# Simplifying Bayesian Persuasion

Elliot Lipnowski[*][†]
*University of Chicago*

Laurent Mathevet
*New York University*

March 22, 2017

**Abstract**

In Bayesian Persuasion (Kamenica and Gentzkow (2011)), the sender's optimal value is characterized by a concave envelope. Since concavification of a general function is notoriously difficult, we propose a method to reduce the problem, using the underlying economic structure of the indirect expected utility. The key observation is that one can find, using the receiver's preferences alone, a small set of posterior beliefs on which some optimal information policy must be supported. This simplifies, sometimes dramatically, the search for optimal information.

*Keywords*: disclosure, posterior cover, outer points, optimal information policy.

*JEL Classification*: D83, D82.

[*]E-mail: `lipnowski@uchicago.edu` and `lmath@nyu.edu`.

[†]We thank an anonymous referee (of another paper), whose comments motivated us to write this paper.

In the model of Bayesian persuasion (Kamenica and Gentzkow (2011)), a receiver must make a decision $a \in A$ in a world where the uncertain state $\omega \in \Omega$ is distributed according to (full-support) $\mu_0 \in \Delta\Omega$. Assume $\Omega$ and $A$ are finite sets. Before making his decision, the receiver obtains additional information about the state from a sender who chooses a signal structure $\sigma : \Omega \to \Delta S$, where $S$ is a rich finite set of messages.[1] In any state $\omega$, the receiver sees a message $s \in S$ drawn according to $\sigma(\cdot|\omega)$, forms a posterior belief $\mu \in \Delta\Omega$ via Bayesian updating, and then acts. The receiver and the sender have preferences given by $u, v : A \times \Omega \to \mathbb{R}$, respectively.

An **(information) policy** is an element of

$$\mathcal{R}(\mu_0) \quad = \quad \left\{ p \in \Delta\Delta\Omega : \quad \int_{\Delta\Omega} \mu(\omega) \, \mathrm{d}p(\mu) = \mu_0(\omega) \text{ for every } \omega \in \Omega \right\}.$$

Any signal induces an information policy by Bayes plausibility, and, conversely, any information policy can be induced by some signal (Aumann and Maschler (1995) and Kamenica and Gentzkow (2011)). Let $A^* : \Delta\Omega \rightrightarrows A$ be the agent's best response correspondence,[2] and let $V : \Delta\Omega \to \mathbb{R}$ be the principal's indirect value function, given by $V(\mu) = \max_{a \in A^*(\mu)} \int_\Omega v(a, \cdot) \, \mathrm{d}\mu$. A policy $p$ is **optimal** if

$$\int_{\Delta\Omega} V \, \mathrm{d}p \geq \int_{\Delta\Omega} V \, \mathrm{d}q$$

for all $q \in \mathcal{R}(\mu_0)$. The concavification result in Kamenica and Gentzkow (2011) provides an abstract characterization of the sender's optimal value. For any prior $\mu_0$, an optimal information policy exists and induces expected indirect utility

$$\overline{V}(\mu_0) = \inf\{\phi(\mu_0) : \quad \phi : \Delta\Omega \to \mathbb{R} \text{ affine}, \phi \geq V\},$$

which is the concave envelope of $V$ (i.e., the pointwise lowest concave function which majorizes $V$), evaluated at the prior.

The purpose of this short paper is to simplify the search for optimal information policies in Bayesian persuasion. If the sender cannot compute the concave envelope or derive qualitative properties of it, then the concavification result cannot be implemented. In general, computing the concave envelope of a function is difficult (Tardella (2008)). In Lipnowski and Mathevet (2017), we develop a method to simplify the computation of an optimal policy in environments with psychological preferences and aligned interests. In this paper, we adapt

[1] In light of Kamenica and Gentzkow (2011, Proposition 1), assume $|S| \geq \min\{|A|, |\Omega|\}$.

[2] So let $A^*(\mu) = \mathrm{argmax}_{a \in A} \int_\Omega u(a, \cdot) \, \mathrm{d}\mu$ be the receiver's set of optimal actions at posterior $\mu$.

our method to standard Bayesian persuasion—where the receiver and the sender are expected utility maximizers with possibly conflicting interests. Section 1 lays out the method. Section 2 illustrates how our method can be applied to compute a policy: first for a large class of examples with a common structure, and second for a worked parametric example. Section 3 concludes. All proofs are in the appendix.

# 1 Posterior Covers

Although the sender might not benefit from giving information everywhere, he is at worst indifferent to it "in between" beliefs at which the receiver's incentives are fixed, in virtue of linearity of expected utility in $\mu$. In such a region, both the sender and the receiver like mean-preserving spreads in beliefs. In other words, from the primitives $\langle A, u \rangle$ (in fact, from $A^*$ alone), we can deduce that the indirect utility $V$ must be locally (weakly) convex on various regions of $\Delta\Omega$. Given a "posterior cover"—a collection of such regions—we may restrict our search for an optimal policy to those leaving the receiver's beliefs elsewhere. This insight reduces the problem to a finite program. When the number of actions is small, it becomes simple to compute an optimal policy and derive the concave envelope.

**Definition 1.** *Given $f : \Delta\Omega \to \mathbb{R}$, an $f$-(posterior) cover is a finite family $\mathcal{C}$ of closed convex subsets of $\Delta\Omega$ such that $f|_C$ is convex for every $C \in \mathcal{C}$.*

A posterior cover is a collection of sets of posterior beliefs, over each of which a given function is convex. Given a posterior cover $\mathcal{C}$, let

$$\text{out}(\mathcal{C}) = \{\mu \in \Delta\Omega : \quad \mu \in \text{ext}(C) \text{ whenever } \mu \in C \in \mathcal{C}\}$$

be its set of **outer points**. That is, outer points are those posterior beliefs that are extreme in any member of $\mathcal{C}$ to which they belong. In particular, any point outside $\bigcup \mathcal{C}$ is an outer point, as is any deterministic belief.

**Theorem 1.**

1. *If $\mathcal{C}$ is any $V$-cover, then there exists an optimal policy $p \in \Delta[out(\mathcal{C})]$ such that $supp(p)$ is affinely independent.*

2. *The collection $\mathcal{C}^* := \{C_a\}_{a \in A}$ is a $V$-cover, where $C_a := \{\mu \in \Delta\Omega : a \in A^*(\mu)\}$ for $a \in A$. Moreover, $out(\mathcal{C}^*)$ is finite.*

For any Bayesian persuasion problem, the theorem names a specific posterior cover $C^*$ and identifies a *finite* set of beliefs out($C^*$) that can support an optimal policy. Using further linear structure, the simplified problem becomes one of finite programming, in which the sender must find the best of *finitely* many combinations of beliefs in out($C^*$) that satisfy Bayes plausibility.[3]

Observe an important consequence of the theorem. To identify a small set of posterior beliefs that support a sender-optimal policy, we may temporarily ignore the sender's preferences. Indeed, the second part of the theorem describes a $V$-cover $C^*$ in terms of $\langle A, u\rangle$ only—in fact, in terms of $A^*$ only—which is possible because at a given optimal action by the receiver, both he and the sender weakly like more information. Therefore, although the sender and the receiver have different preferences, their preferences are weakly aligned in this sense inside every piece of that posterior cover. As a result, an optimal policy can be characterized, and hence $V$ concavified, while computing $V$ *only* on the outer points of $C^*$. To this end, the following result helps compute the outer points of $C^*$:

**Proposition 1.** *out*$(C^*) = \{\mu^* \in \Delta\Theta : B(\mu^*) = \{\mu^*\}\}$, *where*

$$B(\mu^*) := \{\mu \in \Delta\Omega : \quad \mathit{supp}(\mu) \subseteq \mathit{supp}(\mu^*) \text{ and } A^*(\mu) \supseteq A^*(\mu^*)\}\,.$$

This proposition says that the outer points are those beliefs $\mu^*$ whose support cannot be reduced while enlarging the set of agent's optimal actions. That is, they are the most extreme beliefs that support a given set of actions as optimal.

Observe that the theorem separates out the two steps of (i) finding an optimal policy from a $V$-cover and (ii) finding a particular $V$-cover with a small set of outer points. The reason is that, in particular applications, it may be possible to improve upon the second part and find an even better $V$-cover than $C^*$. We see this avenue of work—deriving $V$-covers with small sets of outer points in classes of Bayesian persuasion problems—as a promising one.

# 2 Examples

**Example 1**. A principal will publicly choose an outcome $x \in X$ as a function of the state $\omega \in \Omega$. An agent will observe the principal's decision, and make a binary participation decision $z \in \{0, 1\}$. The principal's objective is $v : X \times \{0, 1\} \times \Omega \to \mathbb{R}$; in particular, the agent's participation may be important to her. The agent does not care directly about the principal's decision, but his participation incentives may depend on what is revealed about

[3] Such a collection of beliefs, being affinely independent, is associated with a unique Bayes plausible information policy.

the state: his value from participation is given by $\hat{u}: \Omega \to \mathbb{R}$ with his non-participation value normalized to zero. On one hand, the principal wants to adapt her decision $x$ to the state $\omega$. On the other, $x$ reveals information about the state that can affect the agent's incentive to participate. This problem is equivalent to a Bayesian persuasion problem with $A = X \times \{0, 1\}$ and $u((x, z), \omega) = z\hat{u}(\omega)$. Let $\Omega_- := \{\omega \in \Omega : \ \hat{u}(\omega) < 0\}$ and $\Omega_+ := \{\omega \in \Omega : \ \hat{u}(\omega) > 0\}$. For any $\omega_- \in \Omega_-$ and $\omega_+ \in \Omega_+$, let

$$\mu_{\omega_-,\omega_+} := \left(\tfrac{\hat{u}(\omega_+)}{\hat{u}(\omega_+)-\hat{u}(\omega_-)}, \ \tfrac{-\hat{u}(\omega_-)}{\hat{u}(\omega_+)-\hat{u}(\omega_-)}\right)$$

be the unique belief supported on $\{\omega_-, \omega_+\}$ at which the receiver is indifferent between participating and not participating. In this environment, it follows directly from Proposition 1 that

$$\text{out}(C^*) = \{\delta_\omega : \ \omega \in \Omega\} \cup \{\mu_{\omega_-,\omega_+} : \ \omega_- \in \Omega_-, \ \omega_+ \in \Omega_+\},$$

where $\delta_\omega$ is the point mass on state $\omega$.

By Theorem 1, there exists an optimal policy $p \in \Delta[\text{out}(C^*)]$, meaning that the principal either (i) almost reveals the state while keeping the receiver indifferent over his participation decision (when $\mu$ is such that $|\text{supp}(\mu)| = 2$), or (ii) fully reveals the state (when $\mu$ is such that $|\text{supp}(\mu)| = 1$).

In this problem, there are no more than $|\Omega| + \frac{1}{4}|\Omega|^2$ outer points, no matter how large $X$ is, which is a substantial reduction when $X$ is large relative to $\Omega$.[4]

Consider further the special case of $\Omega = \{\omega_-, \omega_+\}$ with $\hat{u}(\omega_-) < 0 < \hat{u}(\omega_+)$. Rather than computing $V(\mu) = \max_{a \in A^*(\mu)} \int_\Omega v(a, \cdot) \, \mathrm{d}\mu$ at every distinct $\mu \in \Delta\Omega$,[5] and then seeking to concavify this function, we can simply restrict attention to three beliefs:

$$\begin{aligned}
V(\delta_{\omega_-}) &= \max_{x \in X} v((x, 0), \omega_-) \\
V(\delta_{\omega_+}) &= \max_{x \in X} v((x, 1), \omega_+) \\
V(\mu_{\omega_-,\omega_+}) &= \max_{a \in X \times \{0,1\}} \left\{ \tfrac{\hat{u}(\omega_+)}{\hat{u}(\omega_+)-\hat{u}(\omega_-)} v(a, \omega_-) + \tfrac{-\hat{u}(\omega_-)}{\hat{u}(\omega_+)-\hat{u}(\omega_-)} v(a, \omega_+) \right\}.
\end{aligned}$$

With these three numbers in hand, we can compute an optimal information policy by comparing but two candidates: full information and the policy that maximizes the probability of posterior $\mu_{\omega_-,\omega_+}$ among all information policies.[6]

**Example 2**. Consider a consumer who decides how to spend a sum of money. He can invest

[4] Although we required $|A| < \infty$ in the model, the result is true for any compact $X$ and u.s.c. $v$.

[5] For every $\mu$, this is an optimization problem over $X \times \{0\}$, $X \times \{1\}$, or $X \times \{0, 1\}$.

[6] This latter policy is 'no information' if the prior is $\mu_{\omega_-,\omega_+}$, and supported on $\{\mu_{\omega_-,\omega_+}, \delta_\omega\}$ if the prior is between $\mu_{\omega_-,\omega_+}$ and $\delta_\omega$ for $\omega \in \Omega$.

it in one of two risky assets ($a_1$ or $a_2$) or keep it available to satisfy liquidity needs ($a_\ell$). Let the relevant state come from $\Omega = \{1, 2, \ell\}$, where asset $a_i$ has the highest rate of return in state $i \in \{1, 2\}$, and the agent has an urgent need for liquidity in state $\ell$. Define the consumer's utility as:

| $u$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| 1 | $\alpha$ | $-\beta$ | 0 |
| 2 | $-\gamma$ | $\delta$ | 0 |
| $\ell$ | 0 | 0 | 1 |

where $\alpha, \beta, \gamma, \delta > 0$.

From the theorem, $C^* = \{C_1, C_2, C_\ell\}$ is a $V$-cover, where

$$C_i = \left\{\mu \in \Delta\Omega : \ \mathbb{E}_{\omega\sim\mu}[u(a_i, \omega)] \geq \mathbb{E}_{\omega\sim\mu}[u(a_j, \omega)] \ \text{ for all } \ j\right\}.$$

That is, parametrizing $\mu \in \Delta\Omega$ by $\mu_1 := \mu\{1\}$ and $\mu_2 := \mu\{2\}$:

$$\begin{aligned}
C_1 &= \{(\mu_1, \mu_2) : (\alpha + \beta)\mu_1 \geq (\delta + \gamma)\mu_2, \ (1 + \alpha)\mu_1 \geq 1 + (\gamma - 1)\mu_2\} \\
C_2 &= \{(\mu_1, \mu_2) : (\delta + \gamma)\mu_2 \geq (\alpha + \beta)\mu_1, \ (1 + \delta)\mu_2 \geq 1 + (\beta - 1)\mu_1\} \\
C_\ell &= \{(\mu_1, \mu_2) : 1 + (\gamma - 1)\mu_2 \geq (1 + \alpha)\mu_1, \ 1 + (\beta - 1)\mu_1 \geq (1 + \delta)\mu_2\}.
\end{aligned}$$

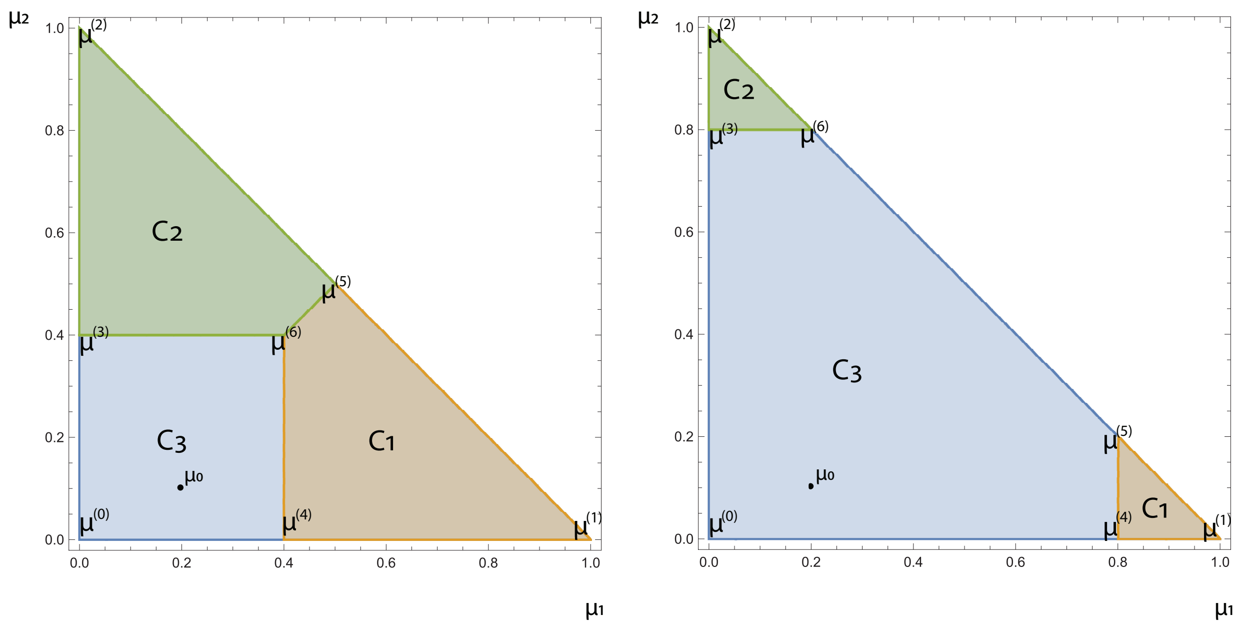
Letting $O := \left\{(0, 0), (1, 0), (0, 1), \left(0, \frac{1}{1+\delta}\right), \left(\frac{1}{1+\alpha}, 0\right)\right\}$, our outer points are

$$\text{out}(C^*) = \begin{cases} O \cup \left\{\left(\frac{\gamma+\delta}{\alpha+\beta+\gamma+\delta}, \frac{\alpha+\beta}{\alpha+\beta+\gamma+\delta}\right), \left(\frac{\gamma+\delta}{\alpha\delta-\beta\gamma+\alpha+\beta+\gamma+\delta}, \frac{\alpha+\beta}{\alpha\delta-\beta\gamma+\alpha+\beta+\gamma+\delta}\right)\right\} & : \ \alpha\delta \geq \beta\gamma \\ O \cup \left\{\left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right), \left(\frac{\delta}{\gamma+\delta}, \frac{\gamma}{\gamma+\delta}\right)\right\} & : \ \alpha\delta \leq \beta\gamma. \end{cases}$$

The specific form of the outer points brings out the tradeoff captured by $\alpha\delta - \beta\gamma$. The qualitative form of the posterior cover $C^*$ depends on whether the agent wants to invest in one of the two risky assets when he is indifferent between them and certain that he has no liquidity needs, which is equivalent to $\alpha\delta \geq \beta\gamma$.

Figure 1 depicts the posterior cover for two parameter specifications (with $\alpha\delta - \beta\gamma$ having different signs). Afterward, we study how optimal disclosure varies between the two given specific receiver preferences.

(i) Figure 1a plots the posterior cover for $\beta = \gamma = 1$ and $\alpha = \delta = \frac{3}{2}$. The outer points are $\mu^{(0)} = (0, 0)$, $\mu^{(1)} = (1, 0)$, $\mu^{(2)} = (0, 1)$, $\mu^{(3)} = (0, \frac{2}{5})$, $\mu^{(4)} = (\frac{2}{5}, 0)$, $\mu^{(5)} = (\frac{1}{2}, \frac{1}{2})$, $\mu^{(6)} = (\frac{2}{5}, \frac{2}{5})$. With this numerical specification and $\mu_0 = (\frac{1}{5}, \frac{1}{10})$, the designer must




(a) $\alpha\delta > \beta\gamma$

(b) $\alpha\delta < \beta\gamma$

Figure 1: $C^*$-Cover

compare eight information policies that are characterized by their support:

$$\begin{array}{rclcrcl} \text{supp}(p_1) & = & \{\mu^{(0)}, \mu^{(4)}, \mu^{(5)}\}, & \quad & \text{supp}(p_5) & = & \{\mu^{(0)}, \mu^{(1)}, \mu^{(5)}\}, \\ \text{supp}(p_2) & = & \{\mu^{(0)}, \mu^{(4)}, \mu^{(6)}\}, & & \text{supp}(p_6) & = & \{\mu^{(0)}, \mu^{(1)}, \mu^{(6)}\}, \\ \text{supp}(p_3) & = & \{\mu^{(0)}, \mu^{(4)}, \mu^{(2)}\}, & & \text{supp}(p_7) & = & \{\mu^{(0)}, \mu^{(1)}, \mu^{(2)}\}, \\ \text{supp}(p_4) & = & \{\mu^{(0)}, \mu^{(4)}, \mu^{(3)}\}, & & \text{supp}(p_8) & = & \{\mu^{(0)}, \mu^{(1)}, \mu^{(3)}\}. \end{array}$$

Whatever the sender's preferences, one of these eight information policies will be optimal.

(ii) Figure 1b plots the posterior cover for $\beta = \gamma = 1$ and $\alpha = \delta = \frac{1}{4}$. Let $\mu^{(0)}$ through $\mu^{(2)}$ be the same as before, but $\mu^{(3)} = (0, \frac{4}{5})$, $\mu^{(4)} = (\frac{4}{5}, 0)$, $\mu^{(5)} = (\frac{4}{5}, \frac{1}{5})$, $\mu^{(6)} = (\frac{1}{5}, \frac{4}{5})$. Given a prior, the sender will choose among finitely many policies supported on these outer points, as in (i).

To illustrate, let $v(a, \theta) = \mathbf{1}_{a=a_1}$, and keep $\mu_0 = (\frac{1}{5}, \frac{1}{10})$. For the specification in Figure 1a, information policy $\frac{1}{2}\left[\mu^{(0)}\right] + \frac{1}{4}\left[\mu^{(4)}\right] + \frac{1}{4}\left[\mu^{(6)}\right]$ is optimal and gives the sender a value of $\frac{1}{2}$. For the specification in Figure 1b, information policy $\frac{13}{20}\left[\mu^{(0)}\right] + \frac{1}{4}\left[\mu^{(4)}\right] + \frac{1}{10}\left[\mu^{(2)}\right]$ is optimal and gives the sender a value of $\frac{1}{4}$. We can make sense of this change in optimal disclosure through the sign of $\alpha\delta - \beta\gamma$. In both situations, the sender at times concedes that the agent needs liquidity. The rest of the time, she can keep asset 1 weakly optimal in the first situation. This possibility disappears in the second situation, because the agent prefers $a_\ell$ when he is indifferent between the two risky assets. In this case, the sender may as well reveal to the agent when asset 2 is optimal.

# 3 Conclusion

The method of posterior cover focuses attention on maximally informative policies, subject to preserving agent incentives. It simplifies Bayesian persuasion by identifying a small set of posterior beliefs (the outer points) that support optimal policies for all priors and all sender preferences. The problem is then one of finite programming, as shown in the examples. This is especially helpful if one wants to study how optimal information varies with the prior or the sender's preferences.

By looking at the outer points, one can sometimes learn about the form of optimal persuasion even before naming an optimal policy. For example, in a binary-state world, if the outer points are $\{0, \mu_0, 1\}$, as may happen in the first example, then the method delivers an interpretable result: either full information or no information is optimal, whatever the sender's preferences are. If the outer points change with respect to some characteristic of the receiver's preferences, as in the second example, then we learn that said characteristic qualitatively shapes the types of information a sender may employ.

# 4 Appendix

Below is the proof of Theorem 1.

*Proof.* Fix a $V$-cover $C$, and let $\mathcal{B} := \Delta[\text{out}(C)]$ and $\mathcal{R} := \mathcal{R}(\mu_0)$, both compact convex subsets of $\Delta\Delta\Omega$. By Krein-Milman, and as

$$\begin{aligned} \mathbb{E}V : \Delta\Delta\Omega &\to \mathbb{R} \\ p &\mapsto \int_{\Delta\Omega} V \, \mathrm{d}p \end{aligned}$$

is affine and continuous, $\mathbb{E}V|_{\mathcal{R}\cap\mathcal{B}}$ is maximized somewhere on $\text{ext}(\mathcal{R}\cap\mathcal{B})$. But $\mathbb{E}V|_{\mathcal{R}}$ is maximized on $\mathcal{R}\cap\mathcal{B}$ by the "optimal support" part of Lipnowski and Mathevet (2017, Theorem 1).[7] Moreover, $\mathcal{B}$ is an extreme subset of $\Delta\Delta\Omega$, so that $\mathcal{R}\cap\mathcal{B}$ is an extreme subset of $\mathcal{R}$, and therefore $\text{ext}[\mathcal{R}\cap\mathcal{B}] = \mathcal{B}\cap\text{ext}(\mathcal{R})$. Accordingly, $\mathbb{E}V|_{\mathcal{R}}$ is maximized on $\mathcal{B}\cap\text{ext}(\mathcal{R})$.

Now, consider arbitrary $p \in \text{ext}(\mathcal{R})$, with the goal of showing that $\text{supp}(p)$ is affinely independent. With $\Delta\Omega$ finite-dimensional, $p \in \overline{\text{co}}[\text{supp}(p)] = \text{co}[\text{supp}(p)]$. Carathéodory's Theorem then delivers an affinely independent set $S \subseteq \text{supp}(p)$ such that $\mu_0 \in \text{co}(S)$.

Now, there is a correspondence $N : S \rightrightarrows \Delta\Omega$ such that, for each $\nu \in S$, the set $N(\nu)$ is a closed convex neighborhood of $\nu$ with $S \cap N(\nu) = \{\nu\}$. Making $(N(\nu))_{\nu\in S}$ smaller if necessary, we may assume that every selector $\eta$ of $N$ has $(\eta(\nu))_{\nu\in S}$ affinely independent and $\mu_0 \in \text{co}\left[\{\eta(\nu)\}_{\nu\in S}\right]$.

It follows from $S \subseteq \text{supp}(p)$ that $p[N(\nu)] > 0$ for every $\nu \in S$. Define, then, $\eta : S \to \Delta\Omega$ which maps $\nu \in S$ to the mean of $\mu \sim p$ conditional on $\mu \in N(\nu)$. More formally, define

$$\begin{aligned} \hat{\eta} : S &\to \Delta\Delta\Omega \\ \nu &\mapsto \tfrac{p[N(\nu)\cap(\cdot)]}{p[N(\nu)]}, \end{aligned}$$

and let $\eta(\nu)$ be the barycentre of $\hat{\eta}(\nu)$ for $\nu \in S$.

As $N$ is compact-convex-valued, $\eta$ is then a selector of $N$. There is therefore some $\gamma \in \Delta S$ such that $\int_S \eta \, \mathrm{d}\gamma = \mu_0$ and $\text{supp}(\gamma) = S$. Now, let

$$q := \int_{\Delta\Delta\Omega} \hat{\eta} \, \mathrm{d}\gamma \in \mathcal{R} \text{ and } \epsilon := \tfrac{1}{2} \wedge \min_{\nu\in S} \tfrac{\gamma(\nu)}{p[N(\nu)]} > 0.$$

By construction, $\frac{p-\epsilon q}{1-\epsilon} \in \mathcal{R}$ and $p \in \text{co}\left\{q, \ \frac{p-\epsilon q}{1-\epsilon}\right\}$. Since $p$ is extreme, we know $q = p$. As this is true even in the limit (holding $S$ fixed) as $N(\nu) \to \{\nu\}$, it must be that $\text{supp}(p) = S$. But then $p$ has affinely independent support, as required. This proves the first part.

[7] Replacing "$U$" with "$V$" and "$\Theta$" with "$\Omega$", the proof applies verbatim.

Toward the second part, note that $C_a$ is closed and convex for every $a \in A$, expected utility being linear and continuous in beliefs. Next, $C_a$ is the intersection of the simplex $\Delta\Omega$ and a collection of half-spaces $\{\mu \in \mathbb{R}^\Omega : \sum_{\omega\in\Omega} \mu_\omega[u(a,\omega) - u(a',\omega)] \geq 0\}_{a'\in A}$. Therefore, $C_a$ has finitely many extreme points. This tells us that $\text{out}(C^*)$ is finite, since:

$$\begin{aligned}
\text{out}(C^*) &= \text{out}(C^*) \cap \bigcup C^* \text{ (as a finite optimization problem has a solution)} \\
&= \left\{\mu \in \bigcup C^* : \ \mu \in \text{ext}(C) \text{ for any } \mu \in C \in \mathcal{C}\right\} \\
&\subseteq \bigcup_{C\in C^*} \{\mu \in C : \ \mu \in \text{ext}(C)\} = \bigcup_{a\in A} \text{ext}(C_a).
\end{aligned}$$

Now, fix an arbitrary $a^* \in A$ and affine $f : [0,1] \to C_{a^*}$; we will show that $V \circ f$ is weakly convex. Define $g : A \times [0,1] \to \mathbb{R}$ via $g(a,x) := \int_\Omega [u(a^*,\cdot) - u(a,\cdot)] \, \text{d}[f(x)]$. Note that $g(a,\cdot)$ is affine for every $a \in A$. For any $x \in [0,1]$, we know $g(a,x) \geq 0$ because $f(x) \in C_{a^*}$. For any $a \in A$, as $g(a,\cdot)$ is affine, it follows that $g(a,\cdot)^{-1}(0) \in \{\emptyset, [0,1], \{0\}, \{1\}\}$. Therefore, there exist disjoint subsets $\hat{A}, A_0, A_1 \subseteq A$ such that

$$A^*(f(x)) = \{a \in A : \ g(a,x) = 0\} = \begin{cases} \hat{A} \cup A_x & \text{if } x \in \{0,1\}, \\ \hat{A} & \text{if } x \in (0,1). \end{cases}$$

First, define $\hat{V} : [0,1] \to \mathbb{R}$ via $\hat{V}(x) := \max_{a\in\hat{A}} \int_\Omega v(a,\cdot) \, \text{d}[f(x)]$; it is convex, as a maximum of affine functions. Next, define $\check{V} : [0,1] \to \mathbb{R} \cup \{-\infty\}$ via $\check{V}(x) := \max_{a\in A_x} \int_\Omega v(a,\cdot) \, \text{d}[f(x)]$ for $x \in \{0,1\}$, and $\check{V}|_{(0,1)} = -\infty$; it is obviously convex. Therefore, $V \circ f = \max\{\hat{V}, \check{V}\}$ is convex. It follows that $V|_{f([0,1])}$ is convex. As $f$ was arbitrary, $V_{C_{a^*}}$ is convex. □

Finally, Proposition 1 follows from Lipnowski and Mathevet (2017, Lemma 4).

# References


Aumann, Robert J. and Michael B. Maschler. 1995. *Repeated Games with Incomplete Information*. Cambridge, MA: MIT Press.

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

Lipnowski, Elliot and Laurent Mathevet. 2017. "Disclosure to a Psychological Audience." Working Paper, Economics Department, New York University.

Tardella, Fabio. 2008. "On The Existence of Polyhedral Convex Envelopes." In *Frontiers In Global Optimization*, edited by C.A. Floudas and P.M. Pardalos. Boston: Kluwer, 149–188.