On The Fly Reduction of Polynomial Differential Equations

John Doe
Aalborg University, Denmark
john.doe@cs.aau.dk

Abstract

Here goes the abstract.

1. Introduction

Here goes the intro.

Related work. Do not forget Boreale's work.

2. Preliminaries

[Max: We need to pick a running example.]

Notation. We assume S to be a finite set of indices ranged over by $x, y, z, \ldots \in S$. For two sets A, B we write $A \to B$ or B^A to denote the set of functions from A to B. A *drift* is a map $f \colon \mathbb{R}^S \to \mathbb{R}^S$ that is totally differentiable. A drift f is *polynomial* when, for all $x \in S$, f_x is a polynomial over S.

Comparison of polynomial ODE systems. Let us fix a polynomial drift $f : \mathbb{R}^S \to \mathbb{R}^S$ and use Newton's dot notation $\dot{v} = f(v)$ to denote the ODE system given by $\dot{v}_x = f_x(v) = f(v)(x)$, where $x \in S$. Given an initial condition $v(0) \in \mathbb{R}^S$, Picard-Lindelöf's theorem ensures that $\dot{v} = f(v)$ has a unique solution $v : dom(v) \to \mathbb{R}^S, t \mapsto v(t)$.

We recall the notion of backward equivalence (BE).

Definition 1 (Backward equivalence)**.** *Let $f \colon \mathbb{R}^S \to \mathbb{R}^S$ be a polynomial drift. An equivalence relation $\mathcal{R} \subseteq S \times S$ is a BE if the implication*

$$\Big(\bigwedge_{(x,y)\in\mathcal{R}} v_x = v_y \Big) \Rightarrow \Big(\bigwedge_{(x,y)\in\mathcal{R}} f_x(v) = f_y(v) \Big)$$

is true for any $v \in \mathbb{R}^S$. To improve readability, we shall say that a vector $v \in \mathbb{R}^S$ respects $\mathcal{R}$ if $v_x = v_y$ for all $x, y \in \mathcal{R}$.

The interest in studying BE is motivated by the fact that backward equivalent indices enjoy identical values if initialized equally [6].

Theorem 1. *Fix a polynomial drift $f : \mathbb{R}^S \to \mathbb{R}^S$ and some equivalence relation $\mathcal{R} \subseteq \mathcal{S} \times \mathcal{S}$. Then, $\mathcal{R}$ is a BE if and only if for any $v(0) \in \mathbb{R}^S$ that respects $\mathcal{R}$, it holds true that the solution of $\dot{v} = f(v)$ with initial condition $v(0)$ respects $\mathcal{R}$ as well, i.e., $v(t)$ respects $\mathcal{R}$ for all $t \in dom(v)$.*

Proof. The claim follows from Theorem 3 in [6] (please note that BE is called uniform lumpability in [6]). □

Example 1. *Let $S = \{x, y\}$ and consider the drift*

$$f_x(v) = -2 \cdot v_x^2 + v_y^2 \qquad f_y(v) = -2 \cdot v_y^2 + v_x^2$$

Then, the partition $\{\{x, y\}\}$ of S is BDE and $v_x(t) = v_y(t)$ for all $t \geq 0$ whenever $v_x(0) = v_y(0)$.

The following result ensures that the largest BE exists and that it can be efficiently computed using partition refinement [4].

Theorem 2. *For any partition $\mathcal{G}$ of S, there exists a largest BE $\mathcal{R}$ such that $\mathcal{S}/\mathcal{R}$ refines $\mathcal{G}$. The BE $\mathcal{R}$ in question can be computed by means of a partition refinement algorithm whose time and space complexity is polynomial in the size of the polynomial drift, i.e., the number of monomials present in all right-hand sides of $\dot{v} = f(v)$.*

Proof. In [4], it is first demonstrated that polynomial ODE systems are equivalent to so-called reaction networks. With this in place, [4] provides a syntactic definition of BE which a) is shown to be equivalent to the conclusion of Theorem 1 and that b) can be computed in polynomial time and space via a partition refinement algorithm. □

The notion of emulation presented next has been introduced in [1] and relates a source ODE system to a target ODE system.

Definition 2 (Emulation)**.** *Fix a source polynomial drift $f : \mathbb{R}^S \to \mathbb{R}^S$ and target polynomial drift $\hat{f} : \mathbb{R}^{\hat{S}} \to \mathbb{R}^{\hat{S}}$. The surjective function $\mu : S \to \hat{S}$ is an emulation if $f_x(\hat{v} \circ \mu) = \hat{f}_{\mu(x)}(\hat{v})$ for all $\hat{v} \in \mathbb{R}^{\hat{S}}$ and $\hat{x} \in \hat{S}$.*

Note that $\hat{v} \circ \mu \in \mathbb{R}^S$ and $(\hat{v} \circ \mu)_x = \hat{v}_{\mu(x)}$ for all $x \in S$. Two ODE systems are related by means of an emulation if the trajectories of the source ODE system coincide with those of the target ODE system whenever the initial conditions of both systems are equal with respect to μ, as has been proven in [1] and is stated next.

Theorem 3. *Let $\mu : S \to \hat{S}$ be an emulation between the source polynomial drift $f : \mathbb{R}^S \to \mathbb{R}^S$ and the target polynomial drift $\hat{f} : \mathbb{R}^{\hat{S}} \to \mathbb{R}^{\hat{S}}$. Then, $v(t) = \hat{v}(t) \circ \mu$ for all $t \in dom(v)$ whenever $v(0) = \hat{v}(0) \circ \mu$.*

Example 2. *Consider $\hat{S} = \{\hat{x}\}$ and $\hat{f}_{\hat{x}}(\hat{v}) = -\hat{v}_{\hat{x}}$. Then the function $\mu(x) = \mu(y) = \hat{x}$ is an emulation between $f : \mathbb{R}^{\{x,y\}} \to \mathbb{R}^{\{x,y\}}$ and $\hat{f} : \mathbb{R}^{\{\hat{x}\}} \to \mathbb{R}^{\{\hat{x}\}}$, where f is as in Example* **??**. *In particular, it holds that $\hat{v}_{\hat{x}}(t) = v_x(t) = v_y(t)$ for all $t \geq 0$ whenever $\hat{v}_{\hat{x}}(0) = v_x(0) = v_y(0)$.*

There is a close relation between emulation and BE, as has been observed in [2, 3].

Proposition 1. *Fix a source polynomial drift $f : \mathbb{R}^S \to \mathbb{R}^S$, a target polynomial drift $\hat{f} : \mathbb{R}^{\hat{S}} \to \mathbb{R}^{\hat{S}}$ with $S \cap \hat{S} = \emptyset$ and a* *surjective function $\mu : S \to \hat{S}$. Then, μ is an emulation if and only if $\{\mu^{-1}(\hat{x}) \cup \{\hat{x}\} \mid \hat{x} \in \hat{S}\}$ are the equivalence classes of a BE.*

Chemical Reaction Networks. A CRN (S, R) is a pair consisting of a finite set of species (indices) S and a finite set of chemical reactions R. A reaction is a triple written in the form $\rho \xrightarrow{\alpha} \pi$, where ρ and π are the multisets of species *reactants* and *products*, respectively, and $\alpha > 0$ is the reaction rate parameter. We denote by $\rho(x)$ the multiplicity of species x in the multiset ρ. The net stoichiometry of a species x in a reaction $r = \rho \xrightarrow{\alpha} \pi$ is the difference between product and reactant multiplicity, times the rate coefficient α, i.e., $\alpha \cdot (\pi(x) - \rho(x))$, and describes the amount of substance x transformed through reaction r in a time unit. A given $\mu : S \to \hat{S}$ can be trivially lifted to multisets over S, e.g., $\mu(x + y) = \mu(x) + \mu(y)$.

The ODE system $\dot{v} = f(v)$ underlying a CRN (S, R) is $f : \mathbb{R}^S \to \mathbb{R}^S$, where each component f_x, with $x \in S$, is given by

$$f_x(v) := \sum_{\rho \to^{\alpha} \pi \in R} (\pi(x) - \rho(x)) \cdot \alpha \cdot \prod_{y \in S} v_y^{\rho(y)}$$

This represents the well-known *mass-action* kinetics, where the reaction rate is proportional to the concentrations of the reactants involved. It readily follows that each CRN induces a polynomial drift.

Example 3. *The network $(\{x, y\}, \{x + x \to^1 y, y + y \to^1 x\})$ induces the drift of Example* **??**.

3. The Backward Relation

We next provide an alternative but equivalent characterisation of BEs phrased in terms of couplings which constitutes the basis for our on-the-fly algorithm.

We denote by $\mathbf{M}[S]$ the set of monomials over S, ranged over by $m, n, \ldots \in \mathbf{M}[S]$ and by $\mathbf{P}[S]$ the set of polynomials over S, ranged over by $p, q, \ldots \in \mathbf{P}[S]$. Throughout the paper we will assume polynomials to be expressed as linear combinations of monomials, i.e., $p = \sum_{i=0}^{k} \alpha_i m_i$ for some $\alpha_i \in \mathbb{R}$. For example, the polynomial $(x + y)^2$ will be assumed to be already expressed in the form $x^2 + 2xy + y^2$. (This is not just for convenience, but will be needed to unambiguously define the size of a polynomial ODE system as the input of our algorithm). Any polynomial p can be decomposed into its *positive* and *negative* parts as $p = p^+ - p^-$, where $p^+ = \sum_{i=0}^{k} \alpha_i m_i$ and $p^- = \sum_{j=0}^{h} \beta_j n_j$, for some $\alpha_i, \beta_j > 0$ positive coefficients. This decomposition is unique once we impose the monomials m_i and n_j to be pair-wise distinct. For a monomial $m = \prod_{i=0}^{k} x_i^{a_i}$ we write $m(x_i)$ for the exponent a_i associated to the index x_i, and for a polynomial $p = \sum_{i=0}^{k} \alpha_i m_i$, we write $p(m_i)$ for the coefficient α_i associated to the monomial m_i. When S is clear from the context, we will simply write $\mathbf{M}$ for $\mathbf{M}[S]$ and $\mathbf{P}$ for $\mathbf{P}[S]$.

Next we introduce the concepts of couplings for pairs of monomials and pairs of polynomials. Intuitively, couplings describe a transportation of positive mass. [Giovanni: We should provide a better intuition of couplings.]

Definition 3 (Monomial coupling)**.** *Let m, n be two monomials over S. We say that $\rho \colon S \times S \to \mathbb{R}_{\geq 0}$ is a* monomial coupling *for (m, n) if the following conditions hold*

(i) $\forall x \in S.\ \sum_{y \in S} \rho(x, y) = m(x)$;
(ii) $\forall y \in S.\ \sum_{x \in S} \rho(x, y) = n(y)$;

Definition 4 (Polynomial coupling)**.** *Let p, q be two polynomials over S. We say that $\omega \colon \mathbf{M} \times \mathbf{M} \to \mathbb{R}_{\geq 0}$ is a* polynomial coupling *for (p, q) if the following conditions hold*

(i) $\forall m \in \mathbf{M}.\ \sum_{n \in \mathbf{M}} \omega(m, n) = (p^+ + q^-)(m)$;
(ii) $\forall n \in \mathbf{M}.\ \sum_{m \in \mathbf{M}} \omega(m, n) = (q^+ + p^-)(n)$;

where $p = p^+ - p^-$ and $q = q^+ - q^-$.

[Giovanni: The decompositions $p = p^+ - p^-$ and $q = q^+ - q^-$ may not be unique. Nevertheless the results do not depend by the particular choice of the decomposition.]

For arbitrary relations $\mathcal{R} \subseteq S \times S$ and $\mathcal{Q} \subseteq \mathbf{M} \times \mathbf{M}$ we define $\mathbf{M}[\mathcal{R}]$ and $\mathbf{P}[\mathcal{Q}]$ as

$$\mathbf{M}[\mathcal{R}] = \{(m, n) \mid \exists \rho \in \Gamma_{\mathbf{M}}(m, n) \text{ such that } \mathit{supp}(\rho) \subseteq \mathcal{R}\} \subseteq \mathbf{M} \times \mathbf{M}\,,$$
$$\mathbf{P}[\mathcal{Q}] = \{(p, q) \mid \exists \omega \in \Gamma_{\mathbf{P}}(p, q) \text{ such that } \mathit{supp}(\omega) \subseteq \mathcal{Q}\} \subseteq \mathbf{P} \times \mathbf{P}\,.$$

We call $\mathbf{M}[\mathcal{R}]$ the lifting of $\mathcal{R}$ over monomials, and $\mathbf{P}[\mathcal{Q}]$ the lifting of $\mathcal{Q}$ over polynomials.

Lemma 1. *For all $\mathcal{R} \subseteq S \times S$ and $v \in \mathbb{R}^S$*

$$\Big(\bigwedge_{(x,y) \in \mathcal{R}} v_x = v_y \Big) \implies \Big(\bigwedge_{(m,n) \in \mathbf{M}[\mathcal{R}]} m(v) = n(v) \Big)\,.$$

Lemma 2. *For all $\mathcal{Q} \subseteq \mathbf{M} \times \mathbf{M}$ and $v \in \mathbb{R}^S$*

$$\Big(\bigwedge_{(m,q) \in \mathcal{Q}} m(v) = n(v) \Big) \implies \Big(\bigwedge_{(p,q) \in \mathbf{P}[\mathcal{Q}]} p(v) = q(v) \Big)\,.$$

The next two lemmas state that the lifting an equivalence relation induces an equivalence relation in the target domain.

Lemma 3. *Let $\mathcal{R} \subseteq S \times S$. If $\mathcal{R}$ is an equivalence relation, then also $\mathbf{M}[\mathcal{R}]$ is.*

Lemma 4. *Let $\mathcal{Q} \subseteq \mathbf{M} \times \mathbf{M}$. If $\mathcal{Q}$ is an equivalence relation, then also $\mathbf{P}[\mathcal{Q}]$ is.*

Corollary 1. *If $\mathcal{R} \subseteq S \times S$ is an equivalence relation, then also $\mathbf{P}[\mathbf{M}[\mathcal{R}]]$ is.*

Definition 5. *Let $f \colon \mathbb{R}^S \to \mathbb{R}^S$ be a polynomial drift. A relation $\mathcal{R} \subseteq S \times S$ is a BR if $(x, y) \in \mathcal{R}$ implies $(f_x, f_y) \in \mathbf{P}[\mathbf{M}[\mathcal{R}]]$.*

Proposition 2. *Let $f \colon \mathbb{R}^S \to \mathbb{R}^S$ be a polynomial drift and $\mathcal{R}$ a BR. Then, for any $v \in \mathbb{R}^S$ the following implication holds*

$$\Big(\bigwedge_{(x,y) \in \mathcal{R}} v_x = v_y \Big) \Rightarrow \Big(\bigwedge_{(x,y) \in \mathcal{R}} f_x(v) = f_y(v) \Big)$$

The following proposition states that a BR can be used as a witness for proving that some indices are related by some BE.

Proposition 3. *Let $f \colon \mathbb{R}^S \to \mathbb{R}^S$ be a polynomial drift and $\mathcal{R} \subseteq S \times S$. If $\mathcal{R}$ is a BR then $\mathcal{R}^*$ is a BE.*

The next result states that the notion of BR generalises the notion of BE.

Proposition 4. *Let $f \colon \mathbb{R}^S \to \mathbb{R}^S$ be a polynomial drift. If $\mathcal{R}$ is a BE then $\mathcal{R}$ is a BR.*

Theorem 4. *The greatest backward equivalence and the greatest backward relation coincide.*

Let $f \colon \mathbb{R}^S \to \mathbb{R}^S$ be a polynomial drift. We define $\mathcal{B} \colon 2^{S \times S} \to 2^{S \times S}$ as follows

$$\mathcal{B}(\mathcal{R}) = \{(x, y) \in S \times S \mid (f_x, f_y) \in \mathbf{P}[\mathbf{M}[\mathcal{R}]]\} \qquad (1)$$

Clearly $\mathcal{B}$ is monotonic w.r.t. $\subseteq$ (i.e., $\mathcal{R} \subseteq \mathcal{R}'$ implies $\mathcal{B}(\mathcal{R}) \subseteq \mathcal{B}(\mathcal{R}')$). By Definition 5, a relation $\mathcal{R}$ is a BR if and only if $\mathcal{R} \subseteq \mathcal{B}(\mathcal{R})$. Hence, by Knaster-Tarski's fixed point theorem we have that the greatest fixed point of $\mathcal{B}$, namely $\bigcup\{\mathcal{R} \mid \mathcal{R} \subseteq \mathcal{B}(\mathcal{R})\}$, corresponds to the greatest BR.

4. Algorithm

On-the-fly-PartRef($Q \subseteq S \times S, \bar{Q} \subseteq S \times S$)

// It is assumed that $Q \cap \bar{Q} = \emptyset$
let Adj be the empty map, and $F = \bar{Q}$
for each $(x, y) \in Q$
  Update(Adj, F, v)
while $\exists v \in dom(Adj).\, Adj(v) \cap F \neq \emptyset$
  Update(Adj, F, v)
return $dom(Adj) \cap (S \times S)$

Update($Adj, F, (x, y) \in S \times S$)

if $\exists \omega \in \Gamma_{\mathbf{P}}(f_x, f_y)\, .supp(\omega) \cap F = \emptyset$
  $Adj(x, y) = supp(\omega)$
  Expand($Adj, F, (x, y)$)
else
  $F = F \cup \{(x, y)\}$
  remove (x, y) from $dom(Adj)$

Update($Adj, F, (m, n) \in \mathbf{M} \times \mathbf{M}$)

if $\exists \rho \in \Gamma_{\mathbf{M}}(m, n)\, .supp(\rho) \cap F = \emptyset$
  $Adj(m, n) = supp(\rho)$
  Expand($Adj, F, (m, n)$)
else
  $F = F \cup \{(m, n)\}$
  remove (m, n) from $dom(Adj)$

Expand(Adj, F, v)

while $\exists u \in Adj(v) \setminus \big(dom(Adj) \cup F\big)$
  Update(Adj, F, u)

Theorem 5. *Let $Q, \bar{Q} \subseteq S \times S$ such that $Q \cap \bar{Q} = \emptyset$ and let $\mathcal{R} =$ On-the-fly-PartRef$(Q, \bar{Q})$. The following statements hold*

(i) $\mathcal{R}$ is a BR such that $\mathcal{R} \cap \bar{Q} = \emptyset$;
(ii) if $(x, y) \in Q$, then $(x, y) \in \mathcal{R}$ iff $(x, y) \in \mathcal{R}'$ for some BR $\mathcal{R}'$ such that $\mathcal{R}' \cap \bar{Q} = \emptyset$.

Proof. TO DO ☐

Corollary 2. *Let $Q \subseteq S \times S$ and $\mathcal{R} =$ On-the-fly-PartRef$(Q, \emptyset)$. The following statements hold*

(i) $\mathcal{R}$ is a BR;
(ii) if $(x, y) \in Q$, then $(x, y) \in \mathcal{R}$ if and only if $(x, y) \in \mathcal{R}'$ for some BR $\mathcal{R}'$.

5. Evaluation

References

[1] Luca Cardelli. Morphisms of reaction networks that couple structure to function. *BMC Systems Biology*, 8(1):84, 2014.

[2] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. Forward and backward bisimulations for chemical reaction networks. In *CONCUR*, pages 226–239, 2015.

[3] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. Comparing chemical reaction networks: A categorical and algorithmic perspective. In *Proceedings of the Thirty-First Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, 2016.

[4] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. Maximal aggregation of polynomial dynamical systems. *Proceedings of the National Academy of Sciences*, 114(38):10029–10034, 2017.

[5] Philippe Clement and Wolfgang Desch. An elementary proof of the triangle inequality for the wasserstein metric. *Proceedings of the American Mathematical Society*, 136(1):333–339, 2008.

[6] Max Tschaikowski and Mirco Tribastone. Approximate Reduction of Heterogenous Nonlinear Models With Differential Hulls. *IEEE Trans. Automat. Contr.*, 61(4):1099–1104, 2016.

Proofs

Proof of Lemma 1. Let $(m,n) \in \mathbf{M}[\mathcal{R}]$ and let $v \in \mathbb{R}^S$ such that $\bigwedge_{(x,y)\in\mathcal{R}} v_x = v_y$. By hypothesis there exists $\rho \in \Gamma_{\mathbf{M}}(m,n)$ such that $supp(\rho) \subseteq \mathcal{R}$. Then, the following hold:

$$\begin{aligned} m(v) &= \prod_{x\in S} v_x^{m(x)} \\ &= \prod_{x\in S} v_x^{\sum_{y\in S}\rho(x,y)} && (\rho \in \Gamma_{\mathbf{M}}(m,n)) \\ &= \prod_{x,y\in S} v_x^{\rho(x,y)} \\ &= \prod_{(x,y)\in\mathcal{R}} v_x^{\rho(x,y)}\,. && (supp(\rho) \subseteq \mathcal{R}) \end{aligned}$$

Analogously, one can prove that $n(v) = \prod_{(x,y)\in\mathcal{R}} v_y^{\rho(x,y)}$. By hypothesis on v we have that $v_x = v_y$ for all $(x,y) \in \mathcal{R}$, hence $m(v) = n(v)$. ☐

Proof of Lemma 2. Let $(p,q) \in \mathbf{P}[\mathcal{Q}]$ and let $v \in \mathbb{R}^S$ such that $\bigwedge_{(m,n)\in\mathcal{Q}} m(v) = n(v)$. By hypothesis there exists $\omega \in \Gamma_{\mathbf{P}}(p,q)$ such that $supp(\omega) \subseteq \mathcal{Q}$. Then, the following hold

$$\begin{aligned} (p^+ + q^-)(v) &= \sum_{m\in\mathbf{M}} (p^+ + q^-)(m)\cdot m(v) \\ &= \sum_{m\in\mathbf{M}} \big(\sum_{n\in\mathbf{M}} \omega(m,n)\big)\cdot m(v) && (\omega \in \Gamma_{\mathbf{P}}(p,q)) \\ &= \sum_{m,n\in\mathbf{M}} \omega(m,n)\cdot m(v) \\ &= \sum_{(m,n)\in\mathcal{Q}} \omega(m,n)\cdot m(v)\,. && (supp(\omega) \subseteq \mathcal{Q}) \end{aligned}$$

Analogously, one can prove that $(p^- + q^+)(v) = \sum_{(m,n)\in\mathcal{Q}} \omega(m,n)\cdot n(v)$. By hypothesis on v we have that $m(v) = n(v)$ for all $(m,n) \in \mathcal{Q}$, therefore we have $(p^+ + q^-)(v) = (p^- + q^+)(v)$, which is equivalent to $p(v) = q(v)$. ☐

Proof of Lemma 3. Reflexivity and symmetry of $\mathbf{M}[\mathcal{R}]$ can be easily proven. We prove transitivity of $\mathbf{M}[\mathcal{R}]$ using a coupling construction similar to [5, Prop. 2.1]. Let $\rho_{1,2}$ and $\rho_{2,3}$ be two monomial couplings witnessing respectively $(m_1,m_2) \in \mathbf{M}[\mathcal{R}]$ and $(m_2,m_3) \in \mathbf{M}[\mathcal{R}]$. Define $\rho\colon S\times S\times S \to \mathbb{R}_{\geq 0}$ as

$$\rho(x,y,z) = \begin{cases} \frac{\rho_{1,2}(x,y)\cdot\rho_{2,3}(y,z)}{m_2(y)} & \text{if } m_2(y) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Let $\rho_{1,3}\colon S\times S \to \mathbb{R}_{\geq 0}$ as $\rho_{1,3}(x,z) = \sum_{y\in S}\rho(x,y,z)$, then for any $z \in S$ the following equalities hold:

$$\begin{aligned} \sum_{x\in S}\rho_{1,3}(x,z) &= \sum_{x\in S}\sum_{y\in S}\rho(x,y,z) \\ &= \sum_{y\in S}\sum_{x\in S}\frac{\rho_{1,2}(x,y)\cdot\rho_{2,3}(y,z)}{m_2(y)} \\ &= \sum_{y\in S}\rho_{2,3}(y,z) && (\textstyle\sum_{x\in S}\rho_{1,2}(x,y) = m_2(y)) \\ &= m_3(z) \end{aligned}$$

Similarly, for any $x \in S$, $\sum_{z\in S}\rho_{1,3}(x,z) = m_1(x)$. Hence $\rho_{1,3} \in \Gamma_{\mathbf{M}}(m_1,m_3)$. Moreover, the following chain of implications hold

$$\begin{aligned} (x,z) \in supp(\rho_{1,3}) &\Rightarrow \rho_{1,3}(x,y) > 0 \\ &\Rightarrow \exists y \in S\,.m_2(y) > 0 \wedge \rho_{1,2}(x,y) > 0 \wedge \rho_{2,3}(y,z) > 0 \\ &\Rightarrow \exists y \in S\,.\rho_{1,2}(x,y) > 0 \wedge \rho_{2,3}(y,z) > 0 \\ &\Rightarrow \exists y \in S\,.(x,y) \in supp(\rho_{1,2}) \wedge (y,z) \in supp(\rho_{2,3}) \\ &\Rightarrow \exists y \in S\,.(x,y) \in \mathcal{R} \wedge (y,z) \in \mathcal{R} \\ &\Rightarrow (x,z) \in \mathcal{R} \end{aligned}$$

This concludes the proof. ☐

Proof of Lemma 4. Reflexivity and symmetry are easy to prove. We show that $\mathbf{P}[\mathcal{Q}]$ is transitive. Note that $(p,q) \in \mathbf{P}[\mathcal{Q}]$ if and only if the optimal value of the following linear program equals 0

$$\begin{aligned} &\min_{\omega} \textstyle\sum_{(m,n)\notin\mathcal{Q}} \omega(m,n) \\ &\quad \textstyle\sum_n \omega(m,n) = (p^+ + q^-)(m) \quad && \forall m \in \mathbf{M} \\ &\quad \textstyle\sum_m \omega(m,n) = (q^+ + p^-)(n) && \forall n \in \mathbf{M} \\ &\quad \omega(m,n) \geq 0 && \forall m,n \in \mathbf{M} \end{aligned} \tag{Primal}$$

Since $\mathcal{Q}$ is an equivalence relation we can formulate the dual of the above linear program as follows.

$$\begin{aligned} &\max_{\alpha} \textstyle\sum_{m\in\mathbf{M}} (p(m) - q(m)) \cdot \alpha(m) \\ &\quad \alpha(m) - \alpha(n) \leq 0 \quad && \forall (m,n) \in \mathcal{Q} \\ &\quad \alpha(m) - \alpha(n) \leq 1 && \forall (m,n) \notin \mathcal{Q} \end{aligned} \tag{Dual}$$

We denote by $F[\mathcal{Q}] \subseteq \mathbb{R}^{\mathbf{M}}$ the set of feasible solutions of the above linear program.

If $(p_1,p_2),(p_2,p_3) \in \mathbf{P}[\mathcal{Q}]$, then the following inequalities hold

$$\begin{aligned} 0 &\leq \min_{\omega\in\Gamma_{\mathbf{P}}(p_1,p_3)} \sum_{(m,n)\notin\mathcal{Q}} \omega(m,n) && (\omega(m,n) \geq 0 \text{ for all } m,n \in \mathbf{M}) \\ &= \max_{\alpha\in F[\mathcal{Q}]} \sum_{m\in\mathbf{M}} (p_1 - p_3)(m) \cdot \alpha(m) && \text{(duality)} \\ &= \max_{\alpha\in F[\mathcal{Q}]} \sum_{m\in\mathbf{M}} (p_1 - p_3 + p_2 - p_2)(m) \cdot \alpha(m) \\ &= \max_{\alpha\in F[\mathcal{Q}]} \left(\sum_{m\in\mathbf{M}} (p_1 - p_2)(m) \cdot \alpha(m) + \sum_{m\in\mathbf{M}} (p_2 - p_3)(m) \cdot \alpha(m) \right) \\ &\leq \left(\max_{\alpha\in F[\mathcal{Q}]} \sum_{m\in\mathbf{M}} (p_1 - p_2)(m) \cdot \alpha(m) \right) + \left(\max_{\alpha\in F[\mathcal{Q}]} \sum_{m\in\mathbf{M}} (p_2 - p_3)(m) \cdot \alpha(m) \right) \\ &= \left(\min_{\omega\in\Gamma_{\mathbf{P}}(p_1,p_2)} \sum_{(m,n)\notin\mathcal{Q}} \omega(m,n) \right) + \left(\min_{\omega\in\Gamma_{\mathbf{P}}(p_2,p_3)} \sum_{(m,n)\notin\mathcal{Q}} \omega(m,n) \right) && \text{(duality)} \\ &= 0 && ((p_1,p_2),(p_2,p_3) \in \mathbf{P}[\mathcal{Q}]) \end{aligned}$$

All the above inequalities are in fact equalities, proving that $(p_1,p_3) \in \mathbf{P}[\mathcal{Q}]$. □

Proof of Proposition 2. For an arbitrary $v \in \mathbb{R}^S$ the following chain of logical implications hold

$$\begin{aligned} \Big(\bigwedge_{(x,y)\in\mathcal{R}} v_x = v_y \Big) &\Rightarrow \Big(\bigwedge_{(m,n)\in\mathbf{M}[\mathcal{R}]} m(v) = n(v) \Big) && \text{(Lemma 1)} \\ &\Rightarrow \Big(\bigwedge_{(p,q)\in\mathbf{P}[\mathbf{M}[\mathcal{R}]]} p(v) = q(v) \Big) && \text{(Lemma 2)} \\ &\Rightarrow \Big(\bigwedge_{(x,y)\in\mathcal{R}} f_x(v) = f_y(v) \Big) && (\mathcal{R} \text{ is a BR}) \end{aligned}$$

This concludes the proof. □

Proof of Proposition 3. For an arbitrary $v \in \mathbb{R}^S$ the following chain of logical implications hold

$$\begin{aligned} \Big(\bigwedge_{(x,y)\in\mathcal{R}^*} v_x = v_y \Big) &\Rightarrow \Big(\bigwedge_{(x,y)\in\mathcal{R}} v_x = v_y \Big) && (\mathcal{R} \subseteq \mathcal{R}^*) \\ &\Rightarrow \Big(\bigwedge_{(x,y)\in\mathcal{R}} f_x(v) = f_y(v) \Big) && \text{(Proposition 2)} \\ &\Rightarrow \Big(\bigwedge_{(x,y)\in\mathcal{R}^*} f_x(v) = f_y(v) \Big) && (= \text{ is an equivalence relation}) \end{aligned}$$

This concludes the proof. □

Lemma 5. *Let $\mathcal{Q} \in \mathbf{M} \times \mathbf{M}$ be an equivalence relation. Then $(p,q) \in \mathbf{P}[\mathcal{Q}]$ iff $\sum_{m\in M} p(m) = \sum_{m\in M} q(m)$ for all $M \in \mathbf{M}/\mathcal{Q}$.*

Proof of Lemma 5. We have already seen in the proof of Lemma 4 that $(p, q) \in \mathbf{P}[\mathcal{Q}]$ iff the optimal value of the following linear program is equal to 0.

$$\begin{aligned} &\max_{\alpha} \textstyle\sum_{m\in\mathbf{M}}(p-q)(m)\cdot\alpha(m) \\ &\qquad \alpha(m)-\alpha(n) \leq 0 \qquad\qquad \forall (m,n)\in\mathcal{Q} \\ &\qquad \alpha(m)-\alpha(n) \leq 1 \qquad\qquad \forall (m,n)\notin\mathcal{Q} \end{aligned}$$

Note that for any $\alpha \in F[\mathcal{Q}]$ we have that if $(m, n) \in \mathcal{Q}$ we have $\alpha(m) = \alpha(n)$. Define for each $M \in \mathbf{M}/\mathcal{Q}$, $\alpha(M) = \alpha(m)$ for some $m \in M$. Then, collecting terms, we rewrite the above linear program as

$$\begin{aligned} &\max_{\alpha} \textstyle\sum_{M\in\mathbf{M}/\mathcal{Q}} \big(\sum_{m\in M}(p-q)(m)\big)\cdot\alpha(M) \\ &\qquad \alpha(M)-\alpha(N) \leq 1 \qquad\qquad\qquad \forall M\neq N\in\mathbf{M}/\mathcal{Q} \end{aligned}$$

From the above formulation it is clear that if $\sum_{m\in M} p(m) = \sum_{m\in M} q(m)$ for all $M \in \mathbf{M}/\mathcal{Q}$ then the optimal value of the above linear program is 0, which implies that $(p, q) \in \mathbf{P}[\mathcal{Q}]$. For the converse, suppose that $\sum_{m\in M} p(m) \neq \sum_{m\in M} q(m)$ for some $M \in \mathbf{M}/\mathcal{Q}$. Without loss of generality suppose that $\sum_{m\in M} p(m) > \sum_{m\in M} q(m)$. Then, we may take $\alpha(M) = 1$ and $\alpha(N) = 0$ for all other classes and obtain a positive lower bound for the above linear program implying that $(p, q) \notin \mathbf{P}[\mathcal{Q}]$. $\square$

Proof of Proposition 4. By hypothesis $\mathcal{R}$ is an equivalence relation. Let's fix a representative index $x_H \in H$ for each equivalence class $H \in S/_{\mathcal{R}}$. For $p \in \mathbf{P}$ we define the backward quotient of p w.r.t. $\mathcal{R}$ as $p^{\mathcal{R}} = \sum_{m\in\mathbf{M}} p(m)\, m^{\mathcal{R}}$, where for arbitrary $m \in \mathbf{M}$, $m^{\mathcal{R}} = \prod_{H\in S/_{\mathcal{R}}} x_H^{\sum_{x\in H} m(x)}$.

For an arbitrary $m \in \mathbf{M}$, one can verify that $\rho\colon S \times S \to \mathbb{R}_{\geq 0}$ defined by

$$\rho(x,y) = \begin{cases} m(x) & \text{if } x\in H \text{ and } y = x_H \text{ for some } H\in S/_{\mathcal{R}} \\ 0 & \text{otherwise} \end{cases}$$

is a monomial coupling witnessing $(m, m^{\mathcal{R}}) \in \mathbf{M}[\mathcal{R}]$. By Lemmas 3 we know that $\mathbf{M}[\mathcal{R}]$ is an equivalence relation, then by Lemma 5 we can conclude that $(p, p^{\mathcal{R}}) \in \mathbf{P}[\mathbf{M}[\mathcal{R}]]$ for any $p \in \mathbf{P}$.

By Lemma 4 also $\mathbf{P}[\mathbf{M}[\mathcal{R}]]$ is an equivalence relation. Pick any $(x, y) \in \mathcal{R}$. For Taylor's theorem, the implication of Definition 1 ensures that $f_x^{\mathcal{R}}(m) = f_y^{\mathcal{R}}(m)$ for all $m \in \mathbf{M}$, that is $f_x^{\mathcal{R}} = f_y^{\mathcal{R}}$. By reflexivity of $\mathbf{P}[\mathbf{M}[\mathcal{R}]]$ we have $(f_x^{\mathcal{R}}, f_y^{\mathcal{R}}) \in \mathbf{P}[\mathbf{M}[\mathcal{R}]]$. As said above, $(f_x, f_x^{\mathcal{R}}), (f_y, f_y^{\mathcal{R}}) \in \mathbf{P}[\mathbf{M}[\mathcal{R}]]$. Hence by symmetry and trasitivity of $\mathbf{P}[\mathbf{M}[\mathcal{R}]]$ we have $(f_x, f_y) \in \mathbf{P}[\mathbf{M}[\mathcal{R}]]$. $\square$