Entangled vs. Separable Choice*

Nail Kashaev ⓡ Martin Plávala ⓡ Victor H. Aguiar†

This version: March, 2024. First version: March, 2024

Abstract We study joint probabilistic choice rules that describe the behavior of two decision makers, each facing a possibly different menu. These choice rules are separable when they can be factored into autonomous choices from each individual solely correlated through their individual probabilistic choice rules. Despite recent interest in studying such rules, a complete characterization of the restrictions on them remains an open question. A reasonable conjecture is that such restrictions on separable joint choice can be factored into individual choice restrictions. We name these restrictions separable and show that this conjecture is true if and only if the probabilistic choice rule of at least one decision maker uniquely identifies the distribution over deterministic choice rules. Otherwise, entangled choice rules exist that satisfy separable restrictions yet are not separable. The possibility of entangled choice complicates the characterization of separable choice since one needs to augment the separable restrictions with the new emerging ones.

JEL classification numbers: C10.

Keywords: correlated choice, stochastic choice, entanglement, partial identification.

---

*The paper subsumes some results from Kashaev, Aguiar, Plávala, and Gauthier (2023). The "ⓡ" symbol indicates that the authors' names are in certified random order, as described by Ray and Robson (2018). We thank Roy Allen, David Freeman, Salvador Navarro, Krishna Pendakur, Wilson Perez, Bruno Salcedo, John Quah, and Lanny Zrill for useful discussions and encouragement. Aguiar and Kashaev gratefully acknowledge financial support from the Social Sciences and Humanities Research Council Insight Development Grant. Plávala acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, project numbers 447948357 and 440958198), the Sino-German Center for Research Promotion (Project M-0294), the German Ministry of Education and Research (Project QuKuK, BMBF Grant No. 16KIS1618K), the DAAD, and the Alexander von Humboldt Foundation. Aguiar thanks USFQ School of Economics for kindly hosting him while writing this paper.

†Kashaev: Department of Economics, University of Western Ontario; nkashaev@uwo.ca. Plávala: Naturwissenschaftlich-Technische Fakultät, Universität Siegen; martin.plavala@uni-siegen.de. Aguiar: Department of Economics, Simon Fraser University; vaguiarl@sfu.ca.

1. Introduction

We define *entanglement of choice* of two decision makers (DMs) and showcase why its possibility poses conceptual and computational challenges to the characterization of separable joint probabilistic choice rules. These rules describe the joint stochastic behavior of two DMs whose choices can be correlated only via their individual choice rules. That is, the DMs could know each other but choose as if they are in separate rooms where each DM chooses according to her individual choice function from her individual menu. All possible interactions of the restrictions on the individual probabilistic choice rules imply the restrictions on the joint one. We call these restrictions *separable*. An observer would expect that, in this separable environment without direct interactions, the separable restrictions should fully characterize separable joint probabilistic choice rules. However, this is not always the case. When the individual probabilistic choice rule, which is a mixture of deterministic choice rules, of each DM does not uniquely identify the distribution over individual deterministic choice rules, the separable restrictions fail to be sufficient, thus, leading to *entangled choice*. Entangled choices are joint probabilistic choice rules that are not separable yet are consistent with the separable restrictions. We show that the separable restrictions are necessary and sufficient to characterize separable choice if and only if the uniqueness property holds.

The absence of the uniqueness property can thwart the efforts towards a general characterization of separable rules because, in that case, separable restrictions do not rule out entangled choice. We connect this difficulty in characterizing separable rules in economic theory with existing results in theoretical physics and computer science that have shown that testing whether a system is separable or entangled is an NP-hard problem (Gurvits, 2003). In that line, we provide a characterization of separable joint probabilistic rules for a simple scenario using Bell inequalities (Bell, 1964) solving an open question posed by Chambers, Masatlioglu, and Turansick (2021) for a simple domain creating a bridge between economic theory and existing results in the study of separable systems (Rosset, Bancal, and Gisin, 2014).

Our results can be used in general finite environments to obtain the separable restrictions and full characterizations under the uniqueness restriction. We generalize recent results in stochastic choice with random utility by Chambers et al. (2021) and Li (2021), solving an open question posed in Strzalecki (2021) for the case of uniqueness. Finally, we justify the importance of uniqueness (point identification) when working with separable joint probabilistic rules: lack of uniqueness leads to the possibility of entangled choice that we argue is an undesirable property of models. A similar phenomenon appears in the analysis of finite games with multiple equilibria. The multiplicity of equilibria, similar to the lack of uniqueness in our setting, often leads to a correlation between the choices of players even after conditioning on available information (De Paula and Tang, 2012).

Separable joint probabilistic rules interest social scientists because of their simplicity in modeling the joint behavior of several DMs as a straightforward interaction of individual behavior, making these joint behaviors compositional and tractable (Kashaev et al., 2023). In particular, in the context of the detection of imitation in group setups (as surveyed in Sacerdote, 2011), *excess variance* in group behavior is used to disentangle imitation from independent behavior. Here, we provide analogous bounds on correlation of choice that detects entanglement. In addition, Ashenfelter and Krueger (1994) have studied a similar primitive in relationship with the returns to education of identical twins. Kashaev et al. (2023) make progress in characterizing separable rules in the context of random utility when entangled choice is present. They interpret multiple DMs as multiple selves in time.

2. A Difficulty with Separable Choice: An Example of Entangled Choice

Consider a hypothetical experimental setting with three distinct alternatives: x, y, and z. A DM, Frodo ($t = 1$), has to pick a single alternative when presented with any of the two menus $\{xy\}$ and $\{yz\}$. The columns of matrix A^t below encode all possible deterministic choice rules or choice patterns of Frodo. The rows of A^t encode all possible pairs of choices and menus that could be observed. The entries of A^t are either 0 or 1. For example, the entry in the first row $(x, \{xy\})$ and the first column is equal to 1 and is interpreted as the deterministic choice rule that picks x from menu $\{xy\}$. Because choice rules are single-valued, the subvector of each column corresponding to the same menu must add up to 1.

$$A^t = \begin{array}{c} \left(\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \end{array} \begin{array}{||l} \text{x,\{xy\}} \\ \text{y,\{xy\}} \\ \text{y,\{yz\}} \\ \text{z,\{yz\}} \end{array}.$$

Frodo can exhibit stochastic behavior, in particular, consider a distribution over the columns on A^t, ν_t, such that the probabilistic choice rule of DM t, ρ_t, is given by

$$\rho_t = A^t \nu_t.$$

The probabilistic choice rule ρ_t is such that all its entries are nonnegative and are such that

$$\rho_t(x, \{xy\}) + \rho_t(y, \{xy\}) = \rho_t(y, \{yz\}) + \rho_t(z, \{yz\}) = 1.$$

We highlight that other than these nonnegativity and adding-up constraints, there are no more restrictions on ρ_t because Frodo randomizes over all possible deterministic choice rules in this setup. In other words, any probabilistic choice rule can be represented as a mixture of all deterministic choice rules.

Next, we consider $T = 2$ DMs. Frodo ($t = 1$) and Sam ($t = 2$) make choices from binary menus. The choices are separated because Frodo and Sam make their choices as if they are in separate isolated rooms with no communication before or during the experiment. However, Frodo and Sam grew up in the same village and are friends. Note that since rooms are isolated and no communication devices are permitted, Frodo and Sam cannot coordinate choices after the start of the experiment, nor can they see what menu the other DM is seeing. However, since Frodo and Sam are friends, their randomization devices over deterministic choice functions can be arbitrarily correlated.

Frodo and Sam face, in each trial, ordered pairs of menus or menu paths (e.g., $\{xy\}$ and $\{yz\}$)). The first menu will be available to Frodo and the second to Sam. We assume that all possible combinations of menus are presented to Frodo and Sam in sequential trials such that there are 4 menu paths. In each menu path, say $\{xy\}, \{yz\}$, the experimenter observes the probability of each of the 4 choice paths such that

$$\rho(x, \{xy\}; y, \{yz\}) + \rho(x, \{xy\}; z, \{yz\}) + \rho(y, \{xy\}; y, \{yz\}) + \rho(y, \{xy\}; z, \{yz\}) = 1$$

Assuming that there are no restrictions on the Sam's individual behavior (i.e., $A^2 = A^1$), we can encode all possible joint choice patterns of Frodo and Sam when faced with a menu path in the columns of matrix A below. For example, column 1 combines column 1 of A^1 and column 1 of A^2, each describing the individual choice rules. The 16 rows of A correspond to all choice and menu paths. For example, the entry of the first column and the first row is equal to 1. This means that both DMs are picking the same alternative out of the same menu because the choice rules that describe their deterministic behavior in this column are the same. Column 2 combines column 1 of A^1 and column 2 of A^2. Hence, its second entry is equal to 1 because Sam chooses y when faced with $\{yz\}$. In total, there are 16 possible deterministic choice patterns for the joint problem faced by the DMs.

$$A = \left(\begin{array}{cccccccccccccccc}
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1
\end{array}\right)
\left\|\begin{array}{l}
x, \{xy\}, x, \{xy\} \\
x, \{xy\}, y, \{xy\} \\
x, \{xy\}, y, \{yz\} \\
x, \{xy\}, z, \{yz\} \\
y, \{xy\}, x, \{xy\} \\
y, \{xy\}, y, \{xy\} \\
y, \{xy\}, y, \{yz\} \\
y, \{xy\}, z, \{yz\} \\
y, \{yz\}, x, \{xy\} \\
y, \{yz\}, y, \{xy\} \\
y, \{yz\}, y, \{yz\} \\
y, \{yz\}, z, \{yz\} \\
z, \{yz\}, x, \{xy\} \\
z, \{yz\}, y, \{xy\} \\
z, \{yz\}, y, \{yz\} \\
z, \{yz\}, z, \{yz\}
\end{array}\right.$$

Note that the vector that collects the choice probabilities over all 4 possible choice paths, ρ, is given (up to permutation) by

$$\rho = A\nu \tag{1}$$

for some probability distribution over the columns of A, ν.

Under this experiment, ρ satisfies nonnegativity, adding-up constraints for each menu path, and a condition that we call *marginality*:

$$\rho(x, \{xy\}; x, \{xy\}) + \rho(x, \{xy\}; y, \{xy\}) = \rho(x, \{xy\}; y, \{yz\}) + \rho(x, \{xy\}; z, \{yz\})$$

and

$$\rho(x, \{xy\}; x, \{xy\}) + \rho(y, \{xy\}; x, \{xy\}) = \rho(y, \{yz\}); x, \{xy\}) + \rho(z, \{yz\}; x, \{xy\}).$$

In general, ρ satisfies marginality if

$$\sum_{m_2 \in \{xy\}} \rho(m_1, M_1; m_2, \{xy\}) = \sum_{m_2 \in \{yz\}} \rho(m_1, M_1; m_2, \{yz\})$$

for all m_1 and M_1 and

$$\sum_{m_1 \in \{xy\}} \rho(m_1, M_1; m_2, \{xy\}) = \sum_{m_1 \in \{yz\}} \rho(m_1, M_1; m_2, \{yz\})$$

for all m_2 and M_2. Marginality follows from Frodo and Sam being in separate rooms and the fact that they randomize over the deterministic choice patterns using a distribution that does not depend on the particular menu path. This observation has been done before by Strzalecki (2021) and echoed in Chambers et al. (2021). In addition, this condition is known as *no signaling* in the theoretical physics literature for systems that are not related to choice (Rosset et al., 2014).

Since the probabilistic choice generated by Frodo by randomizing over her deterministic choice rules does not impose any additional restrictions on individual ρ_t, the observer has a reasonable expectation that there are no more restrictions on joint ρ beyond nonnegativity, adding-up, and marginality constraints. However, as we show in the next section, it is possible to construct ρ that satisfies all these separable restrictions, yet our thought experiment can not generate it. That is, the thought experiment generates new *emerging* restrictions on ρ. We call the discrepancy between the separable behavioral implications and the actual implications of the thought experiment *entanglement of choice*. We say that ρ is entangled if it satisfies the separable restrictions yet fails to be generated by the thought experiment.

2.1. A Characterization of the Thought Experiment

We use Bell's inequalities (Bell, 1964) to provide the emerging restrictions generated by the thought experiment. In particular, we use so-called the CHSH (Clauser, Horne, Shimony, and Holt) inequalities (Rosset et al., 2014). Define 4 measures of choice coordination for 4 different menu paths as

$$\begin{aligned}
E_{\{xy\},\{xy\}} &= \rho(x,\{xy\};x,\{xy\}) + \rho(y,\{xy\};y,\{xy\}) - \rho(x,\{xy\};y,\{xy\}) - \rho(y,\{xy\};x,\{xy\}),\\
E_{\{xy\},\{yz\}} &= \rho(x,\{xy\};y,\{yz\}) + \rho(y,\{xy\};z,\{yz\}) - \rho(x,\{xy\};z,\{yz\}) - \rho(y,\{xy\};y,\{yz\}),\\
E_{\{yz\},\{xy\}} &= \rho(y,\{yz\};x,\{xy\}) + \rho(z,\{yz\};y,\{xy\}) - \rho(z,\{yz\};x,\{xy\}) - \rho(y,\{yz\};y,\{xy\}),\\
E_{\{yz\},\{yz\}} &= \rho(y,\{yz\};y,\{yz\}) + \rho(z,\{yz\};z,\{yz\}) - \rho(y,\{yz\};z,\{yz\}) - \rho(z,\{yz\};y,\{yz\}).
\end{aligned}$$

Note that, $E_{\{xy\},\{xy\}} = 1$ if in menu path $\{xy\},\{xy\}$, Frodo and Sam coordinate on picking the same alternative. At the same time, $E_{\{xy\},\{xy\}} = -1$ of DMs always pick different alternatives. Also, $\left|E_{\{xy\},\{xy\}}\right| \le 1$ by construction. Hence, we can interpret $E_{M,M'}$ as a measure of cooperation between Frodo and Sam in menu path (M, M').

Now, we can define the following CHSH inequalities:

$$\begin{aligned}
-2 &\le E_{\{xy\},\{xy\}} + E_{\{yz\},\{xy\}} + E_{\{xy\},\{yz\}} - E_{\{yz\},\{yz\}} \le 2,\\
-2 &\le E_{\{xy\},\{xy\}} + E_{\{yz\},\{xy\}} - E_{\{xy\},\{yz\}} + E_{\{yz\},\{yz\}} \le 2,\\
-2 &\le E_{\{xy\},\{xy\}} - E_{\{yz\},\{xy\}} + E_{\{xy\},\{yz\}} + E_{\{yz\},\{yz\}} \le 2,\\
-2 &\le -E_{\{xy\},\{xy\}} + E_{\{yz\},\{xy\}} + E_{\{xy\},\{yz\}} + E_{\{yz\},\{yz\}} \le 2.
\end{aligned}$$

The CHSH inequalities provide lower and upper bounds on the cooperation between Frodo and Sam across *all* menu paths in the thought experiment. In particular, at least one of these inequalities is violated if Frodo and Sam fully cooperate in 3 out of 4 menu paths but defect in one. For example, the second inequality is violated if $E_{\{xy\},\{xy\}} = E_{\{yz\},\{xy\}} = E_{\{yz\},\{yz\}} = 1$ and $E_{\{xy\},\{yz\}} = -1$.

Proposition 1. Probabilistic choice rule ρ is consistent with the thought experiment if and only if it satisfies marginality and the CHSH inequalities.

The sufficiency of Proposition 1 follows from Fine (1982) (Proposition 2). Necessity is trivial to verify in this case. The CHSH inequalities are not implied by the interaction of the individual behavioral restrictions, and thus are emerging. In particular, they are distinct from adding-up, nonnegativity, and marginality.

Table 1 depicts a continuum of nondegenerate examples of ρ that satisfy marginality, but do not obey the CHSH inequalities. In particular, $E_{\{xy\},\{xy\}} = E_{\{yz\},\{xy\}} = E_{\{yz\},\{yz\}} = 2(\alpha - \beta)$ and $E_{\{xy\},\{yz\}} = 2(\beta - \alpha)$. Hence, the second CHSH inequality is violated if and only if

$$E_{\{xy\},\{xy\}} + E_{\{yz\},\{xy\}} - E_{\{xy\},\{yz\}} + E_{\{yz\},\{yz\}} = 8(\alpha - \beta) > 2.$$

In other words, for $\frac{3}{8} < \alpha \leq \frac{1}{2}$ and $\beta = \frac{1}{2} - \alpha$ the implied by Table 1 ρ satisfies marginality but cannot be explained by the thought experiment. The violation of CHSH is maximal when $\alpha = \frac{1}{2}$ and $\beta = 0$. That particular configuration was documented in Chambers et al. (2021).

When ρ violates the CHSH inequalities, we say ρ is entangled since the thought experiment cannot explain it, and there must be some form of unobserved communication between Frodo and Sam. The CHSH inequalities in our domain correspond to nonparametric analogues of the excess variance approach to detecting imitation in the peer effects literature (Sacerdote, 2011). In other words, they provide a threshold on correlation of Sam's and Frodo's choices above which choice is not separable and becomes entangled.

Frodo/Sam	$\{x, xy\}$	$\{y, xy\}$	$\{y, yz\}$	$\{z, yz\}$
$\{x, xy\}$	α	β	β	α
$\{y, xy\}$	β	α	α	β
$\{y, yz\}$	α	β	α	β
$\{z, yz\}$	β	α	β	α

Table 1 – Rows correspond to choices and menus for Frodo and columns correspond to choices and menus for Sam. Entries are the joint probabilities over the choice path formed by the corresponding row and column. The parameters are such that $\alpha + \beta = 1/2$ and $\frac{3}{8} < \alpha \leq \frac{1}{2}$.

Next, we present an explanation of this unexpected behavior in the thought experiment and provide a necessary and sufficient condition for it. The condition is based on recent results from mathematical quantum physics that have studied a similar mathematical structure in other domains (Aubrun, Lami, Palazuelos, and Plávala, 2021).

3. Separable Restrictions over Separable Choice

Recall that the thought experiment produces a probabilistic choice rule that satisfies $\rho = A\nu$ for some distribution over columns of A. Every column of A corresponds to a composite type that describes the behavior of each DM. That is, separability in the thought experiment is captured by interactions of different individual types. Formally,

$$A = A^1 \otimes A^2,$$

where $\otimes$ is the Kronecker product of matrices.[1]

When we look at the individual behavioral restrictions implied by $\rho_t = A^t\nu$, we find that the restrictions on ρ_t can be captured in the form of inequality constraints $H^t\rho_t \geq 0$, where

$$H^t = \begin{pmatrix} -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

That is, $\rho_t = A^t\nu$ for some distribution ν if and only if $H^t\rho_t \geq 0$. Rows of H^t capture the restrictions over ρ_t we discussed previously. In particular, rows 1 and 2 combine to provide the adding-up constraint. Rows 3-6 are the nonnegativity restrictions.

[1] If C is an m-by-n matrix and D is a p-by-q matrix, then the Kronecker product $C \otimes D$ is the pm-by-qn block matrix:

$$C \otimes D = \begin{pmatrix} C_{1,1}D & \dots & C_{1,n}D \\ \vdots & \ddots & \vdots \\ C_{m,1}D & \dots & C_{m,n}D \end{pmatrix}.$$

Since the behavior in the thought experiment is generated from the interaction of individual types, we can think of separable restrictions over ρ as those coming from the interaction of the individual restrictions captured by H^t. In other words, the correlation of choice cannot obscure the requirement that each DM chooses from its own menu that is separated from the other DM and that the joint randomization does not depend on the particular menu path. Formally, we define the interaction of such restrictions by

$$(H^1 \otimes H^2)\rho \geq 0,$$

where H^1 and H^2 are the individual restrictions for DM 1 and 2, respectively.

Rows of H^t are restrictions over the behavior captured by ρ_t. The Kronecker product of H^1 and H^2 requires ρ to satisfy the combinations of these restrictions: every restriction of Frodo is combined with every restriction of Sam. These are new restrictions on joint behavior. Importantly, all separable restrictions on joint behavior must arise from the restrictions on individual behavior.

Direct computation of joint restrictions produces adding-up, nonnegativity, and marginality constraints in our thought experiment. In particular, marginality arises from the interactions of the individual adding-up constraints. In this formal sense, we label marginality as a separable restriction over the behavior implied by the thought experiment. Using this formalization we can now say that ρ in Table 1 satisfies the separable restrictions yet it fails to be consistent with the behavior implied by the thought experiment.

Remark 1. In the appendix, we show that separable restrictions are always necessary. They are always implied by the thought experiment and any extension of it for multiple agents and richer finite domains.

4. A Necessary and Sufficient Condition for (Lack of) Entanglement of Choice

Now we consider a slightly more general thought experiment for Frodo and Sam with matrices $A^{1,\Diamond}$ and A^2 describing their behavior. Matrix $A^{1,\Diamond}$ is constructed from columns of A^1. It collects all the *allowable* deterministic choice rules for Frodo in our setup. The experimenter controls the allowable behavior. For example, the experimenter can introduce a dominance relation among alternatives to restrict the behavior of Frodo.

We restrict $A^{1,\lozenge}$ to be *generating*. That is, we require the system of equations $A^{1,\lozenge}\nu = \rho_1$ to have a solution (possibly with negative entries) for every probabilistic choice rule ρ_1. In other words, all signed measures over the columns of $A^{1,\lozenge}$ should be able to generate all possible probabilistic choice rules ρ_1. In the thought experiment, $A^{1,\lozenge}$ is generating if and only if it consists of at least 3 different columns of A^1.

Since $A^{1,\lozenge}$ is a submatrix of A^1, individual probabilistic choice of Frodo is (weakly) more restricted. We collect the restrictions of the behavior of DM 1 in matrix $H^{1,\lozenge}$ such that $H^{1,\lozenge}\rho_1 \geq 0$ if and only if $\rho_1 = A^{1,\lozenge}\nu_1$ for some distribution over columns of $A^{1\lozenge}$, ν_1. The *thought experiment*$^{\lozenge}$ is defined completely analogous to the thought experiment but with A^1 replaced by $A^{1,\lozenge}$.

We also need two more definitions. We say that $A^{1,\lozenge}$ generates a *unique representation* when the system $\rho_1 = A^{1,\lozenge}\nu_1$ has a unique solution for all probabilistic choice rules ρ_1. This is a restriction on the columns of $A^{1,\lozenge}$. In particular, it is satisfied if and only if the columns of $A^{1,\lozenge}$ are linearly independent (i.e., $A^{1,\lozenge}$ has full column rank). For a generating $A^{1,\lozenge}$ the latter happens for if and only if $A^{1,\lozenge}$ has exactly 3 out of 4 columns of A^1. That is, A^1 is generating but does not generate a unique representation.

Finally, we say that the thought experiment$^{\lozenge}$ produces *only separable restrictions* on ρ if

$$\rho = \left(A^{1,\lozenge} \otimes A^2\right)\nu \text{ for some distribution } \nu \iff (H^{1,\lozenge} \otimes H^2)\rho \geq 0.$$

In other words, the thought experiment$^{\lozenge}$ produces only separable restrictions if it cannot generate an entangled ρ. As we demonstrated in the previous section, when $A^{1,\lozenge} = A^1$, the thought experiment$^{\lozenge}$ does not produce only separable restrictions. The next theorem provides a necessary and sufficient condition for entanglement of choice.

Theorem 1. *Thought experiment*$^{\lozenge}$ *produces only separable restrictions on* ρ *if and only if* $A^{1,\lozenge}$ *generates a unique representation.*

In the appendix, using the results in Aubrun et al. (2021), we generalize Theorem 1 to all finite choice sets, arbitrary menu structures, and any finite number of DMs.

Entanglement of choice happens in the thought experiment$^\Diamond$ with $A^{1,\Diamond} = A^1$ because of the lack of uniqueness. The latter leads to emerging restrictions in behavior that allow for the existence of ρ that satisfies the separable restrictions on joint behavior yet cannot come from separable choice.

Our results can be applied to other menu structures or other restrictions on $A^{1,\Diamond}$ (e.g., random utility and beyond), as long as uniqueness is guaranteed for one of the DMs. For example, in the literature of mathematical psychology researchers have studied the random interval and random semiorder models of stochastic choice introduced in Davis-Stober, Doignon, Fiorini, Glineur, and Regenwetter (2018) that generalize random utility. These models have conditions guaranteeing uniqueness (Doignon and Saito, 2023).

Nonuniqness of representation means that the underlying distribution over deterministic choice rules (i.e., columns of $A^{1,\Diamond}$) is partially- or set-identified (see, for instance, Kline and Tamer, 2023). When these distributions for both DMs are set-identified, then DMs may potentially coordinate on the individual distributions they can use. This coordination goes beyond correlation in choice. As a result, partial identification leads to behaviors that are not consistent with the thought experiment. However, when the distribution over deterministic choice rules of at least one DM is point identified, that DM will stick to this distribution, and there is no reason to coordinate with the other DM.

Chambers et al. (2021) provides a characterization of a version of our thought experiment with the random utility model (Block and Marschak, 1960, Falmagne, 1978, McFadden and Richter, 1990) restrictions on $A^{1\Diamond}$ and A^2. They, however, only show the sufficiency of the unique representation property of $A^{1,\Diamond}$. Theorem 1 shows that in their setting, uniqueness is necessary while the random utility restriction is not. Li (2021) considers a setting similar to Chambers et al. (2021) but with finitely many DM and at most three alternatives. As a result, the unique representation property of $A^{1,\Diamond}$ is satisfied. Because of that, the axiomatization in Li (2021) corresponds to the separable conditions generated by the interactions of the individual conditions. Our main result is a generalization of the characterizations in Chambers et al. (2021) and Li (2021).

4.1. An Example of Necessary and Sufficient Separable Restrictions for Separable Choice

Consider the matrix of allowable behavior of Frodo:

$$A^{1,\lozenge} = \begin{array}{c} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \end{array} \left\| \begin{array}{l} \text{x,\{xy\}} \\ \text{y,\{xy\}} \\ \text{y,\{yz\}} \\ \text{z,\{yz\}} \end{array} \right. .$$

In this thought experiment$^{\lozenge}$, a dominance restriction is induced by making y objectively worse than x or z for Frodo. It is direct to verify that the matrix above is generating. Crucially, $A^{1,\lozenge}$ generates a unique representation. For simplicity, we let the behavior of Sam be explained by the same matrix as in the thought experiment. The restrictions on the behavior of Frodo are given by

$$H^{1,\lozenge} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Sam's behavior is the same as in the original thought experiment with behavior captured by H^2. Note that compared to the original thought experiment, $H^{1,\lozenge}$ contains additional rows (rows 1 and 2). These rows correspond to two monotonicity restrictions: $\rho_1(x, \{xy\}) \geq \rho_1(y, \{yz\})$ and $\rho_1(z, \{yz\}) \geq \rho_1(y, \{xy\})$. They appear because of the dominance relation we imposed on y by removing one of the columns from A^1. Interactions of individual monotonicity restrictions with the nonnegativity restrictions in H^2 lead to monotonicity restrictions on ρ:

$$\rho\,(x, \{xy\}; m, M) - \rho\,(y, \{yz\}; m, M) \geq 0,$$

$$\rho\,(z, \{yz\}; m, M) - \rho\,(y, \{xy\}; m, M) \geq 0,$$

for all M and $m \in M$.

In this example, the above monotonicity restrictions are not the only separable restrictions beyond marginality. However, these extra restrictions are implied by monotonicity and marginality (i.e., they are redundant).

Proposition 2. The probabilistic choice rule ρ is generated by the above thought experiment$^\Diamond$ if and only if ρ satisfies marginality and monotonicity.

This proposition is a corollary of Theorem 1 and its proof is omitted for brevity. This proposition demonstrates that marginality and monotonicity are separable restrictions and are necessary and sufficient to describe the joint behavior of Sam and Frodo. This happens because Frodo's stochastic behavior is unique at the individual level. We highlight that in this case, entanglement of choice does not happen, because any ρ that satisfies the separable restrictions can be generated by the thought experiment$^\Diamond$. Note that the CHSH inequalities for the thought experiment$^\Diamond$ are valid. However, they are implied by marginality and monotonicity.

5. Conclusions

We hope our contribution extends a bridge between theoretical quantum physics and economics that could be fruitful to fully understand entangled and separable choice. Our result can be applied to obtain separable necessary conditions for any joint stochastic choice model expressed as a separable finite mixture of choice functions. When the model is generating and unique, we provide a full characterization.

References

Ashenfelter, O. and A. Krueger (1994): "Estimates of the economic return to schooling from a new sample of twins," *The American economic review*, 1157–1173. 1

Aubrun, G., L. Lami, C. Palazuelos, and M. Plávala (2021): "Entangleability of cones," *Geometric and Functional Analysis*, 31, 181–205. 2.1, 4, A.2

Bell, J. S. (1964): "On the einstein podolsky rosen paradox," *Physics Physique Fizika*, 1, 195. 1, 2.1

Block, H. and J. Marschak (1960): "Random Orderings and Stochastic Theories of Responses". in I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, and H. Man (eds) Contributions to Probability and Statistics, Stanford University Press," . 4

Chambers, C. P., Y. Masatlioglu, and C. Turansick (2021): "Correlated Choice," *arXiv preprint arXiv:2103.05084*. 1, 2, 2.1, 4

Davis-Stober, C. P., J.-P. Doignon, S. Fiorini, F. Glineur, and M. Regenwetter (2018): "Extended formulations for order polytopes through network flows," *Journal of mathematical psychology*, 87, 1–10. 4

De Paula, A. and X. Tang (2012): "Inference of signs of interaction effects in simultaneous games with incomplete information," *Econometrica*, 80, 143–172. 1

Doignon, J.-P. and K. Saito (2023): "Adjacencies on random ordering polytopes and flow polytopes," *Journal of Mathematical Psychology*, 114, 102768. 4

Falmagne, J.-C. (1978): "A representation theorem for finite random scale systems," *Journal of Mathematical Psychology*, 18, 52–72. 4

Fine, A. (1982): "Hidden variables, joint probability, and the Bell inequalities," *Physical Review Letters*, 48, 291. 2.1

Gurvits, L. (2003): "Classical deterministic complexity of Edmonds' problem and quantum entanglement," in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, 10–19. 1

Kashaev, N., V. H. Aguiar, M. Plávala, and C. Gauthier (2023): "Dynamic and stochastic rational behavior," *arXiv preprint arXiv:2302.04417*. ∗, 1

Kline, B. and E. Tamer (2023): "Recent developments in partial identification," *Annual Review of Economics*, 15, 125–150. 4

Li, R. (2021): "An Axiomatization of Stochastic Utility," *arXiv preprint arXiv:2102.00143*. 1, 4

McFadden, D. and M. K. Richter (1990): "Stochastic rationality and revealed stochastic preference," *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz, Westview Press: Boulder, CO*, 161–186. 4

Ray, D. and A. Robson (2018): "Certified random: A new order for coauthorship," *American Economic Review*, 108, 489–520. *

Rosset, D., J.-D. Bancal, and N. Gisin (2014): "Classifying 50 years of Bell inequalities," *Journal of Physics A: Mathematical and Theoretical*, 47, 424022. 1, 2, 2.1

Sacerdote, B. (2011): "Peer effects in education: How might they work, how big are they and how much do we know thus far?" in *Handbook of the Economics of Education*, Elsevier, vol. 3, 249–277. 1, 2.1

Strzalecki, T. (2021): *Stochastic Choice*, Mimeo. 1, 2

A. Proof of Theorem 1

We prove Theorem 1 at the highest level of generality. We consider multiple DMs $\mathcal{T} = \{1, \cdots, T\}$ (indexed without loss of generality) with a finite terminal agent $T \geq 1$. Let X^t be a nonempty finite choice set. In each $t \in \mathcal{T}$, there are $J^t < \infty$ distinct menus denoted by

$$B_j^t \in 2^{X^t} \setminus \{\emptyset\}, \quad j \in \mathcal{J}^t = \{1, \ldots, J^t\}.$$

Since X^t is a finite set, we denote the i-th element of menu $j \in \mathcal{J}^t$ as $x_{i|j}^t$. That is, $B_j^t = \{x_{i|j}^t\}_{i \in \mathcal{I}_j^t}$, where $\mathcal{I}_j^t = \{1, 2, \ldots, I_j^t\}$ and I_j^t is the number of elements in menu j.

Define a menu path as an ordered collection of indexes $\mathbf{j} = (j_t)_{t \in \mathcal{T}}$, $j_t \in \mathcal{J}^t$. Menu paths encode menus that are seen by each of the corresponding DMs at index t. Let $\mathbf{J}$ be the set of all *observed* menu paths. Given $\mathbf{j} \in \mathbf{J}$, a *choice path* is an array of alternatives $x_{\mathbf{i}|\mathbf{j}} = \left(x_{i_t|j_t}^t\right)_{t \in \mathcal{T}}$ for some collection of indexes $\mathbf{i} = (i_t)_{t \in \mathcal{T}}$ such that $i_t \in \mathcal{I}^t_{j_t}$ for all t. Similar to a menu path, a choice path encodes the choices of a DM in a given sequence of menus that DM have faced. The set of all possible choice path index sets $\mathbf{i}$, given a menu path $\mathbf{j}$, is denoted by $\mathbf{I_j}$.

Note that every $\mathbf{j} \in \mathbf{J}$ encodes the Cartesian product of menus $\times_{t \in \mathcal{T}} B^t_{j_t} \subseteq \times_{t \in \mathcal{T}} X^t$. Then, for every $\mathbf{j}$, let $\rho_\mathbf{j}$ be a probability measure on $\times_{t \in \mathcal{T}} B^t_{j_t}$. That is, $\rho_\mathbf{j}\left(x_{\mathbf{i}|\mathbf{j}}\right) \geq 0$ for all $\mathbf{i} \in \mathbf{I_j}$ and $\sum_{\mathbf{i} \in \mathbf{I_j}} \rho_\mathbf{j}\left(x_{\mathbf{i}|\mathbf{j}}\right) = 1$. The primitive in our framework is the collection of all observed $\rho_\mathbf{j}$, $\rho = (\rho_\mathbf{j})_{\mathbf{j} \in \boldsymbol{J}}$. We call this collection a *joint probabilistic choice rule.*

Given ρ we can define the generalized thought experiment in an analogous way to our original thought experiment but such that

$$\rho = (\otimes_{t=1}^T A^t)\nu$$

for some $\nu \geq 0$, and where columns of A^t collects all allowable choice functions of DM t.

$\mathcal{H}$- and $\mathcal{V}$-representations

For a single DM we can check if the probabilistic choice rule is consistent with the thought experiment at the individual level by checking that the stochastic choices of DM t belong to the cone

$$\left\{A^t v \,:\, v \geq 0\right\}.$$

This is called the $\mathcal{V}$-representation of the cone. The Weyl-Minkowski theorem states that there exists an equivalent representation of the cone (the $\mathcal{H}$-representation) via some matrix H_t:

$$\left\{z \,:\, H^t z \geq 0\right\}.$$

The $\mathcal{V}$-representation of the cone associated with the thought experiment provides an interpretation of the former as the observed distribution over choices is a finite mixture of deterministic types. The $\mathcal{H}$-representation of the cone associated with the thought experiment corresponds to what is usually called an axiomatization via linear inequalities. Importantly, these inequalities represent facets of the cone.

Proposition 3. If

$$\left\{K^t v \,:\, v \geq 0\right\} = \left\{z \,:\, L^t z \geq 0\right\}$$

for all $t \in \mathcal{T}$, then

$$\left\{\left(\otimes_{t\in\mathcal{T}} K^t\right) v \,:\, v \geq 0\right\} \subseteq \left\{z \,:\, \left(\otimes_{t\in\mathcal{T}} L^t\right) z \geq 0\right\}.$$

We say that the generalized experiment produces only separable restrictions on ρ whenever there exists a $\nu \geq 0$ such that $\rho = (\otimes_{t=1}^T A^t)\nu$ if and only if $(\otimes_{t=1}^T H^t)\rho \geq 0$ and ρ satisfies marginality.

Theorem 2. *The generalized experiment produces only separable restrictions on ρ if and only if for all $t \in \mathcal{T} \setminus \{t'\}$ for some $t' \in \mathcal{T}$ A^t is associated with a unique representation.*

A.1. Proof of Proposition 3

For completeness we provide here the proof of Proposition 3. Let $L_T = \otimes_{t=1}^T L^t$ and $K_T = \otimes_{t=1}^T K^t$. Note that for any v, z and $\otimes_{t=1}^T K^t$ such that $\left(\otimes_{t=1}^T K_t\right) v = z$ is well-defined, we can construct V and Z such that columns of V and Z are subvectors[2] of v and z and

$$\left(\otimes_{t=1}^T K^t\right) v = z \iff K^T V \left(\otimes_{t=1}^{T-1} K^t\right)' = Z.$$

Recall that by definition, $L^t K^t v \geq 0$ for all $v \geq 0$. Hence,

$$\begin{aligned}
&\forall v \geq 0,\ L^1K^1 v \geq 0 \implies \forall V \geq 0,\ L^2K^2V(L^1K^1)' \geq 0 \iff \\
&\forall v \geq 0,\ (L^1K^1 \otimes L^2K^2)v \geq 0 \implies \forall V \geq 0,\ L^3K^3V(L^1K^1 \otimes L^2K^2)' \geq 0 \iff \\
&\forall v \geq 0,\ (\otimes_{t=1}^3 L^tK^t)v \geq 0 \implies \forall V \geq 0,\ L^4K^4V(\otimes_{t=1}^3 L^tK^t)' \geq 0 \implies \\
&\cdots \implies \forall v \geq 0,\ (\otimes_{t=1}^T L^tK^t)v \geq 0 \iff \forall v \geq 0,\ L_TK_Tv \geq 0.
\end{aligned}$$

Hence,

$$\{K_T v \,:\, v \geq 0\} \subseteq \{z \,:\, L_T z \geq 0\}.$$

A.2. Proof of Theorem 2

First we show necessity of marginality. By definition the generalized thought experiment, there exists a distribution over $\mathcal{C}$, the collection of all choice function c that are mappings from the

[2] A vector x is a subvector of $y = (y_j)_{j\in J}$, if $x = (y_j)_{j\in J'}$ for some $J' \subseteq J$.

collection of menus in each t to alternatives, μ, such that

$$\rho\left(\left(x_{i_t|j_t}\right)_{t\in\mathcal{T}}\right) = \int \prod_{t\in\mathcal{T}} \mathbb{1}\left(c(B^t_{j_t}) = x^t_{i_t|j_t}\right) d\mu(c)$$

for all $\mathbf{i},\mathbf{j}$. Fix some $t'\in\mathcal{T}$, $x_{\mathbf{i}|\mathbf{j}}$, and $j_{t'}\in\mathcal{J}^{t'}$. Note that

$$\begin{aligned}
&\sum_{i\in\mathcal{I}^{t'}_{j_{t'}}} \rho\left(x_{\mathbf{i}|\mathbf{j}}\right) = \\
&\sum_{i\in\mathcal{I}^{t'}_{j_{t'}}} \int \mathbb{1}\left(c(B^{t'}_{j_{t'}}) = x^{t'}_{i|j_{t'}}\right) \prod_{t\in\mathcal{T}\setminus\{t'\}} \mathbb{1}\left(c(B^t_{j_t}) = x^t_{i_t|j_t}\right) d\mu(c) = \\
&\int \sum_{i\in\mathcal{I}^{t'}_{j_{t'}}} \mathbb{1}\left(c(B^{t'}_{j_{t'}}) = x^{t'}_{i|j_{t'}}\right) \prod_{t\in\mathcal{T}\setminus\{t'\}} \mathbb{1}\left(c(B^t_{j_t}) = x^t_{i_t|j_t}\right) d\mu(c) = \\
&\int \prod_{t\in\mathcal{T}\setminus\{t'\}} \mathbb{1}\left(c(B^t_{j_t}) = x^t_{i_t|j_t}\right) d\mu(c),
\end{aligned}$$

where the last equality follows from $c(B^{t'}_{j_{t'}})$ being a singleton and $\{x^{t'}_{i|j_{t'}}\}_{i\in\mathcal{I}^{t'}_{j_{t'}}}$ being a partition. The right-hand side of the last expression does not depend on the choice of $j_{t'}$. marginality follows from t' and $x_{\mathbf{i}|\mathbf{j}}$ being arbitrary.

Next, we show that any ρ that satisfies marginality belongs to a linear span of columns of A_T. That is, the system $A_T v = \rho$ always has a solution and the cone generated by A_T is proper when restricted to ρ that satisfies marginality. Hence, Theorem 2 follows from Theorem A in Aubrun et al. (2021).

Consider the following modification of A^t, $t\in\mathcal{T}$. From every menu, except the first one, we pick the last alternative and remove the corresponding row from A^t. Let A^{t*} denote the resulting matrix. Thus, matrix A^t can be partitioned into A^{t*} and A^{t-}, where rows of A^{t-} correspond to alternatives removed from A^t. Consider the first row of A^{t-}. It corresponds to the last alternative from the second menu at time t. Note that the sum of all rows that correspond to the same menu is equal to the row of ones. Hence, the first row of A^{t-} is equal to the sum of the rows that correspond to menu 1 minus the sum of the remaining rows in menu 2. That is, the first row of A^{t-} can be written as

$$(1,\ldots,1,-1,\ldots,-1,0,\ldots,0)A^{t*}.$$

Similarly, the second row of A^{t-} can be written as

$$(1, \ldots, 1, 0, \ldots, 0, -1, \ldots, -1, 0, \ldots, 0)A^{t*}.$$

In matrix notation, $A^{t-} = G^t A^{t*}$, where G^t is the matrix with the k-th row having the elements that correspond to the alternatives from the first menu at time t are equal to 1, the elements that correspond to the alternatives from the k-th menu are equal to -1, and the rest of elements are equal to 0.

Next note that, up to a permutation of rows, A_T can be partitioned into $A^*_T = \otimes_{t\in\mathcal{T}} A^{t*}$ and matrices of the form $\otimes_{t\in\mathcal{T}} C^t$, where $C^t \in \{A^{t*}, A^{t-}\}$, with $C^t = A^{t-}$ for at least one t. We will stack all these matrices into A^-_T. Next, let ρ^* denote the subvector of ρ that corresponds to choice paths that do not contain any of the alternatives removed from A^t, $t \in \mathcal{T}$. Thus, $\rho = (\rho^{*\prime}, \rho^{-\prime})'$, where ρ^- corresponds to all elements of ρ that contain at least one of the removed alternatives. As a result, we can split the original system into two: $A^*_T v = \rho^*$ and $A^-_T v = \rho^-$.

Consider the system $A^*_T v = \rho^*$. Since we only removed one row from each menu except the first one and A^t can generate any ρ_t, A^{t*} has full row rank for all t. Then A^*_T is also of full row rank and, hence, $A^*_T A^{*\prime}$ is invertible and $v^* = A^{*\prime}\left(A^*_T A^{*\prime}\right)^{-1}\rho^*$ solves the system. If, we show that

$$A^-_T v^* = \rho^-,$$

then we prove that $A_T v = \rho$ always has a solution, which will complete the proof.

Note that A^-_T consists of the blocks of the form $\otimes_{t\in\mathcal{T}} C^t$, where $C^t \in \{A^{t*}, A^{t-}\}$ and $C^t = A^{t-}$ for at least one t. Next note that for any A, B, and C, we have that

$$A \otimes (BC) = diag(B)(A \otimes C),$$

where $diag(B)$ is the block-diagonal matrix constructed from B.

Let W^t (with inverse $W^{t,-1}$, which pushes the last element of $\mathcal{T}$ to t-th position) be a transformation that recomputes all objects for the time span where t is pushed to the end. Transformation W^t satisfies the following three properties: $W^t[C] = C$ if C does not depend on $\mathcal{T}$; $W^t[CD] = W^t[C]W^t[D]$ for any matrices C and D; and $W^t[\otimes_{t'\in\mathcal{T}} A^{t'*}] = \otimes_{t'\in\mathcal{T}\setminus\{t\}} A^{t'*} \otimes A^{t*}$. Let Y^t be an operator such that $Y^t[\cdot] = W^{t,-1}\left[diag(G^t)W^t[\cdot]\right]$.

Consider $\otimes_{t\in\mathcal{T}}C^t$, where $C^t \in \{A^{t*}, A^{t-}\}$ and $C^t = A^{t-}$ for only one t. Hence,

$$\otimes_{t'\in\mathcal{T}} C^{t'} v^* = W^{t,-1}\left[diag(G^t)W^t\left[\rho^*\right]\right] = Y^t[\rho^*].$$

Note that because ρ satisfies marginality, $diag(G^T)\rho^*$ is the subvector of ρ^- that corresponds to choice paths that contain one of the removed alternatives from the last DM only. So, $W^t\left[\rho^*\right]$ first pushes DM t to the very end, then $diag(G^t)W^t\left[\rho^*\right]$ computes the elements of ρ^-, and finally $W^{t,-1}\left[diag(G^t)W^t\left[\rho^*\right]\right]$ moves DM t back to her place.

Next, consider $\otimes_{t\in\mathcal{T}}C^t$, where $C^t \in \{A^{t*}, A^{t-}\}$ and $C^t = A^{t-}$ and $C^{t'} = A^{t'-}$ for two distinct t, t'. Similarly to the previous case,

$$\otimes_{t'\in\mathcal{T}} C^{t'} v^* = W^{t,-1}\left[diag(G^t)W^t\left[W^{t',-1}\left[diag(G^{t'})W^{t'}\left[\rho^*\right]\right]\right]\right] = Y^t[Y^{t'}[\rho^*]] = Y^t \circ Y^{t'}[\rho^*],$$

where $Y^t \circ Y^{t'}$ denotes the composite operator. Again, $W^{t',-1}\left[diag(G^{t'})W^{t'}\left[\rho^*\right]\right]$ computes the subvector of ρ^- that corresponds to choice paths where an alternative from only one time t' was missing. Applying to the resulting vector $W^{t,-1}\left[diag(G^t)W^t\left[\cdot\right]\right]$ computes the subvector of ρ^- with alternatives missing from t and t' only. Repeating the arguments for all possible rows of A^-_T, we obtain that

$$\otimes_{t'\in\mathcal{T}} C^{t'} v^* = \circ_{t':C^{t'}=A^{t'-}} Y^{t'}[\rho^*]$$

and, by marginality, $A^-_T v^* = \rho^-$. Hence, v^* is a solution to $A_T v = \rho$.