The Extreme Points of Mean-Preserving Contractions

Andreas Kleiner Benny Moldovanu Philipp Strack

Mark Whitmeyer*

August 31, 2026

Abstract

We study multidimensional mean-preserving contractions (MPC) and their extreme points. Proposition 1 focuses on extreme MPCs of a measure μ: each finitely supported extreme MPC ν induces a partition of X, the domain of μ, into convex sets such that the support of ν on each element of the partition is an affinely independent set, and such that the restriction of ν on each element of the partition is itself an MPC of the restriction of the prior μ on that element. We apply our results to several questions concerning moment persuasion and categorization. In particular, we prove that categories and prototypes (key elements typically assumed to be exogenous in the classical literature on categorization) can endogenously emerge as outcomes of

*We thank participants at ESSET 2022 (Gerzensee), the 2023 SEA conference (New Orleans), and the 2024 SAET conference (Santiago) for their comments. Kun Zhang provided excellent research assistance. Kleiner, Moldovanu, and Whitmeyer acknowledge financial support from the DFG (German Research Foundation) via Germany's Excellence Strategy - EXC 2047/2 - 390685813, EXC 2126/2-390838866 and the CRC TR-224 (projects B01 and B02). Strack was supported by a Sloan Fellowship and the Cowles Foundation. Kleiner: Department of Economics, University of Bonn, andykleiner@gmail.com; Moldovanu: Department of Economics, University of Bonn, mold@uni-bonn.de; Strack: Department of Economics, Yale University, New Haven, philipp.strack@yale.edu; Whitmeyer: Department of Economics, Arizona State University, mark.whitmeyer@gmail.com.

optimization when the decision maker faces limits on the amount of information that she can either acquire or process.

Keywords: mean-preserving contractions; convex order; extreme points; Bayesian persuasion; information design; categorization

1 Introduction

We study the extreme points of the set of mean-preserving contractions (MPCs), also called *fusions*, of a given multidimensional measure. The set of MPCs of a given measure μ—taken here to be absolutely continuous with respect to the Lebesgue measure—is the set of probability distributions that are dominated by μ in the convex stochastic order, and are thus "less variable" than μ (Elton and Hill, 1992).

The burgeoning literature on information design and Bayesian persuasion highlights the importance of MPCs for economic applications: by a famous result that goes back to the work of Blackwell (1953) and Strassen (1965), the set of MPCs of a measure μ is also the set of distributions of posterior means that can be induced by some "signal" (or "experiment") starting from a prior belief μ about the state of the world. Thus, optimization over the set of feasible signals of an objective function that depends continuously on the posterior mean can be reduced to an optimization over the set of MPCs of μ. This set is convex and compact in the weak topology. By Bauer's theorem, the optimal way to reveal information about the state will be achieved at an extreme point of the set of MPCs. These features constitute the primary motivation for our study.

The case where the underlying state of the world is unidimensional has been studied along the above lines by Kleiner et al. (2021) and Arieli et al. (2023): each extreme measure is then characterized by a collection of (convex) intervals where either the prior mass on the interval remains in place (this corresponds to full information revelation of these states in a Bayesian persuasion problem), or where this mass is contracted into a measure with at most two elements in its support (this corresponds to pooling of states where information is only partially or not at all revealed). Here we extend these insights to the multidimensional case.[1]

Our main result, Proposition 1 in Section 2, focuses on finitely supported, extreme MPCs of a measure μ defined on a convex, compact subset X of $\mathbb{R}^n$. It shows that, necessarily, each such extreme MPC, ν, induces a partition of X into convex sets such that the support of ν on each element of the partition is an affinely independent set, and such that the restriction of ν on each element of the partition is itself an MPC of the restriction of the prior μ on that element. The analogy to the unidimensional result sketched above is clear by noting that the maximal cardinality of an affinely independent set in $\mathbb{R}^n$ is $n+1$ and this equals two for states on the real line. The finite-support restriction concerns the induced distribution of posterior means, not the prior, which may be continuous. It is nevertheless substantive: our results do not characterize optimal signals with infinitely many posterior means and therefore do not apply to all continuous-state persuasion applications. This restriction does not affect the applications in Section 4: there, finite support is either implied by the finite-action and strictly convex cost structure or imposed directly through the finite-memory constraint.

Dworczak and Kolotilin (2024) offer an elegant, unified duality approach to Bayesian persuasion.[2] They also prove a structural result of interest here, their Theorem 8: under the assumptions that the objective is continuously differentiable and that the prior has a density with respect to the Lebesgue measure, the moment-persuasion problem has a solution defined by a partition of the state space into convex regions; moreover, for each element of the partition, any induced posterior mean is an extreme point of the (possibly infinite) set of posterior means induced by the optimal signal for that region. While in one dimension Dworczak and Kolotilin's result reduces to the structure of extreme

[1] Kleiner et al. (2021) also characterize the extreme points of mean-preserving spreads of a given unidimensional measure while Roy (1976) and Yang and Yang (2026) look at the multidimensional case.

[2] In particular, they offer a sufficient condition for the optimality of *convex-partitional signals*, where the mass in each element of a partition is contracted into a single point. These special signals correspond to the *monotone-partitional signals* that were analyzed in the unidimensional case by several authors, e.g., Ivanov (2021).

points obtained by Kleiner et al. (2021) and Arieli et al. (2023),[3] it does not imply a cardinality bound in higher dimensions because the relevant set of posterior means may be infinite. Proposition 1 provides such a finite-support geometric bound for extreme MPCs: it shows that, on each region of the associated convex partition, the support is affinely independent and therefore has at most $n+1$ points.

In Section 3 we apply the above insights to the study of *convex-partitional* extreme points and to moment persuasion. An MPC is convex partitional if there is a partition of the state space into convex sets such that, for each cell, the support of the MPC is a singleton. Proposition 2 shows that convex-partitional MPCs are special in the sense that they are the finest (or "more informative") MPCs among those with a fixed number of support points. Proposition 3 uses the extreme-point structure from Proposition 1 plus the duality result of Dworczak and Kolotilin (2024) to show that a solution to a moment-persuasion problem can be canonically decomposed into two parts. First, in the global portion of the problem, the domain is partitioned into finitely many convex sets. Second, in the local portion, the sender solves an unconstrained problem on each cell by concavifying the value function. Whenever the moment-persuasion problem admits a finitely supported optimum, there exists such a canonical solution.

In Section 4 we study *categorization*, that is, the partition of a set of heterogeneous objects into subsets according to some rule. The classical contributions, e.g., Gärdenfors (2004), assume that the various categories or their *prototypes* (or "central" objects) are exogenous. Using our earlier results, we prove that these key elements of categorization can endogenously arise as natural outcomes of an optimization process when the decision maker faces limits on the amount of information that she can either acquire or process.

Appendix A contains the proofs omitted from the main text, and Appendix B complements Proposition 1 by developing some sufficient conditions for extremal mean-preserving contractions.

[3]This follows since in one dimension any convex set has at most two extreme points.

1.1 Preliminaries and Notation

We begin by introducing the notation used below and a few important concepts. Let X be a compact and convex subset of a finite-dimensional Euclidean space. For $A \subseteq X$, $\operatorname{int} A$ denotes the (relative) interior of A, and $\operatorname{ch} A$ denotes the convex hull of A. For a (Borel) measure μ on X, $\operatorname{supp} \mu$ denotes its support, and for measurable $A \subseteq X$, $\mu|_A$ denotes the restriction of μ to A, defined via $\mu|_A(B) := \mu(A \cap B)$ for any measurable B. We denote the set of probability measures on X by $\Delta(X)$; δ_x denotes a Dirac measure concentrated on $x \in X$, and the *barycenter* of a nonzero finite measure μ on X is defined by

$$r(\mu) := \frac{1}{\mu(X)} \int_X x d\mu(x).$$

Definition 1.1. For measures μ and ν, we say that μ *dominates* ν *in the convex order* (or that ν is a *MPC* of μ), denoted by $\mu \gtrsim \nu$, if $\int \phi \, \mathrm{d}\mu \geq \int \phi \, \mathrm{d}\nu$ for all convex functions ϕ such that both integrals exist. We write $\mu \succ \nu$ if μ dominates ν in the convex order and $\mu \neq \nu$. We denote by $F_\mu = \{\nu : \nu \lesssim \mu\}$ the set of *MPC*s of a given measure μ.

A finite collection of vectors $V = \{x_1, \ldots, x_k\}$ in $\mathbb{R}^n$ is *affinely independent* if the unique solution to $\sum_{i=1}^k \lambda_i x_i = 0$ and $\sum_{i=1}^k \lambda_i = 0$ is $\lambda_i = 0$, $i = 1, \ldots, k$. Recall also the well-known equivalence: $x_1, \ldots, x_k$ are affinely independent if and only if $x_2 - x_1, \ldots, x_k - x_1$ are linearly independent. Linear independence implies affine independence, but not vice versa. In particular, the maximal number of affinely independent vectors in $\mathbb{R}^n$ is $n+1$. A finite collection of vectors $V = \{x_1, \ldots, x_k\}$ in $\mathbb{R}^n$ is *convexly independent* if no element $x_i \in V$ lies in $\operatorname{ch}(V \setminus \{x_i\})$.

An *extreme* point of a convex set C is a point $x \in C$ that cannot be represented as a convex combination of two other points in C.[4] The usefulness of extreme points for optimization stems from *Bauer's Maximum Principle*: a convex, upper-semicontinuous functional on a non-empty, compact, and convex set C of a locally convex space attains its maximum at an extreme point of C. The *Krein–Milman theorem* states that any convex and compact set C in a locally convex space is the closed, convex hull of its extreme points. In particular, such a set has extreme points.

[4]Formally, $x \in C$ is an extreme point of C if $x = \alpha y + (1-\alpha)z$, for $z, y \in C$ and $\alpha \in [0,1]$ imply together that $y = x$ or $z = x$.

2 Extreme Mean-Preserving Contractions

Our first main result offers a necessary condition for a finitely supported measure ν to be an extreme point of F_μ. A key feature of any extremal MPC is the partition of the domain X into finitely many convex sets P such that all the original mass $\mu|_P$ remains within P and is "fused" into $\nu|_P$ whose support is an affinely independent set of points.

Proposition 1. *Let $X \subseteq \mathbb{R}^n$ be compact and convex, and let μ be an absolutely continuous probability measure on X.*[5] *Suppose that ν is an extreme point of F_μ that is finitely supported. Then there exists a partition $\mathcal{P}$ of X into finitely many convex sets such that, for each $P \in \mathcal{P}$, $\nu|_P \precsim \mu|_P$ and $\nu|_P$ has affinely independent support.*

Proposition 1 constrains the complexity of extremal MPCs: since an affinely independent set in $\mathbb{R}^n$ has at most $n+1$ elements, each cell of the partition supports at most $n+1$ atoms. This generalizes the "bi-pooling" structure identified in one dimension by Kleiner et al. (2021) and Arieli et al. (2023), where each interval in the partition supports at most two points.

Our result also provides a complementary perspective on Theorem 8 of Dworczak and Kolotilin (2024). Using duality, these authors show that an optimal signal for moment persuasion can be chosen so that the state space is divided into convex pooling regions and such that, within each region, the induced posterior means are extreme points of their own closed convex hull. As mentioned in the Introduction, in one dimension this local extremality condition implies bi-pooling. In higher dimensions, however, it does not imply a cardinality bound: Dworczak and Kolotilin (2024) note that the relevant set of posterior means may be infinite, and that a finest-partition approach might yield a sharper $n+1$ bound. Proposition 1 provides such a finite-support geometric restriction for extreme MPCs, showing that on the regions of the associated convex partition, the support is affinely independent and therefore has at most $n+1$ points.

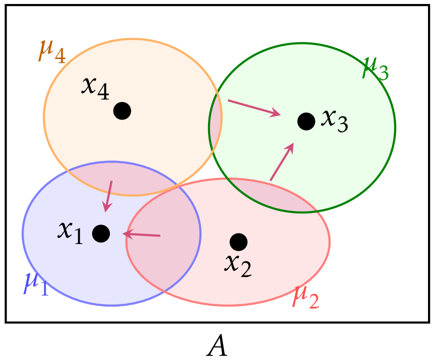
The proof of Proposition 1 (see Appendix A) rests on the observation that affine dependence among the support points within a cell creates a degree of freedom that is

[5] Throughout, absolute continuity is also meant with respect to the Lebesgue measure on $\mathbb{R}^n$.

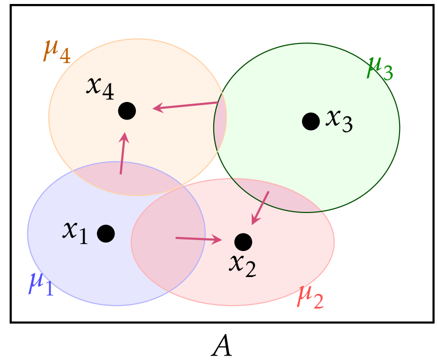
incompatible with extremality. Suppose ν is an extreme MPC of μ. Consider the finest partition of X into convex sets such that on each cell A, the restriction $\nu|_A$ is an MPC of $\mu|_A$. Fix a cell A with support points $\{x_1, \ldots, x_m\}$. By a decomposition result of Cartier et al. (1964) (see Lemma 1), the restriction $\mu|_A$ can be written as a sum of positive component measures $\mu_1 + \cdots + \mu_m$, where each μ_i has barycenter x_i and mass $\nu(\{x_i\})$. Since the partition is the finest possible, these component measures cannot be separated into groups whose aggregate supports have disjoint convex hulls—otherwise a cell A could be split further, contradicting our assumption.

Now suppose the support $\{x_1, \ldots, x_m\}$ is affinely dependent: there exist coefficients β_i, not all zero, satisfying $\sum_i \beta_i = 0$ and $\sum_i \beta_i x_i = 0$. This relation determines how the weights $\nu(\{x_i\})$ can be adjusted while maintaining the overall barycenter. Whenever the convex hulls of the supports of two component measures have a common interior point—which does not require the supports themselves to intersect—a small amount of mass can be transferred from one to the other without changing either barycenter. Figure 1 illustrates this construction for four support points in $\mathbb{R}^2$, where the hulls of neighboring component measures overlap directly. In general, the relative interiors of these convex hulls need not overlap pairwise, and a main step of the proof (Lemmas 2–4) consists in showing that the necessary mass transfers can always be carried out through a chain of pairwise overlaps. Performing this perturbation in both directions produces two distinct MPCs $\tilde{\nu}$ and $\hat{\nu}$ of μ such that $\nu = \frac{1}{2}(\tilde{\nu} + \hat{\nu})$, contradicting the assumption that ν is an extreme point.

Proposition 1 provides the necessary structure for finitely supported extreme MPCs. It is natural to ask about sufficient conditions. We address this question in Appendix B. Take a finite, convex partition $\mathscr{P}$ and suppose that an MPC satisfies the two local properties identified in Proposition 1: for each cell $P \in \mathscr{P}$, the restriction $\nu|_P$ is an MPC of $\mu|_P$, and its support is affinely independent. The remaining obstacle to extremality is of a global nature: even if each cell satisfies the necessary local conditions, it may still be possible to redistribute prior mass across cells, while preserving the mass and barycenter assigned to each cell. Performing such redistributions may allow the representation of a partition satisfying the necessary conditions as a convex combination of other measures



(a) Perturbation $\tilde{\nu}$: mass shifts toward x_1 and x_3



(b) Perturbation $\hat{\nu}$: mass shifts toward x_2 and x_4

Figure 1: Illustration of the proof of Proposition 1 for four support points in $\mathbb{R}^2$ (which are necessarily affinely dependent). The component measures $\mu_1, \ldots, \mu_4$ have barycenters at $x_1, \ldots, x_4$ (ellipses represent the convex hulls of their supports). The affine dependence $\sum \beta_i x_i = 0$ permits two opposite perturbations: **(a)** mass is transferred through the overlaps of these convex hulls toward x_1 and x_3 while maintaining the barycenters, producing $\tilde{\nu}$; **(b)** mass shifts toward x_2 and x_4 while maintaining the barycenters, producing $\hat{\nu}$. Since $\nu = \frac{1}{2}(\tilde{\nu} + \hat{\nu})$, ν is not an extreme point.

in F_μ. In Appendix B we formalize these mass redistributions as feasible flows.[6] We show that if the zero flow is the unique feasible flow for $\mathscr{P}$, then the above local conditions are sufficient for extremality. Conversely, if there is a non-zero feasible flow, one can construct a non-extreme *MPC* satisfying them. Although the material in Appendix B is not needed for our applications below, it clarifies why the properties in Proposition 1 are only necessary, and it identifies an intuitive condition under which the local structure is in fact sufficient.

[6]A feasible flow is a collection of signed measures, one for each cell, describing the mass removed from that cell and reallocated to other cells. We consider flows that are balanced across the whole partition and have zero net mass and zero net first moment for each cell.

3 Convex Partitional MPCs and Moment Persuasion

We say that an MPC ν of μ is *convex partitional* if there is a partition of X into convex sets such that, for each cell P, the support of $\nu|_P$ is a singleton and $\nu|_P \precsim \mu|_P$. An extreme point of F_μ need not be convex partitional, but, as we show below, convex-partitional MPCs are special in the sense that they are the finest MPCs among those with a fixed number of support points. The next proposition generalizes the one-dimensional result, Lemma 1 of Ivanov (2021):

Proposition 2. *Suppose that ν is an MPC of μ with K points in its support. Then there exists a convex-partitional MPC λ of μ with at most K points in its support that satisfies $\nu \precsim \lambda$. If ν is not convex partitional, then $\nu \prec \lambda$.*

In Section 4, we apply this proposition to categorization. Under either processing or memory constraints, where a decision maker is constrained to arrange given data in a finite number of "bins," there always exists an optimal arrangement such that the data is categorized according to a convex partition of the state space.

3.1 Moment Persuasion

In a special class of multidimensional Bayesian persuasion problems, the sender's payoff can be written in reduced form as a function of the receiver's vector of posterior means. Let $X \subseteq \mathbb{R}^d$ be compact and convex, and let $\mu \in \Delta(X)$ be absolutely continuous with full support on X. For any posterior $G \in \Delta(X)$, let

$$m(G) := \int_X x\, dG(x) \in X.$$

Assume that there exists a Lipschitz function $V\colon X \to \mathbb{R}$ such that the sender's payoff under posterior G is $V(m(G))$. Given V and μ, the sender solves

$$\max_{\nu \in F_\mu} \mathbb{E}_\nu[V].$$

To formulate the local problems, for any convex set $P \subseteq X$ with $\mu(P) > 0$, let $r_P(\mu) := r(\mu|_P)$, and let $D_{\mu,P}$ denote the set of probability measures supported on P that are mean-preserving spreads of $\delta_{r_P(\mu)}$. The *unconstrained persuasion problem* on P is

$$\max_{\eta \in D_{\mu,P}} \mathbb{E}_\eta [V].$$

When P is a simplex, this local problem is affinely equivalent to the usual formulation studied by Kamenica and Gentzkow (2011).

We define the upper concave envelope of V on P, $\operatorname{cav}_P V$, to be the smallest concave function that lies pointwise above V on P. Alternatively, we call this the *concavification* of V on P. We say that an optimal solution ν^* to the persuasion problem is *canonical* if there exists a finite convex partition $\mathscr{P}^*$ of X such that, for each cell $P \in \mathscr{P}^*$ with $\mu(P) > 0$, the normalized restriction $\mu(P)^{-1}\, \nu^*|_P$ solves the unconstrained persuasion problem on P.

Proposition 3. *Suppose that the moment-persuasion problem admits a finitely supported optimal solution. Then there exists a canonical solution $\nu^* \in F_\mu$. Moreover, if $\mathscr{P}^*$ is the corresponding partition, then*

$$\mathbb{E}_{\nu^*}[V] = \sum_{P \in \mathscr{P}^*} \mu(P) \operatorname{cav}_P V(r_P(\mu)).$$

In the usual moment-persuasion problem with finitely many receiver actions and ties broken in the sender's favor, a finitely supported optimum exists. Finitely many actions do not, however, imply our Lipschitz assumption: the induced value function is generally only upper semicontinuous. Thus, Proposition 3 applies to such environments only when the induced V is Lipschitz.

Proposition 3 indicates that the optimal solution to a moment persuasion problem with a finitely supported optimum can be decomposed into two parts. First, in the **global** portion of the problem, the domain is partitioned into finitely many convex sets. Second, in the **local** portion of the problem, the sender solves an unconstrained problem on each of these convex sets; on each of these convex sets, the optimizer of the local problem is obtained by concavifying the value function. As Proposition 3 states, whenever the moment-persuasion problem admits a finitely supported optimal solution, there exists such a canonical solution.[7]

[7] Yet again, this result reduces to bi-pooling in one dimension.

Proposition 3 also clarifies how our approach connects to Dworczak and Kolotilin (2024). Their duality theorem supplies the local price-function representation we use in the proof, while Proposition 1 supplies the global finite-support geometry. Proposition 3 combines these two ingredients: our finest-partition argument rules out the further splitting of a cell by the local dual price function; thus, the normalized restriction of the optimal MPC to each cell solves the unconstrained local problem. To elaborate, if the moment-persuasion problem has a finitely supported optimizer, we may select one that is an extreme point of F_{μ}. Proposition 1 shows that, on each cell, the restriction of the optimizer is a mean-preserving contraction of the restricted prior. Global optimality implies that no cell can be improved while the other cells are held fixed, and hence each restriction is locally optimal relative to the restricted prior. The additional content of the proposition is that, after choosing a fine enough partition, this locally constrained problem becomes an unconstrained persuasion problem on the cell. Intuitively, if the local dual price function had a nontrivial kink, its affine pieces would further split the cell into smaller convex regions, contradicting the fact that the partition was finest. Consequently, the local price is affine on each cell, and the restriction of the optimizer attains the concavification value on that cell.

The decomposition described above is primarily a conceptual result rather than an algorithmic simplification of the problem. In particular, Proposition 3 does not characterize the optimal global partition, nor does it imply that the search over partitions can be separated from the local persuasion problems. Moreover, the search for the optimal partition is complex since it depends on the solution of the original problem.

Nevertheless, the decomposition is useful because it identifies where the difficulty lies. Conditional on a candidate partition, the sender's problem on each cell has the familiar concavification form. Thus, the global part of the problem is not an arbitrary search over all information structures, but a search over convex regions on which the local concavification program is conducted. In this sense, the proposition provides a "normal form" for finitely supported optimizers: we can restrict attention to solutions that first divide the state space into convex cells and then solve unconstrained persuasion problems within those cells. This clarifies the economic content of the optimal signal, since posterior means within a cell are generated by local pooling, while the boundaries between cells determine which states are treated as qualitatively distinct.

4 Applications to Categorization

Categorization, or the partition of heterogeneous objects according to some rule, is a central part of machine learning and of our interaction with the world more broadly. An influential work that formally studies this process is Gärdenfors (2004). There, a main construct is the notion of a conceptual space—in economics, a state space—and its division into convex sets, each containing a *prototype*, i.e., a kind of "central" object. For example, one can think of colors: "red" describes a whole family of hues, as does "blue," and so on. Gärdenfors describes two ways of producing such a partition: 1. We can define a *natural property* as covering a convex region of the conceptual space, in which case the prototype of each region is its central object according to some measure of centrality, e.g., its barycenter or some multidimensional median; 2. One can start with a finite set of prototypes and then partition the conceptual space by grouping points together that are "closest" to each prototype. Closeness in the sense of Euclidean distance then yields a *Voronoi diagram*.

Both constructions above seem reasonable and realistic, yet note that the central properties are exogenous: that is, the convexity of the regions in the first construction or the existence of prototypes in the second are assumed.

In this section, Proposition 3 turns out to be especially useful. We use it to show that convex categories and prototypes can emerge as natural outcomes of optimization when the decision maker faces limits on the amount of information she can either acquire or process.

We provide two different micro-foundations for categorization. In each of them, the outcome is a convex partition of the state space with a single prototype per partition element. In the first derivation, we show that such an endogenous structure emerges as information becomes sufficiently cheap, while in the second we show that it emerges when a decision maker must store information in finitely many bins.

4.1 Decision Problems with Flexible Information Acquisition

We continue to work in the reduced-form setting of Section 3.1. Let $X := [0,1]^d$, and let μ be an absolutely continuous probability measure with full support on X. The DM chooses an action from a finite set of actions A, and for each $a \in A$, her expected payoff at posterior mean $x \in X$ is affine:

$$z_a(x) := b_a \cdot x + d_a.$$

Her induced reduced-form value function is therefore

$$V(x) := \max_{a \in A} z_a(x),$$

which is piecewise-affine and Lipschitz. We furthermore assume that each action $a \in A$ is strictly optimal at some $x \in X$.[8]

Prior to her decision, the DM acquires or processes information. We model the cost of information by a functional $C\colon F_\mu \to \mathbb{R}$ of the form

$$C(\nu) := \kappa \int_X c\, d\nu,$$

where $\kappa > 0$ and where the function $c\colon X \to \mathbb{R}$ is strictly convex and twice continuously differentiable. Since the partial derivatives of c are bounded on the compact set X, the mean-value theorem implies that c is Lipschitz-continuous on X. The DM solves the problem

$$\max_{\nu \in F_\mu} \mathbb{E}_\nu[V - \kappa c].$$

Note that this specification of costs is posterior-mean separable: two information structures that induce the same distribution of posterior means have the same cost, and the cost is affine in that induced distribution. This restriction is substantive. It rules out technologies that distinguish between two information structures that induce the same distribution of posterior means but different conditional higher moments, for instance.

[8] As V is continuous, this means that the set of posterior means at which each action is strictly optimal is of full dimension in X. Moreover, this restriction is without loss of generality for our exercise: given an arbitrary finite menu, a decision maker who can acquire information as in this section would never choose actions without this property.

However, it is exactly the class of costs for which the information-acquisition problem can be written directly over the feasible set F_μ of posterior-mean distributions.[9]

In the categorization interpretation, a posterior mean is the prototype attached to a category, so $C(\nu)$ can be understood here as the expected cost of producing, storing, or processing the prototypes induced by the information policy.

Proposition 4. *For any decision problem, cost function, and prior of the above form, there exists a $\bar{\kappa} > 0$ such that, if $\kappa \leq \bar{\kappa}$, there exists an optimal solution to the DM's information acquisition problem that is convex partitional, i.e., with a single prototype per category.*

If a canonical cell contains multiple posterior means, then at least two different actions must be induced on that cell. At an optimum, the difference between the corresponding slope vectors must be matched by the scaled difference in marginal information costs. In the proof, the number of actions enters through a parameter β, defined as the smallest non-zero slope difference that can arise between the action-specific slope vectors $\{b_a\}_{a \in A}$, while a parameter α bounds the largest possible difference in marginal costs. The sufficient bound $\bar{\kappa} := \frac{\beta}{2\alpha}$ gets tighter when two action-specific payoff hyperplanes have similar slopes. Thus, enlarging the set of actions A (weakly) shrinks β, because it adds new slope vectors and possibly new slope differences, thus (weakly) reducing the range of costs for which the convex partitional conclusion is guaranteed.

4.2 Decision Problems with Finite Memory

We continue to work in the reduced-form setting of Section 3.1, with $X := [0,1]^d$ and μ being an absolutely continuous probability measure with full support on X. Likewise, V remains the reduced-form value function from Section 4.1, but we replace the smooth information cost assumed there with a hard cap on the number of categories.

[9] Posterior-mean separable costs have recently been given a revealed-preference foundation by Mensch and Malik (2023). They study stochastic choice data in environments where payoffs depend on beliefs only through posterior means and characterize when observed behavior can be rationalized by such an information cost. Accordingly, when posterior means are payoff relevant, our assumption corresponds to a behaviorally testable restriction on information acquisition.

Specifically, the DM has access to $K \in \mathbb{N}$ bins or categories, and she chooses a stochastic mapping that assigns states to bins. When choosing her action, the DM learns only to which bin the realized state was mapped, not the realized state itself. Importantly, we do not assume that the K bins correspond to a partition, much less a convex one. For instance, the DM could assign any $x \in X$ to one of the bins uniformly at random. The DM's problem becomes

$$\max_{\nu \in F_\mu} \mathbb{E}_\nu [V] \quad \text{s.t.} \quad |\operatorname{supp} \nu| \leq K.$$

Because V is the maximum of finitely many affine functions, it is convex. This means that information is beneficial to the DM and Proposition 2 implies the following:

Proposition 5. *There is an optimal categorization that is convex partitional. If the decision problem has at least K actions such that each one of them is strictly optimal at some $x \in X$, then any optimal categorization must be convex partitional.*

4.3 Decision-Making with Complexity Constraints

We now move to a finite-state environment in which the DM's value need not be a function of posterior means alone. Suppose there is a finite state space Θ. There is a Bayesian DM with a finite set of actions A, with a continuous utility function $u \colon A \times \Theta \to \mathbb{R}$ and with a prior $\eta_0 \in \operatorname{int} \Delta(\Theta)$.

The DM observes information before making a decision. Formally, before choosing an action $a \in A$, the DM observes the realization of a signal $\pi \colon \Theta \to \Delta(S)$, where S is a compact set of signal realizations. Let $\rho \in \Delta(\Delta(\Theta))$ denote the distribution over posteriors η induced by the prior together with the signal.

We assume that the DM optimizes over decision rules $\sigma \colon S \to \Delta(A)$ whose image has at most $K \in \mathbb{N}$ points. We call such a decision rule *simplicity-constrained*. It is immediate that this is equivalent to the DM being restricted to choosing an MPC ν of ρ that is supported on at most K points before choosing an arbitrary unconstrained decision rule. We say that a simplicity-constrained decision rule is convex partitional if ν is a convex-partitional MPC of ρ. Then Proposition 5 implies the following:

Corollary 1. *There is an optimal simplicity-constrained decision rule that is convex partitional. If (i) the distribution over posteriors is absolutely continuous and has full support on* $\Delta(\Theta)$ *and (ii) the decision problem has at least* K *actions that are strictly optimal at some posterior, then any optimal simplicity-constrained decision rule must be convex partitional.*

A Omitted Proofs

A.1 Proof of Proposition 1

We first need several auxiliary lemmas:

Lemma 1. *(Cartier et al. (1964)): Let* μ, ν *be positive measures on a compact, convex set* X *such that* $\nu = \sum_{i=1}^{k} \alpha_i \delta_{x_i}$ *where* $\sum_{i=1}^{k} \alpha_i = 1$ *and* $\alpha_i > 0$ *for each* i*. Then* $\nu \precsim \mu$ *if and only if there exist positive measures* $\mu_1, \ldots, \mu_k$ *such that, for all* i*, the barycenter of* μ_i *equals* x_i*, and such that* $\mu = \sum_{i=1}^{k} \alpha_i \mu_i$ *(in particular* $\delta_{x_i} \precsim \mu_i$ *for all* i*).*

For any two measures μ and ν on X, $\mu \leq \nu$ denotes the pointwise order defined by $\mu(A) \leq \nu(A)$ for all measurable $A \subseteq X$.

Lemma 2. *Let* μ *be a positive measure supported on* X *and let* $y \in \mathrm{int}(\mathrm{ch}(\mathrm{supp}(\mu)))$*. Then there exists a positive measure* π *such that* $\pi \leq \mu$*,* $\pi(X) > 0$*, and* $r(\pi) = y$*.*

Proof. Choose a finite set $\{y_1, \ldots, y_k\} \subseteq \mathrm{supp}\,\mu$ such that y lies in the interior of the convex hull of $\{y_1, ..., y_k\}$. Since the convex hull operator is continuous (e.g., Sertel (1989)), there is $\varepsilon > 0$ such that for all $z_i \in B_\varepsilon(y_i)$, $y \in \mathrm{ch}(z_1, ..., z_k)$. Choosing $\varepsilon > 0$ small enough, we can assume that $B_\varepsilon(y_i)$ is disjoint from $B_\varepsilon(y_j)$ whenever $i \neq j$.

For each i, let $m_i := \mu(B_\varepsilon(y_i))$ and $z_i := r(\mu|_{B_\varepsilon(y_i)})$. Choose $\theta_i \geq 0$ with $\sum_i \theta_i = 1$ and $y = \sum_i \theta_i z_i$, and set

$$c := \min_{\theta_i > 0} \frac{m_i}{\theta_i}, \qquad \pi := c \sum_i \frac{\theta_i}{m_i} \mu|_{B_\varepsilon(y_i)}.$$

Then $\pi \leq \mu$, $\pi(X) = c > 0$, and $r(\pi) = y$. ∎

Remark A.1. Observe that the measure π constructed in the proof of the above Lemma satisfies $\pi(X) \geq \min_i\{\mu(B_\varepsilon(y_i))\}$.

Lemma 3. *Let μ_1, μ_2 be positive measures on X and set $C_i := \mathrm{ch}(\mathrm{supp}(\mu_i))$. Fix $z \in \mathrm{int}(C_1) \cap \mathrm{int}(C_2)$ and put $L_i := \mathrm{span}(C_i - C_i)$. For every $s \in \mathbb{R}$ and $d \in \mathbb{R}^n$ such that $d - sz \in L_1 + L_2$, there exist a signed measure λ and $\bar{\varepsilon} > 0$ such that $\lambda(X) = s$, $\int x \,\mathrm{d}\lambda = d$, and, for every $\varepsilon \in (0, \bar{\varepsilon})$, $\pi := \varepsilon\lambda$ satisfies*

$$\mu_1 + \pi \geq 0, \qquad \mu_2 - \pi \geq 0.$$

The condition requires the prescribed mass and first moment to be compatible with the two affine hulls and is automatic if $L_1 + L_2 = \mathbb{R}^n$. Because $\pi = \varepsilon\lambda$, its total variation vanishes as $\varepsilon \to 0$, as required in Lemma 4.

Proof. Choose $p_i \in L_i$ such that $d - sz = p_2 - p_1$. For all sufficiently large $m > |s| + 1$, set

$$z_1 := z + \frac{p_1}{m - s} \in \mathrm{int}(C_1), \qquad z_2 := z + \frac{p_2}{m} \in \mathrm{int}(C_2).$$

Lemma 2 yields positive measures $\zeta_0 \leq \mu_1$ and $\xi_0 \leq \mu_2$ with barycenters z_1 and z_2, respectively. Put

$$\zeta := \frac{m - s}{\zeta_0(X)}\zeta_0, \qquad \xi := \frac{m}{\xi_0(X)}\xi_0, \qquad \lambda := \xi - \zeta.$$

Then $\zeta \leq c_1\mu_1$ and $\xi \leq c_2\mu_2$, where $c_1 := (m - s)/\zeta_0(X)$ and $c_2 := m/\xi_0(X)$. Moreover,

$$\lambda(X) = s, \qquad \int x \,\mathrm{d}\lambda = mz_2 - (m - s)z_1 = sz + p_2 - p_1 = d.$$

Set $\bar{\varepsilon} := 1/\max\{c_1, c_2\} > 0$. For $\varepsilon \in (0, \bar{\varepsilon})$ and $\pi := \varepsilon\lambda$ we have $\varepsilon\xi \leq \varepsilon c_2\mu_2 \leq \mu_2$ and $\varepsilon\zeta \leq \varepsilon c_1\mu_1 \leq \mu_1$, so that

$$\mu_1 + \pi = (\mu_1 - \varepsilon\zeta) + \varepsilon\xi \geq 0 \qquad \text{and} \qquad \mu_2 - \pi = (\mu_2 - \varepsilon\xi) + \varepsilon\zeta \geq 0.$$

Finally, $\pi(X) = \varepsilon\lambda(X) = \varepsilon s$ and $\int x \,\mathrm{d}\pi = \varepsilon \int x \,\mathrm{d}\lambda = \varepsilon d$. ■

Corollary 2. *Let μ_1, μ_2 be positive measures with barycenters x_1 and x_2, respectively. If*

$$C = \mathrm{int}(\mathrm{ch}(\mathrm{supp}(\mu_1))) \cap \mathrm{int}(\mathrm{ch}(\mathrm{supp}(\mu_2))) \neq \emptyset,$$

1. *there is a signed measure π satisfying $\pi(X) > 0$, $\mu_1 + \pi \geq 0$, $\mu_2 - \pi \geq 0$ and such that the barycenter of $\tilde{\mu}_1 := \mu_1 + \pi$ is x_1.*

2. *there is a signed measure π satisfying $\pi(X) < 0$, $\mu_1 + \pi \geq 0$, $\mu_2 - \pi \geq 0$ and such that the barycenter of $\tilde{\mu}_1 := \mu_1 + \pi$ is x_1.*

The corollary shows that mass can be moved in either direction while preserving the barycenter of μ_1.

Proof. Fix $z \in C$ and let L_1 be as in Lemma 3. For $s \in \{-1,1\}$ set $d = sx_1$. Since $d - sz = s(x_1 - z) \in L_1$, the lemma yields, for small $\varepsilon > 0$, a signed measure π with $\pi(X) = \varepsilon s$, $\int x\,\mathrm{d}\pi = \varepsilon s x_1$, and the required positivity properties. Hence $r(\mu_1 + \pi) = x_1$; taking $s = 1$ and $s = -1$ proves the two claims. ■

For a set $A \subseteq \mathbb{R}^n$, let $\operatorname{aff} A$ denote the affine span of A:

$$\operatorname{aff} A = \{\sum_{i=1}^{k} \alpha_i x_i : k > 0, x_i \in A, \alpha_i \in \mathbb{R}, \sum_{i=1}^{k} \alpha_i = 1\}.$$

Lemma 4. *Let $\{\mu_i \mid i \in I\}$ be a finite collection of positive measures with $r(\mu_i) = x_i$ and $\operatorname{aff}(\operatorname{supp} \mu_i) = \mathbb{R}^n$ for every $i \in I$. Let $\mu = \sum_{i \in I} \mu_i$, and let $V = \{J_1, \ldots, J_k\}$ be a partition of I.*

Consider the graph G whose vertex set is V and which contains an edge between J and J' if

$$\operatorname{int}(\operatorname{ch}(\operatorname{supp}(\sum_{j \in J} \mu_j))) \cap \operatorname{int}(\operatorname{ch}(\operatorname{supp}(\sum_{j \in J'} \mu_j))) \neq \emptyset.$$

Suppose that G is connected, and let λ be a finite signed measure such that $\mu + \varepsilon\lambda$ is positive for all sufficiently small $\varepsilon > 0$ and such that, for some reals $(\gamma_i)_{i \in I}$,

$$\int x\,\mathrm{d}\lambda = \sum_{i \in I} \gamma_i x_i \qquad \text{and} \qquad \lambda(X) = \sum_{i \in I} \gamma_i.$$

Then, for all $\varepsilon > 0$ small enough we can decompose $\tilde{\mu} := \mu + \varepsilon\lambda$ into positive measures $\tilde{\mu}_J$ with $\sum_{J \in V} \tilde{\mu}_J = \tilde{\mu}$ and $r(\tilde{\mu}_J) \in \operatorname{aff}\{x_j | j \in J\}$.

Proof. Let $\mu_J = \sum_{i \in J} \mu_i$, and let $\{\lambda_J\}_{J \in V}$ be such that $\lambda = \sum_{J \in V} \lambda_J$, $\hat{\mu}_J := \mu_J + \varepsilon\lambda_J \geq 0$, and $\operatorname{supp} \mu_J \subseteq \operatorname{supp} \hat{\mu}_J$.[10] By hypothesis there are reals $(\gamma_i)_{i \in I}$ such that $\int x\,\mathrm{d}\lambda = \sum_{i \in I} \gamma_i x_i$ and $\sum_{i \in I} \gamma_i = \lambda(X)$.

---

[10]For all sufficiently small $\varepsilon > 0$, such a decomposition exists with the λ_J independent of ε. Let $f_J := \frac{\mathrm{d}\mu_J}{\mathrm{d}\mu} \in [0,1]$ (so $\sum_J f_J = 1$ μ-a.e.) and write $\lambda = \lambda^+ - \lambda^-$. Since $\mu + \varepsilon\lambda \geq 0$ for all sufficiently small $\varepsilon > 0$, the negative part of λ satisfies $\lambda^- \leq K\mu$ for some finite K (one may take $K = 1/\varepsilon_0$ for any ε_0 with $\mu + \varepsilon_0\lambda \geq 0$). Fixing any J_0, set $\lambda_{J_0} := \lambda^+ - f_{J_0}\lambda^-$ and $\lambda_J := -f_J\lambda^-$ for $J \neq J_0$. Then $\sum_J \lambda_J = \lambda$, and each negative part satisfies $\lambda_J^- = f_J\lambda^- \leq K f_J \mu = K\mu_J$, so $\hat{\mu}_J = \mu_J + \varepsilon\lambda_J \geq (1 - \varepsilon K)\mu_J \geq 0$ and $\operatorname{supp}\mu_J \subseteq \operatorname{supp}\hat{\mu}_J$ whenever $\varepsilon \leq 1/K$. Since

Consider a spanning tree T of G. If T has at most two vertices, skip the following construction and proceed directly to the corresponding terminal case below (with $L(\hat{J}) = L(\bar{J}) = \emptyset$). Otherwise, for each leaf J of T, define a weight $w_J = \sum_{i\in J}\gamma_i - \lambda_J(X)$ and let

$$\tilde{\mu}_J = \hat{\mu}_J + \pi_J,$$

where π_J is a measure chosen such that $\pi_J(X) = \varepsilon w_J$, $\tilde{\mu}_J \geq 0$, $\pi_J \leq \frac{1}{|I|+1}\hat{\mu}_{J'}$ (where J' is the neighbor of J in T), and

$$\varepsilon\int x\,\mathrm{d}\lambda_J + \int x\,\mathrm{d}\pi_J = \varepsilon\sum_{i\in J}\gamma_i x_i.$$

Such π_J exist by Lemma 3 whenever ε is chosen small enough.

Proceeding inductively, consider the tree obtained by deleting all leaves from the previous tree. For each leaf J of the new tree, denote by $L(J)$ the set of all neighbors of J in T that were deleted in earlier rounds (each deleted vertex thus lies in $L(J)$ for exactly one J, its unique neighbor at the time of deletion). Note that for each leaf J of the new tree, $\hat{\mu}_J - \sum_{K\in L(J)}\pi_K \geq 0$ and $\mathrm{supp}(\hat{\mu}_J) \subseteq \mathrm{supp}(\hat{\mu}_J - \sum_{K\in L(J)}\pi_K)$. Define the weight $w_J = \sum_{K\in L(J)} w_K + \sum_{i\in J}\gamma_i - \lambda_J(X)$ for each leaf J of the new tree. As long as there are at least 3 vertices remaining, define

$$\tilde{\mu}_J = \hat{\mu}_J - \sum_{K\in L(J)}\pi_K + \pi_J$$

where the π_J's are chosen so that $\pi_J(X) = \varepsilon w_J$, $\tilde{\mu}_J \geq 0$, $\pi_J \leq \frac{1}{|I|+1}\hat{\mu}_{J'}$ (where J' is the neighbor of J in the remaining tree), and

$$\varepsilon\int x\,\mathrm{d}\lambda_J - \sum_{K\in L(J)}\int x\,\mathrm{d}\pi_K + \int x\,\mathrm{d}\pi_J = \varepsilon\sum_{i\in J}\gamma_i x_i. \tag{A1}$$

Such a choice exists by Lemma 3 whenever $\varepsilon > 0$ is small enough. We repeat this procedure until we obtain a tree with at most two vertices.

---

the λ_J are independent of ε and of finite total variation, the weights w_J and the barycenter targets in the construction below do not depend on ε; the shifts π_J then have mass and first moment proportional to ε, hence vanishing as $\varepsilon \to 0$, and so exist by Lemma 3 once ε is small enough.

If only two nodes are left in the remaining tree (denote them by $\hat{J}$ and $\bar{J}$), define

$$\tilde{\mu}_{\hat{J}} = \hat{\mu}_{\hat{J}} - \sum_{K \in L(\hat{J})} \pi_K + \pi_{\hat{J}}$$
$$\tilde{\mu}_{\bar{J}} = \hat{\mu}_{\bar{J}} - \sum_{K \in L(\bar{J})} \pi_K - \pi_{\hat{J}}$$

where $\pi_{\hat{J}}$ is chosen so that $\pi_{\hat{J}}(X) = \varepsilon w_{\hat{J}}$, $\tilde{\mu}_{\hat{J}} \geq 0$, $\tilde{\mu}_{\bar{J}} \geq 0$, and

$$\varepsilon \int x \, \mathrm{d}\lambda_{\hat{J}} - \sum_{K \in L(\hat{J})} \int x \, \mathrm{d}\pi_K + \int x \, \mathrm{d}\pi_{\hat{J}} = \varepsilon \sum_{i \in \hat{J}} \gamma_i x_i.$$

If only one node, denoted by $\bar{J}$, is left in the remaining tree, define

$$\tilde{\mu}_{\bar{J}} = \mu_{\bar{J}} + \varepsilon \lambda_{\bar{J}} - \sum_{K \in L(\bar{J})} \pi_K.$$

In either case, it follows that $\sum_{J \in V} \tilde{\mu}_J = \sum_{J \in V} \hat{\mu}_J$ by construction. Moreover, for all $J \neq \bar{J}$, the barycenter of $\tilde{\mu}_J$ is in $\mathrm{aff}\{x_i | i \in J\}$.[11]

It remains to verify that the barycenter of $\tilde{\mu}_{\bar{J}}$ is in $\mathrm{aff}\{x_i | i \in \bar{J}\}$. For $J \in V$, let $b_J = \int x \, \mathrm{d}\pi_J$. For any node $J \neq \bar{J}$ we obtain from (A1) that

$$b_J = \varepsilon \sum_{i \in J} \gamma_i x_i - \varepsilon \int x \, \mathrm{d}\lambda_J + \sum_{K \in L(J)} b_K, \tag{A2}$$

where $L(J) = \emptyset$ if J is a leaf of T. Suppose $\hat{J}$ and $\bar{J}$ are the remaining vertices. Therefore

$$\begin{aligned}
\left[\mu_{\bar{J}}(X) + \varepsilon \lambda_{\bar{J}}(X) - \varepsilon \sum_{K \in L(\bar{J})} w_K - \varepsilon w_{\hat{J}}\right] r(\tilde{\mu}_{\bar{J}}) &= r(\tilde{\mu}_{\bar{J}}) \tilde{\mu}_{\bar{J}}(X) \\
&= r(\mu_{\bar{J}}) \mu_{\bar{J}}(X) + \varepsilon \int x \, \mathrm{d}\lambda_{\bar{J}} - \sum_{K \in L(\bar{J})} b_K - b_{\hat{J}} \\
&= r(\mu_{\bar{J}}) \mu_{\bar{J}}(X) + \varepsilon \int x \, \mathrm{d}\lambda - \varepsilon \sum_{i \notin \bar{J}} \gamma_i x_i,
\end{aligned}$$

where the first equality follows from the definition of π_K and the third equality follows from using (A2) repeatedly.

[11] This holds because $r(\mu_J) \in \mathrm{aff}\{x_i | i \in J\}$, $\tilde{\mu}_J(X) = \mu_J(X) + \varepsilon \sum_{i \in J} \gamma_i$ and $\int x \, \mathrm{d}\tilde{\mu}_J = \mu_J(X) r(\mu_J) + \varepsilon \sum_{i \in J} \gamma_i x_i$.

Since $\sum_{K\in L(\bar{J})} w_K + w_{\hat{J}} = \sum_{i\notin \bar{J}} \gamma_i - \sum_{J\neq \bar{J}} \lambda_J(X)$, and $\int x \,\mathrm{d}\lambda = \sum_i \gamma_i x_i$, this implies

$$r(\tilde{\mu}_{\bar{J}}) = \frac{1}{\mu_{\bar{J}}(X) + \varepsilon \sum_{i\in \bar{J}} \gamma_i} \left[\mu_{\bar{J}}(X) r(\mu_{\bar{J}}) + \varepsilon \sum_{i\in \bar{J}} \gamma_i x_i \right].$$

Since the barycenter of $\mu_{\bar{J}}$ lies in $\operatorname{aff}\{x_i | i \in \bar{J}\}$, we conclude that the barycenter of $\tilde{\mu}_{\bar{J}}$ lies in $\operatorname{aff}\{x_i | i \in \bar{J}\}$.

If $\bar{J}$ is the only remaining vertex, then $\sum_{K\in L(\bar{J})} w_K = \sum_{i\notin \bar{J}} \gamma_i - \sum_{J\neq \bar{J}} \lambda_J(X)$, which again implies

$$r(\tilde{\mu}_{\bar{J}}) = \frac{1}{\mu_{\bar{J}}(X) + \varepsilon \sum_{i\in \bar{J}} \gamma_i} \left[\mu_{\bar{J}}(X) r(\mu_{\bar{J}}) + \varepsilon \sum_{i\in \bar{J}} \gamma_i x_i \right].$$

It follows that the barycenter of $\tilde{\mu}_{\bar{J}}$ lies in $\operatorname{aff}\{x_i | i \in \bar{J}\}$. ■

Proof of Proposition 1. Let $\mathcal{P}$ be a partition of X such that each $A \in \mathcal{P}$ is convex, satisfies $\mu(A) > 0$, and $\nu|_A \precsim \mu|_A$. Moreover, assume there is no finer partition satisfying these properties.[12]

Suppose that for some $A \in \mathcal{P}$, the support of $\nu|_A$ is not affinely independent. Let $\{x_i | i \in I\}$ denote the support of $\nu|_A$. Also, let μ_i satisfy $r(\mu_i) = x_i$, $\mu_i(X) = \nu(\{x_i\})$, and $\mu|_A = \sum_i \mu_i$ (such μ_i exist by Lemma 1). Since $\mu_i \leq \mu|_A$ and μ is absolutely continuous, $\operatorname{aff}(\operatorname{supp} \mu_i) = \mathbb{R}^n$ for every $i \in I$.

The idea behind the iterative construction below is as follows. We seek a subset $C \subseteq I$ of support-point indices such that $\{x_i \mid i \in C\}$ is affinely dependent—providing a direction along which the weights $\nu(\{x_i\})$ can be perturbed—and such that the component measures $\{\mu_i\}_{i\in C}$ are 'connected' through chains of pairwise overlaps—two measures overlapping if the convex hulls of their supports have a common interior point—which enables the actual mass transfer. To find such a subset, we progressively merge groups of component measures whose aggregate supports have convex hulls with a common interior point. At each stage, the vertices of the new graph represent coarser clusters, and an edge indicates

[12]The trivial partition $\mathcal{P} = \{X\}$ satisfies all conditions. Moreover, there is an upper bound on the number of partition elements in any such partition since the support of ν is finite. Therefore, there is a partition that satisfies all conditions and cannot be refined.

that two clusters can exchange mass. We stop as soon as one of these clusters is large enough to be affinely dependent.

Construct a graph G_1 with vertex set I and an edge between i and j if

$$\operatorname{int}(\operatorname{ch}(\operatorname{supp}(\mu_i))) \cap \operatorname{int}(\operatorname{ch}(\operatorname{supp}(\mu_j))) \neq \emptyset.$$

Recursively, given a graph G_n construct a new graph G_{n+1} with vertex set being the set of connected components in G_n (for each connected component in G_n there is one vertex in G_{n+1}; each vertex is labeled by a subset of I corresponding to the set of vertices it represents). Add an edge between vertices C and C' in G_{n+1} if

$$\operatorname{int}(\operatorname{ch}(\bigcup_{i\in C}\operatorname{supp}(\mu_i))) \cap \operatorname{int}(\operatorname{ch}(\bigcup_{i\in C'}\operatorname{supp}(\mu_i))) \neq \emptyset.$$

Stop this procedure of constructing graphs once it yields a graph G_N with a vertex C such that $\{x_i | i \in C\}$ is affinely dependent.

We first claim that the procedure always stops with such a graph G_N. Suppose the procedure does not stop. Up to some iteration M, each iteration reduces the number of vertices by at least one; after that, the number of vertices stays constant, which implies that there are no edges. Let G_M denote this graph without edges. We claim that G_M has exactly one vertex. Suppose not and let $J_1,\dots,J_m$ with $m \geq 2$ denote its vertices. For $k = 1,\dots,m$, $A_k := \operatorname{ch}(\bigcup_{i\in J_k}\operatorname{supp}(\mu_i))$ is convex and satisfies $\mu(A_k) > 0$. Since there are no edges in G_M, $\operatorname{int}(A_j) \cap \operatorname{int}(A_k) = \emptyset$ for $j,k = 1,\dots,m$ with $j \neq k$. Since every hyperplane has μ-measure 0, this implies $\nu|_{A_k} \precsim \mu|_{A_k}$. This implies that the collection $\{A_1,\dots,A_m\}$ gives rise to a partition $\mathscr{P}'$ of X that is finer than $\mathscr{P}$, contradicting our initial hypothesis. We conclude that G_M has one vertex; since $\{x_i | i \in I\}$ is affinely dependent, this contradicts the hypothesis that the procedure did not stop.

This implies that vertex C of G_N corresponds to a connected component of G_{N-1}; let T be a spanning tree of this component. Every vertex J of T has $\{x_i | i \in J\}$ affinely independent, since otherwise the procedure would already have stopped at G_{N-1}. We claim we may choose T together with a leaf J of T so that *both* $\{x_i | i \in J\}$ and $\{x_i | i \in C \setminus J\}$ are affinely independent. To see this, pick any leaf J of T. If $\{x_i | i \in C \setminus J\}$ is affinely independent we are done. Otherwise, delete the leaf J from T and *replace C by $C \setminus J$*: the resulting tree spans a connected component of G_{N-1} whose vertex-union $C \setminus J$ is still affinely dependent, and we repeat the argument on it. Each step removes one vertex, so the procedure terminates; it must terminate with the desired leaf, since a tree with a single vertex has affinely independent support. Henceforth C, T and J denote the (possibly reduced) objects produced by this pruning, and the component measures μ_i for indices i removed from C are left unchanged in the construction below.

Now let $\{J_1, ..., J_m\}$ denote the vertices of T, choose a leaf J of T (without loss of generality, assume $J = J_1$), and let $\mu_{C\setminus J} = \sum_{i\in C\setminus J} \mu_i$. Since $\{x_i | i \in C\}$ is affinely dependent, there are coefficients $(\beta_i)_{i\in C}$, not all zero, such that $\sum_{i\in C} \beta_i = 0$ and $\sum_{i\in C} \beta_i x_i = 0$. This implies

$$s := \sum_{i\in J} \beta_i = -\sum_{i\in C\setminus J} \beta_i \quad \text{and} \quad d := \sum_{i\in J} \beta_i x_i = -\sum_{i\in C\setminus J} \beta_i x_i.$$

By Lemma 3 (applied with the s and d above) there is a fixed signed measure λ, independent of ε, with $\lambda(X) = s$ and $\int x \, \mathrm{d}\lambda = d$, such that for any $\varepsilon > 0$ small enough the measure $\pi := \varepsilon\lambda$ satisfies

$$\mu_J + \pi \geq 0, \quad \mu_{C\setminus J} - \pi \geq 0, \quad \pi(X) = \varepsilon s \quad \text{and} \int x \, \mathrm{d}\pi = \varepsilon d.$$

It follows that

$$r(\mu_J + \pi) = \frac{1}{\mu_J(X) + \pi(X)} \left[\int x \, \mathrm{d}\mu_J + \int x \, \mathrm{d}\pi \right] = \frac{1}{\mu_J(X) + \pi(X)} \Big[\mu_J(X) r(\mu_J) + \varepsilon d \Big]$$

$$r(\mu_{C\setminus J} - \pi) = \frac{1}{\mu_{C\setminus J}(X) - \pi(X)} \left[\int x \, \mathrm{d}\mu_{C\setminus J} - \int x \, \mathrm{d}\pi \right] = \frac{1}{\mu_{C\setminus J}(X) - \pi(X)} \Big[\mu_{C\setminus J}(X) r(\mu_{C\setminus J}) - \varepsilon d \Big].$$

Therefore, $r(\mu_J + \pi) \in \mathrm{aff}\{x_i | i \in J\}$ and $r(\mu_{C\setminus J} - \pi) \in \mathrm{aff}\{x_i | i \in C \setminus J\}$. Recall that $\pi = \varepsilon\lambda$ with λ independent of ε, so $\mu_{C\setminus J} - \pi = \mu_{C\setminus J} + \varepsilon(-\lambda)$. We apply Lemma 4 to the collection $\{\mu_i\}_{i\in C\setminus J}$, partitioned according to $\{J_2, \dots, J_m\}$, with the perturbation $-\lambda$: its hypothesis holds because $\mu_{C\setminus J} - \varepsilon\lambda \geq 0$ for all small ε and, taking $\gamma_i := \beta_i$ for $i \in C \setminus J$, we have $-\lambda(X) = -s = \sum_{i\in C\setminus J} \beta_i$ and $\int x \, \mathrm{d}(-\lambda) = -d = \sum_{i\in C\setminus J} \beta_i x_i$. Hence, for ε small enough, $\mu_{C\setminus J} - \pi = \sum_{k=2}^m \tilde{\mu}_{J_k}$ with each $\tilde{\mu}_{J_k} \geq 0$ and $r(\tilde{\mu}_{J_k}) \in \mathrm{aff}\{x_i | i \in J_k\}$. We decompose $\mu_J + \pi = \mu_J + \varepsilon\lambda$ in the same way, now with $\gamma_i = \beta_i$ for $i \in J$ (so that $\lambda(X) = s = \sum_{i\in J} \beta_i$ and $\int x \, \mathrm{d}\lambda = d = \sum_{i\in J} \beta_i x_i$). Applying these decompositions repeatedly to the graphs G_{N-2}, G_{N-3}, etc., yields measures $\{\tilde{\mu}_i\}_{i\in C}$ with $\sum_{i\in C} \tilde{\mu}_i = \sum_{i\in C} \mu_i$ and $r(\tilde{\mu}_i) = x_i$. We can then define $\tilde{\nu}$ to be a positive measure with the same support as $\nu|_A$ that satisfies $\tilde{\nu}(\{x_i\}) = \tilde{\mu}_i(X)$ for $i \in C$ and $\tilde{\nu}(\{x_i\}) = \nu(\{x_i\})$ for $i \in I \setminus C$. This is an MPC of $\mu|_A$ by Lemma 1.

Repeating the construction with $-\beta$ in place of β—equivalently, with $-s$ and $-d$ in place of s and d—yields a second MPC $\hat{\nu}$ of $\mu|_A$, again with $\hat{\nu}(\{x_i\}) = \nu(\{x_i\})$ for $i \in I \setminus C$.

We claim that $\tilde{\nu}(\{x_i\}) = \nu(\{x_i\}) + \varepsilon\beta_i$ for every $i \in C$. Indeed, on the leaf J the decomposition gives $\sum_{i\in J} \tilde{\mu}_i = \mu_J + \pi$, so

$$\sum_{i\in J} \tilde{\nu}(\{x_i\}) = \mu_J(X) + \varepsilon s \qquad \text{and} \qquad \sum_{i\in J} \tilde{\nu}(\{x_i\})\, x_i = \int x \, \mathrm{d}\mu_J + \varepsilon d.$$

Since $\{x_i | i \in J\}$ is affinely independent, the vectors $\{(x_i, 1) : i \in J\}$ are linearly independent, so this pair of (moment and mass) equations has a unique solution. As $\nu(\{x_i\}) + \varepsilon\beta_i$ solves it—using $\sum_{i\in J} \beta_i = s$ and $\sum_{i\in J} \beta_i x_i = d$, together with the fact that $\sum_{i\in J} \nu(\{x_i\})\delta_{x_i}$ is an MPC of μ_J—we obtain $\tilde{\nu}(\{x_i\}) = \nu(\{x_i\}) + \varepsilon\beta_i$ for $i \in J$. The same argument on $C \setminus J$ (with $\sum_{i\in C\setminus J} \beta_i = -s$ and $\sum_{i\in C\setminus J} \beta_i x_i = -d$) gives the identity for $i \in C \setminus J$. By symmetry of the construction, $\hat{\nu}(\{x_i\}) = \nu(\{x_i\}) - \varepsilon\beta_i$ for $i \in C$.

Consequently $\frac{1}{2}(\tilde{\nu} + \hat{\nu}) = \nu|_A$, while $\tilde{\nu} \neq \nu|_A$ because $\beta \neq 0$ implies $\tilde{\nu}(\{x_i\}) = \nu(\{x_i\}) + \varepsilon\beta_i \neq \nu(\{x_i\})$ for some $i \in C$. This implies that ν is not an extreme point of F_μ, contradicting our initial hypothesis. We conclude that $\mathrm{supp}(\nu|_A)$ is affinely independent for each $A \in \mathscr{P}$. ■

A.2 Proofs for Section 3

Proof of Proposition 2. Let G denote the set of probability measures that are MPC of μ and have at most K points in their support. We show first using Zorn's lemma that there is a maximal measure in G according to the convex order: The convex order is a partial order. Let $\{g_i\}_{i\in I}$ be a totally ordered subset of G, which can be viewed as a net by setting $i < j$ if $g_i \precsim g_j$. Since G is compact (e.g., in the weak*-topology), there is a subnet with limit g_∞. Then $g_i \precsim g_\infty$ because for any convex continuous function f and $i < j$, $\int f \, \mathrm{d}g_i \leq \int f \, \mathrm{d}g_j$. Therefore, every totally ordered subset of G has an upper bound in G. It follows from Zorn's lemma that there is a maximal element in G.

Let λ denote the maximal element of G, let $\{x_1, \ldots, x_L\}$ denote its support points, where $L \leq K$, and suppose λ is not convex partitional. Using Lemma 1 there is a decomposition of μ into positive measures $\{\lambda_i\}$ such that $\mu = \sum_{i=1}^{L} \lambda_i$, $r(\lambda_i) = x_i$, and $\lambda_i(X) = \lambda(\{x_i\})$. Since λ is not convex partitional, for some i, j, $\mathrm{int}(\mathrm{ch}(\mathrm{supp}(\lambda_i))) \cap \mathrm{int}(\mathrm{ch}(\mathrm{supp}(\lambda_j))) \neq \emptyset$. For z in this intersection, $x_j - x_i = (x_j - z) + (z - x_i) \in L_1 + L_2$ in Lemma 3, applied with $(\mu_1, \mu_2) = (\lambda_j, \lambda_i)$. Hence that lemma yields a measure α and $\varepsilon > 0$ such that $\alpha(X) = 0$, $-\lambda_j \leq \alpha \leq \lambda_i$, and $\int x \mathrm{d}\alpha = \varepsilon(x_j - x_i)$. If we fuse for each $k \neq i, j$ the measure λ_k to its barycenter and fuse $\lambda_i - \alpha \geq 0$ and $\lambda_j + \alpha \geq 0$ to their barycenters (if two of these measures have the same barycenter, decrease the value of ε slightly), we obtain an element of G that is larger than λ in the convex order, a contradiction. ∎

Proof of Proposition 3. Let $\bar{\nu} \in F_\mu$ be a finitely supported optimal solution. Let $S := \mathrm{supp}\, \bar{\nu}$, and let

$$\mathcal{O}_S := \Big\{\lambda \in F_\mu \colon \lambda \text{ is optimal and } \mathrm{supp}\, \lambda \subseteq S\Big\}.$$

As a subset of the finite-dimensional simplex of probability measures supported on S, $\mathcal{O}_S$ is nonempty, compact, and convex, so it has an extreme point; fix one and call it ν^*. We claim that ν^* is an extreme point of F_μ. Suppose that

$$\nu^* = t\lambda_1 + (1-t)\,\lambda_2$$

for some $t \in (0,1)$ and $\lambda_1, \lambda_2 \in F_\mu$. By the linearity of the objective, λ_1 and λ_2 are optimal. Since $\nu^*(X \setminus S) = 0$, we have

$$\lambda_1\,(X \setminus S) = \lambda_2\,(X \setminus S) = 0.$$

Thus, $\lambda_1, \lambda_2 \in \mathcal{O}_S$, and by the extremality of ν^* in $\mathcal{O}_S$, $\lambda_1 = \lambda_2 = \nu^*$. We conclude that ν^* is an extreme point of F_μ.

By Proposition 1, there exists a convex partition $\mathcal{P}^*$ of X corresponding to ν^*. Among all such partitions, choose $\mathcal{P}^*$ to be a finest up to μ-null refinements.[13]

Next, fix $P \in \mathcal{P}^*$ with $\mu(P) > 0$, and define

$$\tilde{\mu}_P := \mu(P)^{-1}\,\mu|_P \quad \text{and} \quad \tilde{\nu}_P := \mu(P)^{-1}\,\nu^*|_P.$$

[13]That is, there is no partition corresponding to ν^* that strictly refines $\mathcal{P}^*$ on a set of positive μ-measure.

By the global optimality of ν^*, the normalized restriction $\tilde{\nu}_P$ is optimal for the moment-persuasion problem on P with prior $\tilde{\mu}_P$, as otherwise we could strictly improve the global objective by modifying the signal only on P. Denote $S_P := \operatorname{supp} \tilde{\nu}_P$.

As $\tilde{\nu}_P \precsim \tilde{\mu}_P$, Strassen's theorem implies that there is a martingale coupling $\pi_P \in \Delta(P \times P)$ whose first marginal is $\tilde{\nu}_P$ and whose second marginal is $\tilde{\mu}_P$. Because the objective depends only on the first marginal and $\tilde{\nu}_P$ is optimal, any such martingale coupling π_P is an optimal solution to the normalized moment-persuasion problem. Applying Dworczak and Kolotilin (2024, Theorem 5) to this normalized problem on P, there exist a convex Lipschitz price function $\varphi_P \colon P \to \mathbb{R}$ and a measurable function $Q_P \colon P \to \mathbb{R}^d$ such that

$$\varphi_P(z) = \sup_{x \in P} \{V(x) + Q_P(x) \cdot (z - x)\}$$

for all $z \in P$, $\varphi_P \geq V$ on P, and the complementary-slackness condition

$$\varphi_P(z) = V(x) + Q_P(x) \cdot (z - x) \quad \text{for } \pi_P\text{-a.e. } (x, z). \tag{CS}$$

For each $x \in S_P$, define the affine function

$$\ell_x(z) := V(x) + Q_P(x) \cdot (z - x),$$

and set $\psi_P(z) := \max_{x \in S_P} \ell_x(z)$. As S_P is finite, ψ_P is convex and piecewise affine. Moreover, $\psi_P \leq \varphi_P$ on P. Complementary slackness (CS) implies that $\psi_P(z) = \varphi_P(z)$ for $\tilde{\mu}_P$-almost every $z \in P$. As both functions are continuous and $\tilde{\mu}_P$ has full support relative to P (after modifying cells on μ-null boundaries), we have $\psi_P = \varphi_P$ on P and so $\psi_P \geq V$ on P.

We also have $\psi_P(x) = V(x)$ for every $x \in S_P$. Indeed, since $\psi_P = \varphi_P$, $\psi_P \geq V$. On the other hand,

$$\int_P \psi_P \, d\tilde{\nu}_P = \int_P \varphi_P \, d\tilde{\nu}_P \leq \int_P \varphi_P \, d\tilde{\mu}_P = \int_P V \, d\tilde{\nu}_P,$$

where the inequality follows from $\tilde{\nu}_P \precsim \tilde{\mu}_P$ and the convexity of φ_P, and the final equality is the primal-dual value equality from (Dworczak and Kolotilin, 2024, Theorem 5) applied to the normalized problem on P. Since $\psi_P \geq V$ and $\tilde{\nu}_P$ is supported on S_P, equality of the integrals implies $\psi_P(x) = V(x)$ for every $x \in S_P$.

We claim that ψ_P is affine on P. Suppose not. For each $x \in S_P$, define

$$C_x := \left\{z \in P \colon \psi_P(z) = \ell_x(z)\right\}.$$

The sets C_x, $x \in S_P$, are convex and cover P. As $\psi_P = \varphi_P$, (CS) implies that the conditional distribution of z given x under π_P is supported on C_x. Since x is the barycenter of this conditional distribution and C_x is convex, we have $x \in C_x$. Moreover, because $x \in S_P$ has positive $\tilde{\nu}_P$-mass and $\tilde{\mu}_P$ is absolutely continuous, C_x has nonempty relative interior in P.

If two such sets have overlapping relative interiors, then the corresponding affine functions agree on a relative open set and, therefore, agree everywhere. Thus, after identifying support points that generate the same affine function and assigning common boundaries, which are μ-null, the distinct sets C_x define a finite convex partition of P, where we assign the boundary pieces so that each support point $x \in S_P$ belongs to the cell generated by C_x.

For each cell R of this partition, let $S_R := \{x \in S_P \colon C_x = R\}$. The conditional distribution of z given any $x \in S_R$ is supported on R, and x is the barycenter of this conditional distribution. Aggregating over $x \in S_R$, the restriction $\tilde{\nu}_P|_R$ is a mean-preserving contraction of $\tilde{\mu}_P|_R$. Scaling by $\mu(P)$, replacing P by these cells yields a convex partition corresponding to ν^*. Since ψ_P is not affine on P, this partition is strictly finer than $\mathscr{P}^*$, a contradiction. Therefore, ψ_P is affine on P.

In sum, ψ_P is affine, $\psi_P \geq V$ on P, and $\psi_P = V$ on S_P. For every $\eta \in D_{\mu,P}$, both η and $\tilde{\nu}_P$ have barycenter $r_P(\mu)$. Therefore,

$$\mathbb{E}_\eta[V] \leq \mathbb{E}_\eta\left[\psi_P\right] = \psi_P(r_P(\mu)) = \mathbb{E}_{\tilde{\nu}_P}\left[\psi_P\right] = \mathbb{E}_{\tilde{\nu}_P}[V],$$

whence we conclude that $\tilde{\nu}_P$ solves the unconstrained persuasion problem on P; and, as $P \in \mathscr{P}^*$ was arbitrary, ν^* is canonical. We obtain the stated payoff by summing the values of the unconstrained problems for each cell: $\mathbb{E}_{\nu^*}[V] = \sum_{P \in \mathscr{P}^*} \mu(P) \operatorname{cav}_P V(r_P(\mu))$. ∎

A.3 Proofs for Section 4

Proof of Proposition 4. Fix a decision problem, cost function, and prior of the form described above. The strict convexity of κc means that any optimal solution must be finitely supported. Moreover, if V is affine, then the unique optimum is degenerate on the prior's barycenter $\delta_{r(\mu)}$. That solution is convex partitional with a single prototype, so the conclusion is immediate. We therefore assume that V is not affine.

For each $j = 1,\ldots,d$, let

$$c_j(x) := \frac{\partial c}{\partial x_j}(x),$$

and define

$$\alpha := \max_{x,y\in X;j=1,\ldots,d} \left|c_j(x) - c_j(y)\right|.$$

Because the partial derivatives of c are bounded and c is strictly convex, $\alpha \in \mathbb{R}_{++}$.

We write $z_a(x) = b_a \cdot x + d_a$ for each $a \in A$, with $b_a = (b_{a,1},\ldots,b_{a,d})$. Define

$$\beta := \min\left\{\left|b_{a,j} - b_{a',j}\right| : a,a' \in A,\ j = 1,\ldots,d,\ b_{a,j} \neq b_{a',j}\right\}.$$

Because V is not affine and because the actions in A are undominated, the finite set of slope vectors $\{b_a\}_{a\in A}$ contains at least two distinct elements and $\beta \in \mathbb{R}_{++}$.

Let $\bar{\kappa} := \beta/(2\alpha)$, and fix $\kappa \leq \bar{\kappa}$. Since $V - \kappa c$ is Lipschitz on X and the optimum is finitely supported, Proposition 3 applies and so there exists a canonical solution ν with corresponding partition $\mathscr{P}$. Suppose for the sake of contradiction that for some $P \in \mathscr{P}$, the restriction $\nu|_P$ has support $\{x_1,\ldots,x_m\}$ with $m \geq 2$.

Since ν is canonical, the normalized restriction $\mu(P)^{-1}\,\nu|_P$ solves the unconstrained persuasion problem on P for the objective $V - \kappa c$. Accordingly, there exists an affine function $h\colon P \to \mathbb{R}$ such that $h \geq V - \kappa c$ on P and

$$h(x_i) = V(x_i) - \kappa c(x_i)$$

for each $i = 1,\ldots,m$.

For each i, choose $a_i \in A$ such that $V(x_i) = z_{a_i}(x_i)$. Then,

$$\frac{z_{a_i}(x) - h(x)}{\kappa} \leq c(x)$$

for all $x \in P$, with equality at x_i. Since c is differentiable, the affine function

$$x \mapsto \frac{z_{a_i}(x) - h(x)}{\kappa}$$

is a supporting hyperplane to c at x_i. If $a_i = a_k$ for all i,k, then the same affine function

$$x \mapsto \frac{z_{a_1}(x) - h(x)}{\kappa}$$

would support c at the distinct points $x_1, \ldots, x_m$, contradicting the strict convexity of c.

Thus, there exist $i,k \in \{1,\ldots,m\}$ such that $a_i \neq a_k$. For such i and k,

$$b_{a_i} - b_{a_k} = \kappa(\nabla c(x_i) - \nabla c(x_k)).$$

Choose $j \in \{1,\ldots,d\}$ such that $b_{a_i,j} \neq b_{a_k,j}$. Then

$$\kappa = \frac{\left|b_{a_i,j} - b_{a_k,j}\right|}{\left|c_j(x_i) - c_j(x_k)\right|} \geq \frac{\beta}{\alpha} > \bar{\kappa},$$

contradicting $\kappa \leq \bar{\kappa}$. We conclude that every cell $P \in \mathscr{P}$ contains a single support point of ν, as desired. ■

Proof of Proposition 5. The first statement follows from Proposition 2 and the convexity of V, so only the second statement needs proving. Let $A_0 \subseteq A$ be a set of K actions that are strictly optimal somewhere. For $a \in A_0$, define

$$U_a := \{x \in X : z_a(x) > z_b(x) \quad \forall\, b \neq a\}.$$

Each U_a is relatively open and convex. Since a is strictly optimal somewhere and μ has full support, $\mu(U_a) > 0$.

First, we observe that every optimum has exactly K support points. Indeed, suppose for the sake of contradiction that ν is optimal and $|\operatorname{supp}\nu| < K$. Since the sets U_a, $a \in A_0$, are pairwise disjoint and there are K of them, there is some $a \in A_0$ such that $U_a \cap \operatorname{supp}\nu = \emptyset$. If π is a martingale coupling with first marginal ν and second marginal μ, then for some $y \in \operatorname{supp}\nu$, the conditional distribution π_y assigns mass $p \in (0,1)$ to U_a. Replacing y by the two barycenters $y_a = r(\pi_y|_{U_a}) \in U_a$ and $y_R = r(\pi_y|_{X \setminus U_a})$, with weights $p\nu(\{y\})$ and $(1-p)\nu(\{y\})$, yields an MPC with at most $|\operatorname{supp}\nu| + 1 \leq K$ support points and a strictly higher payoff, because a is uniquely optimal at y_a, while the affinity of payoffs implies that a is not optimal at y_R. This contradicts the optimality of ν.

Second, suppose for the sake of contradiction that some optimal ν with $|\operatorname{supp}\nu| \leq K$ is not convex partitional. As we have just argued, it must be that $|\operatorname{supp}\nu| = K$. Furthermore

Proposition 2 delivers a convex-partitional MPC λ of μ, with $|\operatorname{supp}\lambda| \leq K$, such that $\nu \prec \lambda$. Since V is convex, $\mathbb{E}_\lambda[V] \geq \mathbb{E}_\nu[V]$. If the inequality were strict, ν would not be optimal. Thus, equality must hold, and λ is also optimal. Moreover, $|\operatorname{supp}\lambda| = K$.

Let γ be a martingale coupling with first marginal ν and second marginal λ. Since $\nu \prec \lambda$, some conditional distribution γ_y is nondegenerate. Moreover,

$$0 = \mathbb{E}_\lambda[V] - \mathbb{E}_\nu[V] = \sum_{y \in \operatorname{supp}\nu} \nu(\{y\}) \left[\int V(z)\, d\gamma_y(z) - V(y) \right],$$

and each term in brackets is nonnegative by Jensen's inequality. Therefore, Jensen's inequality holds with equality for the nondegenerate conditional distribution γ_y, so V is affine on $\operatorname{ch}\operatorname{supp}\gamma_y$. In particular, there exist two distinct points $z_1, z_2 \in \operatorname{supp}\lambda$ such that V is affine on the segment $[z_1, z_2]$.

Let $m_i = \lambda(\{z_i\})$, $i = 1, 2$, and define

$$\bar{z} := \frac{m_1 z_1 + m_2 z_2}{m_1 + m_2}.$$

Let ρ be obtained from λ by replacing the two atoms z_1, z_2 by the single atom $\bar{z}$ with mass $m_1 + m_2$. Then $\rho \precsim \lambda \precsim \mu$ so $\rho \in F_\mu$, and $|\operatorname{supp}\rho| < K$.

Since V is affine on $[z_1, z_2]$, $\mathbb{E}_\rho[V] = \mathbb{E}_\lambda[V]$, and so ρ is an optimal solution with fewer than K support points, a contradiction. Therefore, every optimal solution must be convex partitional. ■

B Sufficient Conditions for Extremal MPCs

We explore here sufficient conditions for a finitely supported measure ν to be an extreme point of F_μ.

Definition B.1. Given a positive measure μ and a partition $\mathcal{P}$ of X into finitely many convex subsets, we say that a collection of (signed) measures $\{u_P\}_{P\in\mathcal{P}}$ is a *feasible flow for*

$\mathcal{P}$ if it satisfies

$$0 \le u_P|_{X\setminus P} \le \mu|_{X\setminus P} \tag{B.1}$$

$$-\mu|_P \le u_P|_P \le 0 \tag{B.2}$$

$$u_P(X) = 0 \tag{B.3}$$

$$\int_X x d u_P = 0 \tag{B.4}$$

$$\sum_P u_P = 0 \tag{B.5}$$

Proposition 6. *Fix an absolutely continuous measure μ and let $\mathcal{P}$ be a partition of X into finitely many convex sets.*

1. *Suppose that $\nu \in F_\mu$ satisfies $\nu|_P \precsim \mu|_P$ for all $P \in \mathcal{P}$ and that the support of $\nu|_P$ is affinely independent. If $u_P \equiv 0$ for all $P \in \mathcal{P}$ is the unique feasible flow for $\mathcal{P}$ then ν is an extreme point of F_μ.*

2. *Suppose there exists a non-zero feasible flow for $\mathcal{P}$. Then there exists an MPC $\nu \in F_\mu$ that is not an extreme point of F_μ and that satisfies*

   (a) *For each $P \in \mathcal{P}$, $\nu|_P \precsim \mu|_P$; and*

   (b) *The support of $\nu|_P$ is affinely independent.*

The proof of Part 2 uses the following lemma.

Lemma 5. *Let μ be absolutely continuous and let $P \subseteq X$ be convex with $\mu(P) > 0$. Then there is $\bar{\delta} > 0$ such that for every $\delta \in (0, \bar{\delta})$ there are affinely independent points $y_1, \ldots, y_{n+1} \in$ $\mathrm{int}(P) \cap B_\delta(r(\mu|_P))$ with the following property. For every signed measure w of finite total variation such that $\mu|_P + \varepsilon w \ge 0$ for all sufficiently small $\varepsilon > 0$, there is $\bar{\varepsilon} > 0$ such that for all $\varepsilon \in (0, \bar{\varepsilon})$ one can write $\mu|_P + \varepsilon w = \sum_{j=1}^{n+1} \lambda_j^\varepsilon$ with each λ_j^ε positive and $r(\lambda_j^\varepsilon) = y_j$. Consequently the unique measure supported on $\{y_1, \ldots, y_{n+1}\}$ with the same mass and barycenter as $\mu|_P + \varepsilon w$ is positive and is an MPC of $\mu|_P + \varepsilon w$.*

Proof. Let $M := \mu(P)$, $z := r(\mu|_P)$ and $\Sigma_P(z, \mu) := \int_P (x - z)(x - z)^\top \, \mathrm{d}\mu(x)$. Thus, $\Sigma_P(z, \mu)$ is the unnormalized covariance matrix of $\mu|_P$.

Since μ is absolutely continuous and $\mu(P) > 0$, the measure $\mu|_P$ is not carried by a hyperplane. Hence P has nonempty interior and the matrix $\Sigma_P(z,\mu)$ is positive definite. Moreover $z \in \operatorname{int}(\operatorname{ch}(\operatorname{supp}(\mu|_P)))$: otherwise a supporting hyperplane at z with normal a would give $\langle a, x-z\rangle \geq 0$ on $\operatorname{supp}(\mu|_P)$ while $\int \langle a, x-z\rangle \,\mathrm{d}\mu|_P = 0$, forcing $\mu|_P$ to be carried by that hyperplane. Since $\operatorname{ch}(\operatorname{supp}(\mu|_P)) \subseteq \bar{P}$ and $\operatorname{int}(\bar{P}) = \operatorname{int}(P)$ for a convex set with nonempty interior, $z \in \operatorname{int}(P)$.

Let $v_1, \dots, v_{n+1}$ be the vertices of a regular simplex in $B_1(0) \subseteq \mathbb{R}^n$ centered at the origin, so that $\sum_j v_j = 0$ and $z + v_1, \dots, z + v_{n+1}$ are affinely independent. For $\delta > 0$ set

$$y_j := z + \delta v_j, \qquad \psi_j(x) := \tfrac{1}{n+1} + \tfrac{M\delta}{n+1}\Big\langle \Sigma_P(z,\mu)^{-1} v_j,\, x - z\Big\rangle, \qquad \lambda_j := \psi_j \cdot \mu|_P.$$

Since $\sum_j v_j = 0$ we have $\sum_j \psi_j \equiv 1$ and hence $\sum_j \lambda_j = \mu|_P$. Since $\int (x-z)\,\mathrm{d}\mu|_P = 0$ we have $\lambda_j(X) = M/(n+1)$, and since $\int (x-z)\langle b, x-z\rangle \,\mathrm{d}\mu|_P = \Sigma_P(z,\mu) b$ for every $b \in \mathbb{R}^n$,

$$\int x \,\mathrm{d}\lambda_j = \tfrac{M}{n+1} z + \Sigma_P(z,\mu)\Big(\tfrac{M\delta}{n+1} \Sigma_P(z,\mu)^{-1} v_j\Big) = \tfrac{M}{n+1}\,(z + \delta v_j),$$

so that $r(\lambda_j) = y_j$. Because X is compact, $|\psi_j(x) - \frac{1}{n+1}| \leq \frac{M\delta}{n+1} \|\Sigma_P(z,\mu)^{-1}\| \max_j \|v_j\| \operatorname{diam}(X)$ for all $x \in X$. Hence there is $\bar{\delta} > 0$ such that, for all $\delta \in (0, \bar{\delta})$ and all j, $y_j \in \operatorname{int}(P)$ and $\psi_j \geq \frac{1}{2(n+1)}$, so that $\lambda_j \geq \frac{1}{2(n+1)} \mu|_P$. This proves the assertion for $w = 0$.

Fix such a δ and let w be as in the statement. Since $\lambda_j \geq \frac{1}{2(n+1)} \mu|_P$ for every j, all the λ_j have the same support as $\mu|_P$ and hence $\operatorname{aff}(\operatorname{supp} \lambda_j) = \mathbb{R}^n$. Consequently the graph in Lemma 4, for the singleton partition, is complete and connected. Furthermore, $\{y_1, \dots, y_{n+1}\}$ is affinely independent and therefore spans $\operatorname{aff}(P) = \mathbb{R}^n$, so the vectors $(1, y_j) \in \mathbb{R}^{n+1}$ form a basis and there are reals $\gamma_1, \dots, \gamma_{n+1}$ with $w(X) = \sum_j \gamma_j$ and $\int x \,\mathrm{d}w = \sum_j \gamma_j y_j$. Lemma 4 thus applies and yields, for all $\varepsilon > 0$ small enough, positive measures λ_j^ε with $\sum_j \lambda_j^\varepsilon = \mu|_P + \varepsilon w$ and $r(\lambda_j^\varepsilon) \in \operatorname{aff}\{y_j\} = \{y_j\}$.

Finally, let ν^ε be the measure supported on $\{y_1, \dots, y_{n+1}\}$ with $\nu^\varepsilon(\{y_j\}) = \lambda_j^\varepsilon(X)$. By Lemma 1, $\nu^\varepsilon \precsim \mu|_P + \varepsilon w$, and by construction ν^ε has the same mass and barycenter as $\mu|_P + \varepsilon w$. As $\{y_1, \dots, y_{n+1}\}$ is affinely independent, it is the only measure supported on that set with this mass and barycenter. ∎

Proof of Proposition 6. Part 1: Assume ν is not an extreme point. Then there are $\nu_1, \nu_2 \in F_\mu$ such that $\nu_1 \neq \nu_2$ and $\nu = \frac{1}{2}\nu_1 + \frac{1}{2}\nu_2$. Let $\{x_1, ..., x_m\}$ denote the support of ν and let μ_j^i

be positive measures such that, for $i = 1, 2$, $\mu = \sum_j \mu_j^i$, $\nu_i(\{x_j\}) = \mu_j^i(X)$, and $r(\mu_j^i) = x_j$ (such measures exist by Lemma 1). Define $\mu_P^i := \sum_{j: x_j \in P} \mu_j^i$ and $u_P := \frac{1}{2}\mu_P^1 + \frac{1}{2}\mu_P^2 - \mu|_P$. It follows that $\sum_P u_P = 0$. Moreover, u_P is positive on $X \setminus P$ and, since $\mu_P^i \le \mu$, it satisfies (B.1). Analogous arguments establish (B.2). Finally, $u_P(X) = 0$ and $\int x \, \mathrm{d}u_P = 0$. Since $\nu_1 \neq \nu \neq \nu_2$, $u_P \neq 0$ for some P. Hence, there is a non-zero feasible flow.

Part 2: Conversely, suppose there is a solution $\{u_P\}$ to (B.1)–(B.5) satisfying $u_Q \neq 0$ for some $Q \in \mathscr{P}$. Enumerate the elements of $\mathscr{P}$ as $\mathscr{P} = \{P_1, ..., P_m\}$. Note first that $\mu(P) > 0$ whenever $u_P \neq 0$: if $\mu(P) = 0$, then (B.2) forces $u_P|_P = 0$, and since $u_P(X) = 0$ and $u_P|_{X \setminus P} \geq 0$ by (B.1), also $u_P|_{X \setminus P} = 0$. In particular $\mu(Q) > 0$.

We first construct ν. For each $P \in \mathscr{P}$ with $\mu(P) > 0$, let $\bar{\delta}_P > 0$ be as in Lemma 5, fix $\delta \in \left(0, \min\{\bar{\delta}_P : \mu(P) > 0\}\right)$, and let $y_1^P, ..., y_{n+1}^P$ be the points that the Lemma associates with P and δ. Let $\nu|_P$ be the measure supported on $\{y_1^P, ..., y_{n+1}^P\}$ with the same mass and barycenter as $\mu|_P$, and let $\nu|_P := 0$ if $\mu(P) = 0$. By the case $w = 0$ of Lemma 5, $\nu|_P$ is a positive measure satisfying $\nu|_P \precsim \mu|_P$, so $\nu := \sum_{P \in \mathscr{P}} \nu|_P$ is an *MPC* of μ; moreover, the support of $\nu|_P$ is affinely independent, spans the affine hull of P, and is contained in the δ-ball around the barycenter of $\nu|_P$. Note that δ is fixed at this point; the parameter ε below is chosen small only afterwards.

Choose $\varepsilon \in (0, 1)$ small enough and define, for $i = 1, ..., m$,

$$\begin{aligned} \mu_{P_i}^i &:= \mu|_{P_i} + \varepsilon u_{P_i}|_{X \setminus P_i} \\ \mu_P^i &:= \mu|_P - \varepsilon u_{P_i}|_P \text{ for } P \neq P_i. \end{aligned}$$

By construction, μ_P^i is a positive measure for all $P \in \mathscr{P}$ and $\sum_{P \in \mathscr{P}} \mu_P^i = \mu$. Write $\mu_P^i = \mu|_P + \varepsilon w_P^i$, where $w_{P_i}^i := u_{P_i}|_{X \setminus P_i} \geq 0$ and $w_P^i := -u_{P_i}|_P \geq -\mu|_P$ for $P \neq P_i$, both inequalities by (B.1). Hence $\mu|_P + \varepsilon w_P^i \geq 0$ for all $\varepsilon \in (0, 1)$, so Lemma 5 applies to each of the finitely many pairs (P, w_P^i): for ε small enough, the measure ν_P^i with support $\mathrm{supp}(\nu) \cap P$ and the same mass and barycenter as μ_P^i is well defined, positive, and an *MPC* of μ_P^i. Consequently $\nu^i := \sum_{P \in \mathscr{P}} \nu_P^i$ is an *MPC* of μ.

We claim that $\frac{1}{m} \sum_{i=1}^m \nu^i = \nu$, which implies that ν is not an extreme point. To establish the claim, we argue that $\frac{1}{m} \sum_{i=1}^m \nu^i(P) = \nu(P)$ and $\frac{1}{m} \sum_{i=1}^m \int_P x d\nu^i = \int_P x d\nu$ for all $P \in \mathscr{P}$. Since the support of ν^i and ν coincides on each P and is affinely independent, this implies that $\frac{1}{m}\sum_{i=1}^{m}\nu^i = \nu$. That is,

$$
\begin{aligned}
\frac{1}{m}\sum_{i=1}^{m}\nu^i(P_j) &= \mu(P_j) + \frac{1}{m}\varepsilon u_{P_j}|_{X\setminus P_j}(X) - \frac{1}{m}\varepsilon\sum_{i:i\neq j} u_{P_i}|_{P_j}(X) \\
&= \mu(P_j) + \frac{1}{m}\varepsilon u_{P_j}|_{X\setminus P_j}(X) + \frac{1}{m}\varepsilon u_{P_j}|_{P_j}(X) \text{ (since } \sum_i u_{P_i}|_{P_j} = 0 \text{)} \\
&= \mu(P_j) + \frac{1}{m}\varepsilon u_{P_j}(X) \\
&= \nu(P_j)
\end{aligned}
$$

Also,

$$
\begin{aligned}
\frac{1}{m}\sum_{i=1}^{m}\int_{P_j} x d\nu^i &= \frac{1}{m}\sum_{i=1}^{m}\int_{P_j} x d\mu^i_{P_j} \\
&= \int_{P_j} x d\mu + \varepsilon/m \int_{X\setminus P_j} x d u_{P_j} - \varepsilon/m \sum_{i:i\neq j}\int_{P_j} x d u_{P_i} \\
&= \int_{P_j} x d\mu + \varepsilon/m \int_{X\setminus P_j} x d u_{P_j} + \varepsilon/m \int_{P_j} x d u_{P_j} \\
&= \int_{P_j} x d\nu.
\end{aligned}
$$

■

Lemma 6. *Suppose that $P \in \mathcal{P}$ is the intersection of a half-space and X, i.e., $P = H \cap X$ for some half-space H of $\mathbb{R}^n$. If $\{u_P\}$ is a feasible flow then $u_P|_{P'} = u_{P'}|_P = 0$ for all $P' \neq P$.*

Proof. If $\{u_P\}$ is a feasible flow then $\int x\,\mathrm{d}u_P = 0$ and therefore $\int x\,\mathrm{d}[-u_P|_P] = \sum_{P'\neq P}\int x\,\mathrm{d}[u_P|_P']$. The result follows since if $u_P|_{P'} \neq 0$ then the barycenter of $-u_P|_P$ lies in the interior of P, hence in the interior of the half-space H, whereas the barycenter of $\sum_{P'\neq P} u_P|_{P'}$ lies in the complement of this half-space. This violates $\int x\,\mathrm{d}u_P = 0$, therefore $u_P|_{P'} = 0$. ■

Lemma 7. *Let μ be an absolutely continuous measure with full support on X.*

Let $\mathcal{P}$ be a partition of X into finitely many convex sets and suppose that there is a non-zero feasible flow $\{u_P\}_{P\in\mathcal{P}}$ for $\mathcal{P}$. Let $m_{PQ} := u_P|_Q(X)$ and $x_{PQ} := \int x d u_P|_Q$. For all $P, Q \in \mathcal{P}$ with $m_{PQ} > 0$, let $b_{PQ} := \frac{x_{PQ}}{m_{PQ}}$.

If $\mathscr{P}'$ is a partition of X that approximates $\mathscr{P}$ in the sense that for each $P \in \mathscr{P}$ there is $P' \in \mathscr{P}'$ such that $P' \subseteq P$ and, for all $Q \in \mathscr{P}$, $b_{QP} \in \operatorname{int} P'$ then there is a non-zero feasible flow for $\mathscr{P}'$.

Proof. Let k denote the number of partition elements in $\mathscr{P}$ and fix $\delta > 0$ small enough. For all $P \in \mathscr{P}$ and all $Q \in \mathscr{P} \setminus \{P\}$, let P' and Q' be the corresponding partition elements in $\mathscr{P}'$ that approximate P and Q, respectively. Let $\tilde{u}_{Q'P'}$ be a non-negative measure with $\tilde{u}_{Q'P'} \leq \frac{1}{k}\mu|_{P'}$, $\tilde{u}_{Q'P'}(X) = \delta m_{QP}$, and barycenter $\frac{1}{\delta m_{QP}} \int x d\tilde{u}_{Q'P'} = b_{QP}$ (if $m_{QP} > 0$). Such measures exist by Lemma 2 whenever $\delta > 0$ is chosen small enough. For all $P', Q' \in \mathscr{P}'$ for which $\tilde{u}_{P'Q'}$ has not been defined this way, let $\tilde{u}_{P'Q'}$ be the zero measure.

For any $P' \in \mathscr{P}'$, define $\tilde{u}_{P'P'} = -\sum_{Q' \in \mathscr{P}' \setminus \{P'\}} \tilde{u}_{Q'P'}$ and define $\tilde{u}_{P'} = \sum_{Q' \in \mathscr{P}'} \tilde{u}_{P'Q'}$.

Then $\sum_{P' \in \mathscr{P}'} \tilde{u}_{P'} = \underline{0}$ by construction. Moreover, since $\sum_{Q \in \mathscr{P}} u_Q|_P$ is the zero measure, we obtain for any $P \in \mathscr{P}$ and its approximating partition cell $P' \in \mathscr{P}'$,

$$\delta u_P|_P(X) = -\delta \sum_{Q \in \mathscr{P} \setminus \{P\}} u_Q|_P(X) = -\sum_{Q' \in \mathscr{P}' \setminus \{P'\}} \tilde{u}_{Q'P'}(X) = \tilde{u}_{P'P'}(X).$$

This yields

$$\tilde{u}_{P'}(X) = \sum_{Q' \in \mathscr{P}'} \tilde{u}_{P'Q'}(X) = \sum_{Q \in \mathscr{P}} \delta u_P|_Q(X) = \delta u_P(X) = 0.$$

Analogous arguments establish that $\int x d\tilde{u}_{P'} = 0$. Since $0 \leq \tilde{u}_{P'}|_{X \setminus P'} \leq \mu|_{X \setminus P'}$ and $-\mu|_{P'} \leq \tilde{u}_{P'}|_{P'} \leq 0$, it follows that $\{\tilde{u}\}$ is a non-zero feasible flow for $\mathscr{P}'$. ■

References

Itai Arieli, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita. Optimal persuasion via bi-pooling. *Theoretical Economics*, 18(1):15–36, 2023.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953.

Pierre Cartier, J. M. G. Fell, and Paul-André Meyer. Comparaison, des mesures portées par un ensemble convexe compact. *Bulletin de la Société Mathématique de France*, 92: 435–445, 1964.

Piotr Dworczak and Anton Kolotilin. The persuasion duality. *Theoretical Economics*, 19 (4):1701–1755, 2024.

J. Elton and T. P. Hill. Fusions of a probability distribution. *The Annals of Probability*, 20 (1):421–454, 1992.

Peter Gärdenfors. *Conceptual Spaces: The Geometry of Thought*. A Bradford book. MIT Press, 2004.

Maxim Ivanov. Optimal monotone signals in Bayesian persuasion mechanisms. *Economic Theory*, 72(3):955–1000, 2021.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

Andreas Kleiner, Benny Moldovanu, and Philipp Strack. Extreme points and majorization: Economic applications. *Econometrica*, 89(4):1557–1593, 2021.

Jeffrey Mensch and Komal Malik. Posterior-mean separable costs of information acquisition. *arXiv preprint arXiv:2311.09496*, 2023.

A. Roy. Extremal dilations and their associated extremal measures. *Journal of Functional Analysis*, 22:8–31, 1976.

Murat Sertel. On the continuity of closed convex hull. *Mathematical Social Sciences*, 18(3): 297–299, 1989.

Volker Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, 1965.

Frank Yang and Kai Hao Yang. Stochastic optimization and coupling. *arXiv preprint arXiv:2603.11448*, 2026.