(t')} - x_{(t')}) \end{aligned}$$

provided that $(t) = i$, $(t') = j$. Therefore

$$C^\mu(\mathbf{x} - \mathbf{e}_i + \mathbf{e}_j) - C^\mu(\mathbf{x}) = \Delta_{(t')}\mu(H_{t'}) - \Delta_{(t)}\mu(H_t). \tag{C.1}$$

In conclusion, $C^\mu(\mathbf{x} - \mathbf{e}_i + \mathbf{e}_j) - C^\mu(\mathbf{x}) < 0$ if and only if $\Delta_{(t)}\mu(H_t) > \Delta_{(t')}\mu(H_{t'})$. This inequality holds due to temporal-antibuoyancy, noting that $j = (t') > (t) = i$ and $H_{t'} \subsetneq H_t$ ($i, j \in H_t$, $i \notin H_{t'}$). We use (T.2) with $A = H_t$, $B = H_{t'} \setminus \{j\}$.

For further use, we note that using the figures defined in Equations (15) and (16) we can establish

$$C^\mu(\mathbf{x}) - C^\mu(\mathbf{x} - \mathbf{e}_i + \mathbf{e}_j) \geqslant \epsilon^\mu_{ij} \geqslant \epsilon^\mu \geqslant \rho^\mu > 0.$$

Case 2) $x_i = x_j + 1$.

As $\mathbf{x} = (x_1, \ldots, x_T) \in \mathbb{N}^T$, we can fix a permutation σ such that $\mathbf{x}_\sigma = \mathbf{x}_\nearrow = (x_{(1)}, \ldots, x_{(T)})$, $\sigma(t) = j$ and $\sigma(t+1) = i$. Then by Lemma C.2, if τ is the transposition of i and j then $\mathbf{x}'_{\tau\circ\sigma}$ is a non-decreasing ordering of $\mathbf{x}'$. We note that

$$(\tau\circ\sigma)(l) = \sigma(l) \text{ when } l \in N \setminus \{t, t+1\},$$

$$(\tau\circ\sigma)(t) = \tau(j) = i = \sigma(t+1) \text{ and } (\tau\circ\sigma)(t+1) = \tau(i) = j = \sigma(t).$$

Therefore

$$\begin{aligned} x'_{(\tau\circ\sigma)(t)} &= x'_{\sigma(t+1)} = x'_i = x_i - 1 = x_j, \\ x'_{(\tau\circ\sigma)(t+1)} &= x'_{\sigma(t)} = x'_j = x_j + 1 = x_i, \text{ and} \\ x'_{(\tau\circ\sigma)(l)} &= x_{\sigma(l)} \text{ if } l \in N \setminus \{t, t+1\}. \end{aligned}$$

We keep the convention $x_{\sigma(0)} = x'_{(\tau\circ\sigma)(0)} = 0$, and denote $H_i = \{(i), \ldots, (T)\}$, $H'_l = \{(\tau\circ\sigma)(l), \ldots, (\tau\circ\sigma)(T)\}$ for all l.

Eq. (4) produces

$$\begin{aligned}C^{\mu}(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j)-C^{\mu}(\mathbf{x}) &= \sum_{l=1}^{T}\Delta_{(\tau\circ\sigma)(l)}\mu(H'_l)x'_{(\tau\circ\sigma)(l)}-\sum_{l=1}^{T}\Delta_{\sigma(l)}\mu(H_l)x_{\sigma(l)}\\ &=\sum_{l=1}^{t-1}\Big(\Delta_{\sigma(l)}\mu(H'_l)-\Delta_{\sigma(l)}\mu(H_l)\Big)x_{\sigma(l)}\\ &\quad+\Delta_{(\tau\circ\sigma)(t)}\mu(H'_t)x'_{(\tau\circ\sigma)(t)}\\ &\quad-\Delta_{\sigma(t)}\mu(H_t)x_{\sigma(t)}+\Delta_{(\tau\circ\sigma)(t+1)}\mu(H'_{t+1})x'_{(\tau\circ\sigma)(t+1)}\\ &\quad-\Delta_{\sigma(t+1)}\mu(H_{t+1})x_{\sigma(t+1)}\\ &\quad+\sum_{l=t+2}^{T}\Big(\Delta_{\sigma(l)}\mu(H'_l)-\Delta_{\sigma(l)}\mu(H_l)\Big)x_{\sigma(l)}\end{aligned}$$

and, noting $H'_l=H_l$ when $l\in N\backslash\{t,t+1\}$, both summations in this expression vanish. Therefore

$$\begin{aligned}C^{\mu}(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j)-C^{\mu}(\mathbf{x}) =&\Delta_{(\tau\circ\sigma)(t)}\mu(H'_t)x'_{(\tau\circ\sigma)(t)}-\Delta_{\sigma(t)}\mu(H_t)x_{\sigma(t)}\\ &+\Delta_{(\tau\circ\sigma)(t+1)}\mu(H'_{t+1})x'_{(\tau\circ\sigma)(t+1)}\\ &-\Delta_{\sigma(t+1)}\mu(H_{t+1})x_{\sigma(t+1)}=\\ \Big(\Delta_i\mu(H'_t)-\Delta_j\mu(H_t)\Big)x_j&+\Big(\Delta_j\mu(H'_{t+1})-\Delta_i\mu(H_{t+1})\Big)x_i.\end{aligned}$$

Note that

$$\Delta_j\mu(H'_{t+1})-\Delta_i\mu(H_{t+1})=\Delta_j\mu(H_t\setminus\{i\})-\Delta_i\mu(H_{t+1})=\Delta_j\mu(H_t)-\Delta_i\mu(H_t).$$

Substituting above, and using $H'_t=H_t$ and $x_i=x_j+1$, one has

$$\begin{aligned}C^{\mu}(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j)-C^{\mu}(\mathbf{x}) =&\Big(\Delta_i\mu(H_t)-\Delta_j\mu(H_t)\Big)(x_j-x_i)\\ =&\Delta_j\mu(H_t)-\Delta_i\mu(H_t).\end{aligned}$$

The inequality $C^{\mu}(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j)<C^{\mu}(\mathbf{x})$ holds due to property (T.1) of temporal-antibuoyancy.

For further use, we note that resorting to the figures defined in Equations (15) and (16) we can establish

$$C^{\mu}(\mathbf{x})-C^{\mu}(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j)=\Delta_i\mu(H_t)-\Delta_j\mu(H_t)\geqslant\delta^{\mu}_{ij}\geqslant\delta^{\mu}\geqslant\rho^{\mu}>0. \tag{C.2}$$

In conclusion, this first step has demonstrated that when $\mathbf{x}=(x_1,\dots,x_T)\in\mathbb{N}^T$, $C^{\mu}(\mathbf{x})-C^{\mu}(\mathbf{x}-\mathbf{e}_i+\mathbf{e}_j)\geqslant\rho^{\mu}>0$. Now a recursive argument proves that $C^{\mu}(\mathbf{x})-C^{\mu}(\mathbf{x}-h\mathbf{e}_i+h\mathbf{e}_j)\geqslant h\rho^{\mu}>0$ follows from the previous case when $\mathbf{x}=(x_1,\dots,x_T)\in\mathbb{N}^T$, $j>i$, $h\in\{1,2,\dots,x_i\}\in\mathbb{N}$, and $hx_i\geqslant 0$.

Once the conclusion is proven when $\mathbf{x}=(x_1,\dots,x_T)\in\mathbb{N}^T$ and for integer delays, we can extend the property to $\cup_{k\in\mathbb{N}\cup\{0\}}X_k$ where $X_k=\{x\in\mathbb{R}_+ \mid 10^k x\in\mathbb{N}\}$ for each $k=0,1,2,\dots$. Note that X_k is the set of rational numbers with at most k decimal places. This property follows from positive homogeneity, because in fact, in this case $\mathbf{x} \in (\cup_{k\in\mathbb{N}\cup\{0\}} X_k)^T$ and $h \in \{1, 2, \ldots, 10^k\}$ means that for a fixed k,

$$\begin{aligned} C^{\mu}(\mathbf{x}) - C^{\mu}(\mathbf{x} - h\mathbf{e}_i + h\mathbf{e}_j) =& \frac{1}{10^k}\Big(C^{\mu}(10^k\mathbf{x}) - C^{\mu}(10^k\mathbf{x} - 10^k h\mathbf{e}_i + 10^k h\mathbf{e}_j)\Big) \\ >& \frac{1}{10^k}(10^k h\rho^{\mu}) \geqslant \rho^{\mu} > 0. \end{aligned}$$

The next step proves the conclusion when $\mathbf{x} = (x_1, \ldots, x_T) \in \mathbb{R}^T$ and for integer delays. We can restrict ourselves to unit delays, as explained above. For each $k \in \mathbb{N}$ and $x \in \mathbb{R}_+$, let $\lfloor x \rfloor_k$ denote the number that truncates x to k number of digits. So $\lfloor x \rfloor_k$ keeps k decimal places after the decimal point of x's decimal expansion. Correspondingly, associated with $\mathbf{x} = (x_1, \ldots, x_T) \in \mathbb{R}^T_+$ and k we denote $\lfloor \mathbf{x} \rfloor_k = (\lfloor x_1 \rfloor_k, \ldots, \lfloor x_T \rfloor_k) \in X_k$. If $\mathbf{x}' = \mathbf{x} - \mathbf{e}_i + \mathbf{e}_j$ with $j > i$, then $\lfloor \mathbf{x}' \rfloor_k = \lfloor \mathbf{x} \rfloor_k - \mathbf{e}_i + \mathbf{e}_j \in X_k$. The previous step guarantees

$$C^{\mu}(\lfloor \mathbf{x} \rfloor_k) - C^{\mu}(\lfloor \mathbf{x}' \rfloor_k) \geqslant \rho^{\mu} > 0.$$

As this is true for all $k \in \mathbb{N}$, we get $C^{\mu}(\mathbf{x}) - C^{\mu}(\mathbf{x}') > 0$ by the continuity of the Choquet integral with respect to the Euclidean topology.

Finally, once the conclusion is proven when $\mathbf{x} = (x_1, \ldots, x_T) \in \mathbb{R}^T_+$ and for integer delays only, the general statement of strong delay-aversion follows from positive homogeneity. □

Acknowledgements The author is grateful to the Guest Editors of this Special Issue and two anonymous referees for helpful comments and suggestions that improved the manuscript considerably.

Funding Open access funding provided by FEDER European Funds and the Junta de Castilla y León under the Research and Innovation Strategy for Smart Specialization (RIS3) of Castilla y León 2021-2027.

Data availability This article does not analyse or generate any datasets.

Declarations

Conflict of interest The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Alcantud, J. C. R. (2024). Sugeno capacities for the extension of exponential and hyperbolic discounting to fuzzy integrals. *Punjab University Journal of Mathematics, 56*, 102–111.

Alcantud, J. C. R., & de Andrés Calle, R. (2017). The problem of collective identity in a fuzzy environment. *Fuzzy Sets and Systems, 315*, 57–75. https://doi.org/10.1016/j.fss.2016.02.011

Alfonsi, A. (2015). A simple proof for the convexity of the Choquet integral. *Statistics & Probability Letters, 104*, 22–25.

Anger, B. (1977). Representation of capacities. *Mathematische Annalen, 229*, 245–258. https://doi.org/10.1007/BF01391470

Aristondo, O., García-Lapresta, J. L., Lasso de la Vega, C., & Marques Pereira, R. A. (2013). Classical inequality indices, welfare and illfare functions, and the dual decomposition. *Fuzzy Sets and Systems, 228*, 114–136. https://doi.org/10.1016/j.fss.2013.02.001

Arnold, B. C. (2012). *Majorization and the Lorenz order: A brief introduction* vol. 43. Springer Science & Business Media.

Banerjee, A. K. (2014). A multidimensional Lorenz dominance relation. *Social Choice and Welfare, 42*, 171–191. https://doi.org/10.1007/s00355-014-0808-9

Beliakov, G. (2022). On random generation of supermodular capacities. *IEEE Transactions on Fuzzy Systems, 30*, 293–296. https://doi.org/10.1109/TFUZZ.2020.3036699

Beliakov, G., Cabrerizo, F. J., Herrera-Viedma, E., & Wu, J.-Z. (2022). Random generation of *k*-interactive capacities. *Fuzzy Sets and Systems, 430*, 48–55. https://doi.org/10.1016/j.fss.2020.12.010

Beliakov, G., & Divakov, D. (2020). On representation of fuzzy measures for learning Choquet and Sugeno integrals. *Knowledge-Based Systems, 189*, Article 105134. https://doi.org/10.1016/j.knosys.2019.105134

Beliakov, G., & James, S. (2021). Choquet integral optimisation with constraints and the buoyancy property for fuzzy measures. *Information Sciences, 578*, 22–36. https://doi.org/10.1016/j.ins.2021.07.032

Beliakov, G., & James, S. (2022). Choquet integral-based measures of economic welfare and species diversity. *International Journal of Intelligent Systems, 37*, 2849–2867. https://doi.org/10.1002/int.22609

Beliakov, G., James, S., & Wu, J. (2020). *Discrete fuzzy measures*. *Computational Aspects volume 382 of seriesStudies in Fuzziness and Soft Computing*. addressCham, Switzerland: Springer Nature Switzerland AG.

Beliakov, G., Pradera, A., & Calvo, T. (2007). *Aggregation Functions: A Guide for Practitioners*. *Volume 221 of seriesStudies in Fuzziness and Soft Computing*. addressHeidelberg: Springer-Verlag Berlin Heidelberg.

Buja, A. (1984). Simultaneously least favorable experiments. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 65*, 367–384. https://doi.org/10.1007/BF00533742

Choquet, G. (1954). Theory of capacities. *Annales de l'institut Fourier, 5*, 131–295.

Couceiro, M., & Marichal, J.-L. (2011). Axiomatizations of quasi-Lovász extensions of pseudo-Boolean functions. *Aequationes Mathematicae, 82*, 213–231. https://doi.org/10.1007/s00010-011-0091-0

Couceiro, M., & Marichal, J.-L. (2013). Discrete integrals based on comonotonic modularity. *Axioms, 2*, 390–403. https://doi.org/10.3390/axioms2030390

de Campos Ibáñez, L. M., & Bolaños Carmona, M. J. (1989). Representation of fuzzy measures through probabilities. *Fuzzy Sets and Systems, 31*, 23–36. https://doi.org/10.1016/0165-0114(89)90064-X

Denneberg, D. (2013). *Non-additive measure and integral*. *Volume 27 of seriesTheory and Decision Library B*. addressDordrecht: Springer Science & Business Media.

Dimuro, G. P., Fernández, J., Bedregal, B., Mesiar, R., Sanz, J. A., Lucca, G., & Bustince, H. (2020). The state-of-art of the generalizations of the Choquet integral: From aggregation and pre-aggregation to ordered directionally monotone functions. *Information Fusion, 57*, 27–43.

Even, Y., & Lehrer, E. (2014). Decomposition-integral: unifying Choquet and the concave integrals. *Economic Theory, 56*, 33–58.

Fisher, I. (1930). *The Theory of Interest* vol. 43. The Macmillan Company.

Gilboa, I. (1989). Expectation and variation in multi-period decisions. *Econometrica*, 1153–1169.

Grabisch, M. (1995). A new algorithm for identifying fuzzy measures and its application to pattern recognition. In *Proceedings of 1995 IEEE International Conference on Fuzzy Systems* (pp. 145–150 vol.1). volume 1. https://doi.org/10.1109/FUZZY.1995.409673.

Grabisch, M. (1997). *k*-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems, 92*, 167–189. https://doi.org/10.1016/S0165-0114(97)00168-1

Grabisch, M. (2016). *Set functions, games and capacities in decision making*. *volume 46 of seriesTheory and Decision Library C*. addressCham: Springer. https://doi.org/10.1007/978-3-319-30690-2.

Grabisch, M., & Labreuche, C. (2010). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *Annals of Operations Research, 175*, 247–290. https://doi.org/10.1007/s10479-009-0655-8

Grabisch, M., Marichal, J.-L., Mesiar, R., & Pap, E. (2009). *Aggregation Functions*. addressCambridge: Cambridge University Press.

Green, L., Myerson, J., & Mcfadden, E. (1997). Rate of temporal discounting decreases with amount of reward. *Memory & Cognition, 25*, 715–723. https://doi.org/10.3758/BF03211314

Horanská, L., & Siposová, A. (2018). A generalization of the discrete Choquet and Sugeno integrals based on a fusion function. *Information Sciences, 451–452*, 83–99. https://doi.org/10.1016/j.ins.2018.03.059

Huber, P. J. (2004). Robust statistics volume 523. *John Wiley & Sons*. https://doi.org/10.1002/0471725250

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General, 126*, 54.

Leszczyński, K., Penczek, P., & Grochulski, W. (1985). Sugeno's fuzzy measure and fuzzy clustering. *Fuzzy Sets and Systems, 15*, 147–158. https://doi.org/10.1016/0165-0114(85)90043-0

Li, G., Law, R., Vu, H. Q., & Rong, J. (2013). Discovering the hotel selection preferences of Hong Kong inbound travelers using the Choquet integral. *Tourism Management, 36*, 321–330.

Loewenstein, G., & Thaler, R. H. (1989). Anomalies: intertemporal choice. *Journal of Economic Perspectives, 3*, 181–193.

Marichal, J.-L. (2000). An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems, 8*, 800–807. https://doi.org/10.1109/91.890347

Marshall, A. W., Olkin, I., & Arnold, B. C. (2011). *Inequalities: theory of majorization and its applications, 2nd edition*. Springer Series in Statistics. Springer New York. https://doi.org/10.1007/978-0-387-68276-1.

Martino, R., Torrecillas, M. M., & Ventre, V. (2024). Discrete Choquet integral aggregation applied to intertemporal choices: financial literacy components and gender in anomalies and uncertainty aversion. *Quality & Quantity*. https://doi.org/10.1007/s11135-024-01957-z

Mesiar, R., & Mesiarová, A. (2008). Fuzzy integrals–what are they? *International Journal of Intelligent Systems, 23*, 199–212. https://doi.org/10.1002/int.20264

Narukawa, Y., Taha, M., & Torra, V. (2023). On the definition of probabilistic metric spaces by means of fuzzy measures. *Fuzzy Sets and Systems, 465*, Article 108528. https://doi.org/10.1016/j.fss.2023.108528

Ok, E. A., & Zhou, L. (2000). The Choquet bargaining solutions. *Games and Economic Behavior, 33*, 249–264. https://doi.org/10.1006/game.1999.0778

Ontkovičová, Z., & Torra, V. (2024). Computation of choquet integrals: Analytical approach for continuous functions. *Information Sciences, 679*, Article 121105. https://doi.org/10.1016/j.ins.2024.121105

Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty, 23*, 5–32.

Ruggeri, K., Panin, A., Vdovic, M., Većkalov, B., Abdul-Salaam, N., Achterberg, J., Akil, C., Amatya, J., Amatya, K., Andersen, T. L., et al. (2022). The globalizability of temporal discounting. *Nature Human Behaviour, 6*, 1386–1397.

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies, 4*, 155–161.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley Publications in Statistics.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society, 97*, 255–261.

Sparkman, G., Lee, N. R., & Macdonald, B. N. (2021). Discounting environmental policy: The effects of psychological distance over time and space. *Journal of Environmental Psychology, 73*, Article 101529. https://doi.org/10.1016/j.jenvp.2020.101529

Struk, P. (2006). Extremal fuzzy integrals. *Soft Computing, 10*, 502–505. https://doi.org/10.1007/s00500-005-0525-5

Sugeno, M. (1974). *Theory of fuzzy integrals and applications*. Ph.D. thesis Tokyo Institute of Technology.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters, 8*, 201–207.

Torra, V., & Narukawa, Y. (2008). The *h*-index and the number of citations: Two fuzzy integrals. *IEEE Transactions on Fuzzy Systems, 16*, 795–797.

Torra, V., & Narukawa, Y. (2016). Numerical integration for the Choquet integral. *Information Fusion, 31*, 137–145. https://doi.org/10.1016/j.inffus.2016.02.007

Türkarslan, E., & Torra, V. (2022). Measure identification for the Choquet integral: A Python module. *International Journal of Computational Intelligence Systems, 15*, 89. https://doi.org/10.1007/s44196-022-00146-w

Vitali, G. (1925). Sulla definizione di integrale delle funzioni di una variabile. *Annali di Matematica Pura ed Applicata, 2*, 111–121. https://doi.org/10.1007/BF02409934

Wakker, P. (1990). A behavioral foundation for fuzzy measures. *Fuzzy Sets and Systems, 37*, 327–350.

Van der Wal, A. J., Schade, H. M., Krabbendam, L., & Van Vugt, M. (2013). Do natural landscapes reduce future discounting in humans? *Proceedings of the Royal Society B: Biological Sciences, 280*, 20132295.

Wang, Z., & Klir, G. J. (1992). *Fuzzy Measure Theory*. Springer.

Willis-Moore, M. E., Haynes, J. M., Frye, C. C. J., Johnson, H. M., Cousins, D. J., Bamfo, H. D., & Odum, A. L. (2024). Recent experience affects delay discounting: Evidence across temporal framing, signs, and magnitudes. *Perspectives on Behavior Science, 47*, 365–392. https://doi.org/10.1007/s40614-024-00412-6

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man and Cybernetics, 18*, 183–190. https://doi.org/10.1109/21.87068

Yuan, C., Wang, E., Zhai, Q., & Yang, F. (2015). Temporal discounting in life cycle assessment: A critical review and theoretical framework. *Environmental Impact Assessment Review, 51*, 23–31. https://doi.org/10.1016/j.eiar.2015.01.001