A CHARACTERISATION OF COMPATIBLE STATE SPACE AGGREGATIONS FOR DISCRETE DYNAMICAL SYSTEMS

CHRIS GOOD, DAVID PARKER, MATE PULJIZ, AND JONATHAN E. ROWE

Abstract. We study coarse grainings of discrete time dynamical systems on $\mathbb{R}^n$. In particular we are interested in state space (or dimensional) aggregations. We give a complete characterisation of compatible aggregations for analytic maps on $\mathbb{R}^n$. We show how our results apply to artificial chemistries, Random Heuristic Search models, and related finite population models. We also show that deciding whether a Markov chain has a non-trivial aggregation is NP-complete.

1. Introduction

Modelling real life dynamical systems is a challenging problem. Whether we are exploring social dynamics on a large scale, or trying to understand biological interactions in a cell at the molecular level, the number of variables involved to describe these processes is often too large for any meaningful computer simulation or analysis to take place in a reasonable amount of time. One then has to resort to various techniques of simplifying the underlying model, thus reducing its dimension and enabling computation.

One typical approach is that of coarse graining. Given a (continuous) map $T\colon \mathbb{R}^n \to \mathbb{R}^n$, the map $\Xi\colon \mathbb{R}^n \to \mathbb{R}^m$ is a *coarse graining* of T provided the following diagram commutes.

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{T} & \mathbb{R}^n \\ \downarrow{\scriptstyle \Xi} & & \downarrow{\scriptstyle \Xi} \\ \mathbb{R}^m & \xrightarrow{S} & \mathbb{R}^m \end{array}$$

The point here is that the dynamics are preserved by Ξ in the sense that $\Xi\big(T^k(x)\big) = S^k\big(\Xi(x)\big)$ for any k (in topological dynamics, such a Ξ is called a semi-conjugacy). Typically, of course, in practical applications one wants m to be very much smaller than n.

In this paper we look at a systematic approach for finding those coarse grainings which are a result of aggregating or lumping dimensions. Our approach stems from analysing state space aggregations in the Random Heuristic Search framework developed by Vose [20], but extends to provide a complete characterization of such coarse grainings for discrete time Markov processes and analytic maps on $\mathbb{R}^n$. The theory provides an algorithm for determining such coarse grainings. However, we also show that the problem of deciding whether a Markov chain has a non-trivial state space aggregation is NP-complete.

Consider the following problem from evolutionary biology. Suppose that for a set of n states, which we might call genotypes, G (for simplicity we frequently assume that $G = \{1, 2, \ldots, n\} = [n]$) we are given a set of rules that determines which

genotype is dominant in any given pair, i.e. the genotype inherited by the offspring. Equivalently, the rules specify an orientation for each edge of the complete graph over the set of n vertices G. We write $k \to i$ to mean that mating between k and i produces an offspring of type i. One would then like to predict how a population of individuals with these genotypes evolves over time.

The first problem one faces is how to model this evolution. One common approach is to assume that the population is large and well-mixed. This leads to a model whose dynamics is given by a quadratic polynomial $T\colon \mathbb{R}^n \to \mathbb{R}^n$, $T = (T_1, \ldots, T_n)$, where

$$T_i(p) = T_i(p_1, \ldots, p_n) = p_i \left(p_i + 2 \sum\nolimits_{k \to i} p_k \right), \text{ for any } 1 \leq i \leq n, \tag{1}$$

and where the sum runs over all the genotypes k which are dominated by i. The function T is the *heuristic* in the aforementioned Random Heuristic Search framework, but we shall often say that T is a *continuous*, or *infinite population model*. The interpretation is as follows: suppose that we start with a population in which the proportions of different genotypes are given by $p = (p_1, \ldots, p_n)$, then $T_i(p)$ gives the probability that the offspring of two individuals chosen at random has genotype i. In particular, note that the non-negative portion ($p \geq 0$) of the hyperplane $p_1 + \cdots + p_n = 1$ is invariant under T. We remark that this construction generalises *binary tournaments* given in [16, 17].

Another approach is to consider finite population models. Assume that our population is of fixed size r and that this total number does not change over time. Let $\mathbb{N}_0$ denote the set of non-negative integers and let X_r^n be the set of all vectors in $\mathbb{N}_0^n$ whose terms sum to r, so that $v = (v_1, \ldots, v_n) \in X_r^n$ represents a population of r individuals v_i of whom have genotype i, for each $i \leq n$. Clearly X_r^n has cardinality $C(n, r) := \binom{r+n-1}{r}$. We then define a discrete-time Markov chain with the state space X_r^n where the transition probabilities are given by

$$\mathbb{P}\left[v \to w\right] = \frac{r!}{w_1! \cdots w_n!} \left(T_1(v/r)\right)^{w_1} \cdots \left(T_n(v/r)\right)^{w_n} = \frac{r!}{w!} \left(T\left(v/r\right)\right)^{w}, \tag{2}$$

where T is as in (1). The equation above implies that, starting from a population $v \in X_r^n$, the transition probabilities are given by a multinomial random variable with the expectation $T\left(v/r\right)$ and the parameter r. In other words, if we start with the population v, in order to obtain the generation in the next time step we first compute $T\left(v/r\right)$, this represents the proportions of different labels in an urn from which we draw r of them independently with replacement to form the new generation w.

Alternatively, one can think of this Markov chain purely in algorithmic terms. Starting with a population represented by $v \in X_r^n$, one randomly chooses two individuals from it with replacement. They then produce an offspring which inherits the dominant genotype of the parents (if the parents have the same genotype or if the same parent is chosen twice, then the offspring inherits that same genotype). This 'mating' process is repeated r times and results in a new generation of r offspring represented by w. It was shown in [20] that this leads to the same transition probabilities as given in (2).

Once a particular model has been chosen, one might try to reduce the state space in order to facilitate computations. This clustering of portions of the state space should be done in such a way that the dynamics on these 'higher level' states is still well-defined. In the present article we shall be interested only in reductions

that work across the range of derived models we mentioned above. In Section 2 we argue that the only simplifications to be considered under this requirement are state space aggregations that correspond to the changes of variables of the form $P_i = p_{i_1} + \cdots + p_{i_k}$ where each of the variables p_j occurs exactly once in exactly one of P_i's. This effectively gives a partition of the set of variables. It was shown by Vose in [20] that for aggregations of this type it suffices to find those that are compatible with the continuous model as they naturally extend to aggregations for the derived models described above. We further show that they are also compatible with the dynamics of derived models from a larger class (Theorem 2.1) reinforcing the view that it is only aggregations of the continuous model that matter.

It turns out that for the map T as in (1) it is possible to characterise all the valid aggregations efficiently using the notion of *contiguous partitions*. We say that an equivalence relation (or the corresponding partition) $\equiv$ on the set G is *contiguous* with respect to an orientation $\to$ on G if for all $i, j, k \in G$ we have

$$i \equiv j \equiv k \quad \text{whenever} \quad i \equiv k \text{ and } k \to j \to i.$$

In words, any two genotypes i and k that are in the same block of the partition must for each individual j from any other block either both dominate or both be dominated by it. In Section 3 we prove

Theorem 3.1. *Let T be a heuristic as in (1). An equivalence relation on G is compatible (i.e. gives a coarse graining) with T if and only if it is contiguous with respect to $\to$.*

This can be seen as an extension of the well-known condition for lumping states of a Markov chain (see e.g. [9]) which says that any two states that are lumped together must have exactly the same outgoing transition probabilities towards any of the blocks in the partition. Theorem 3.1 itself can further be seen as a Corollary of Theorem 6.1 where, more generally, we give a criterion for finding compatible aggregations in terms of the Taylor coefficients of an analytic map.

Section 4 is about the computational complexity of our problem and there we show that even for Markov chains the problem of deciding whether a non-trivial aggregation of the system exists is NP-hard (Theorem 4.1). It is then perhaps somewhat surprising to learn that in that same section we give a polynomial algorithm deciding the very same question in the class of weighted binary tournaments. Theorem 3.1 above and its analogue, Theorem 3.2, play a decisive role in constructing this algorithm.

In Section 5 we show how the systems we are studying can be interpreted in the context of artificial chemistries. Indeed, the techniques developed here have been used to model the synchronization of kinetochore spindles across the cell during mitosis [14]. And lastly Section 6 deals with aggregations of analytic maps. This can be seen as a general model encompassing all those discussed in the sections preceding it. The contiguity criterion manifests itself there as a condition on certain Taylor series coefficients, see (8).

Coarse graining has been previously studied by Vose and his collaborators and some useful criteria have been devised that guarantee existence of coarse grainings in certain cases [1, 15–17]. Other authors, most prominently Rabitz and collaborators [10–12], and more recently Jacobi [6], and Tomlin et al. [19] explored the coarse grainings of continuous systems given by differential equations. It is not hard to

see that the theory of linear coarse grainings coincides in both the discrete and continuous systems.

We would like to stress that there are other approaches to simplifying the dynamics of iterated dynamical systems. These model reduction techniques include, most notably, singular value decomposition and variants thereof (see, for example, [5,7]). However, in general these do not preserve the inner dynamics of the system, which is our primary concern.

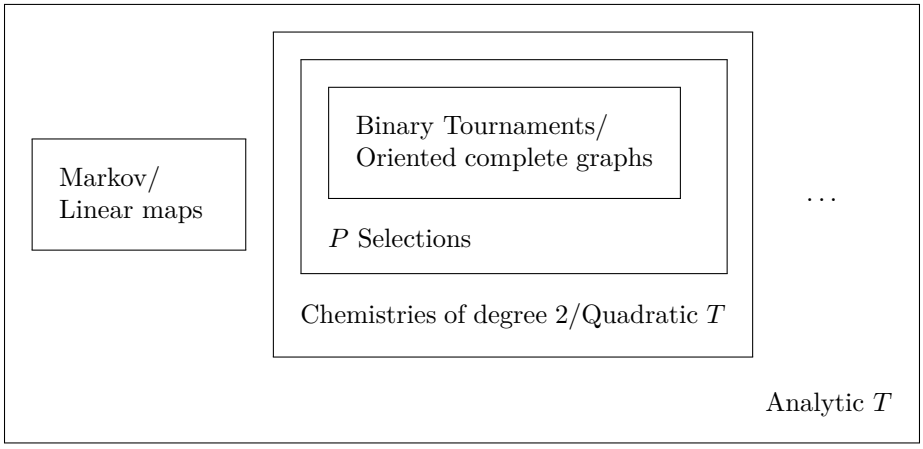

FIGURE 1. Maps for which we characterise aggregations

2. Different models

In this section we describe a general method for generating finite population Markov chain models akin to those we discussed in the introduction. Our starting point is an infinite model given by a map $T\colon \mathbb{R}^n \to \mathbb{R}^n$. We further assume that T maps the unit simplex $\Delta^n = \{p \in \mathbb{R}^n \mid \sum_{i=1}^n p_i = 1 \text{ and } p_i \geq 0\}$ into itself.

Let

$$\left\{ Z_\alpha^{k,i}, Y_\beta^{k,i} \mid k, i \in \mathbb{N} \text{ and } \alpha, \beta \in \Delta^n \right\}$$

be a family of independent random variables where

$$Z_\alpha^{k,i} \sim \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix} \quad \text{and} \quad Y_\beta^{k,i} \sim \begin{pmatrix} 1 & 2 & \dots & n \\ \beta_1 & \beta_2 & \dots & \beta_n \end{pmatrix}.$$

Let $r \in \mathbb{N}$, the population size, be fixed and recall that for our purposes $G = \{1, 2, \dots, n\} = [n]$. The family above acts as a stock of independent random variables we use to define a Markov chain $(S_k\colon k \in \mathbb{N})$ over the state space X_r^n. The evolution of this chain is prescribed by a transition function

$$F\colon X_r^n \times [n]^\infty \times [n]^\infty \to X_r^n,$$

and the rule

$$S_{k+1} = F(S_k; Z_{T(S_k/r)}^{k,1}, Z_{T(S_k/r)}^{k,2}, \dots; Y_{S_k/r}^{k,1}, Y_{S_k/r}^{k,2}, \dots).$$

The idea is that by varying F one can emulate the effect of changing the simulation algorithm by means of which the finite population model evolves. This will better be understood on an example.

If we let F be

$$F_1(v; i_1, i_2, \dots; j_1, j_2, \dots) = \sum_{l=1}^{r} e_{i_l},$$

where e_i is the i^{th} vector of the canonical basis in $\mathbb{R}^n$, then the chain $(S_k \colon k \in \mathbb{N})$ satisfies the following transition rule

$$S_{k+1} = \sum_{l=1}^{r} e_{Z^{k,l}_{T(S_k/r)}}.$$

In particular, this is precisely the same chain as the one described in the introduction. Each new generation S_{k+1} is formed by drawing r genes with replacement from an urn in which the distribution of different genes is given by $T(S_k/r)$.

Another useful choice is taking

$$F_2(v; i_1, \dots; j_1, \dots) = v + e_{i_1} - e_{j_1}.$$

In this case our chain $(S_k \colon k \in \mathbb{N})$ will satisfy

$$S_{k+1} = S_k + e_{Z^{k,1}_{T(S_k/r)}} - e_{Y^{k,1}_{S_k/r}}$$

and the next generation S_{k+1} is produced by throwing out one individual uniformly chosen from the current population (the third term accounts for this) in order to free up one space for another individual carrying a gene randomly chosen from $[n]$ with weights $T(S_k/r)$ (the second term). When T is quadratic as in (1) this evolution can, perhaps more naturally, be explained by saying that in each step a couple[1] is chosen that produces an offspring with the gene inherited from the dominant parent while at the same time, independently, one individual dies.

Clearly F_1 and F_2 represent two extreme approaches to modelling this system. The first is generational, as all the individuals get replaced at each step, while the other represents one change at a time evolution. There are various other possibilities in between which can be modelled by choosing a different F.

We now wish to show that any aggregation of genes that coarse grains the dynamics of T also works for the models induced by F_1 and F_2. Indeed in Theorem 2.1 below we give a sufficient condition on F for this to happen.

Recall that an aggregation is a partition of the set $G = [n]$. By choosing some ordering on the blocks of the partition we can identify this with a function $\pi \colon [n] \to [m]$ where $m \leq n$. We set Ξ_π to be an $m \times n$ matrix associated to this aggregation where

$$\Xi_\pi(i, j) = \begin{cases} 1, & \text{if } \pi(j) = i, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, matrix Ξ_π is the change of variable transformation corresponding to the aggregation π. Asking that π is compatible with a map $T \colon \mathbb{R}^n \to \mathbb{R}^n$ amounts to asking that there exists map $\tilde{T} \colon \mathbb{R}^m \to \mathbb{R}^m$ such that Ξ_π semi-conjugates T and $\tilde{T}$, i.e.

$$\Xi_\pi \circ T = \tilde{T} \circ \Xi_\pi.$$

Note that π aggregates $G = \{1, 2, \dots, n\}$ but the induced models we now wish to consider have X_r^n for the state space. There is, however, a natural partition on X_r^n that π induces and it is given by Ξ_π. In particular Ξ_π maps X_r^n to X_r^m.

[1] As mentioned before, it could happen that a couple is in fact the same parent chosen twice.

Theorem 2.1. *Let $F\colon X_r^n \times [n]^\infty \times [n]^\infty \to X_r^n$ be a modelling scheme which is compatible with any aggregation, i.e. such that for any $m \leq n$ and any $\pi\colon [n] \to [m]$ the map $\tilde{F}\colon X_r^m \times [m]^\infty \times [m]^\infty \to X_r^m$ is well-defined by the formula*

$$\tilde{F}(\Xi_\pi(v); \pi(i_1), \dots; \pi(j_1), \dots) = \Xi_\pi(F(v; i_1, \dots; j_1, \dots)).$$

Then given any system $T\colon \mathbb{R}^n \to \mathbb{R}^n$ all the aggregations compatible with T induce coarse grainings of the Markov model $(S_k\colon k \in \mathbb{N})$ obtained via F.

Proof. Let $\pi\colon [n] \to [m]$ be an aggregation compatible with T. Employing the usual criteria for matching outgoing probabilities, it suffices to see that for any three states $v, u, w \in X_r^n$ the equation

$$\mathbb{P}\left[\Xi_\pi(S_{k+1}) = \Xi_\pi(w) \mid S_k = v\right] = \mathbb{P}\left[\Xi_\pi(S_{k+1}) = \Xi_\pi(w) \mid S_k = u\right]$$

holds as soon as $\Xi_\pi(v) = \Xi_\pi(u)$. The LHS of this expression is

$$\begin{aligned}
&\mathbb{P}\left[\Xi_\pi(S_{k+1}) = \Xi_\pi(w) \mid S_k = v\right] = \\
&\qquad = \mathbb{P}\left[\Xi_\pi\left(F\left(S_k; Z^k_{T(S_k/r)}; Y^k_{S_k/r}\right)\right) = \Xi_\pi(w) \mid S_k = v\right] = \\
&\qquad = \mathbb{P}\left[\tilde{F}\left(\Xi_\pi(S_k); \pi\left(Z^k_{T(S_k/r)}\right); \pi\left(Y^k_{S_k/r}\right)\right) = \Xi_\pi(w) \mid S_k = v\right] = \\
&\qquad\qquad = \mathbb{P}\left[\tilde{F}\left(\Xi_\pi(v); \pi\left(Z^k_{T(v/r)}\right); \pi\left(Y^k_{v/r}\right)\right) = \Xi_\pi(w)\right] = \\
&\qquad\qquad\qquad = \mathbb{P}\left[\tilde{F}\left(\Xi_\pi(v); Z^k_{\Xi_\pi(T(v/r))}; Y^k_{\Xi_\pi(v)/r}\right) = \Xi_\pi(w)\right].
\end{aligned}$$

It remains to notice that the last expression above will not change if we substitute u instead of v. This is because $\Xi_\pi(v) = \Xi_\pi(u)$, and hence also by assumption $\Xi_\pi(T(v/r)) = \Xi_\pi(T(u/r))$. Now backtracking the same steps we get the RHS. □

It is easy to check that both F_1 and F_2, regardless of the population size r, satisfy the hypothesis of the theorem above. Thus, we have just shown that the aggregations that work for the heuristic T, also coarse grain two associated infinite families of Markov chains. This justifies our primary concern of coarse graining heuristics.

At the end of this section we provide a simple example showing that even linear coarse grainings that are not induced by an aggregation of variables need not lead to a coarse graining of finite population models.

Example 2.2. Consider a map $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by $T(p_1, p_2, p_3) = (p_1, p_3, p_2)$. The projection map $\Xi(p_1, p_3, p_2) = p_2 - p_3$ is clearly a coarse graining albeit not induced by an equivalence relation. The transition matrix[2] of the associated Markov chain constructed via F_2 with $r = 2$ over the state space $X_2^3 = \{(2,0,0), (0,2,0), (0,0,2), (1,1,0), (1,0,1), (0,1,1)\}$ is

$$\begin{bmatrix}
1 & \cdot & \cdot & 1/4 & 1/4 & \cdot \\
\cdot & \cdot & 1 & \cdot & 1/4 & 1/4 \\
\cdot & 1 & \cdot & 1/4 & \cdot & 1/4 \\
\cdot & \cdot & \cdot & \cdot & 1/2 & \cdot \\
\cdot & \cdot & \cdot & 1/2 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1/2
\end{bmatrix}.$$

[2]The dots stand for zeros.

This matrix is clearly not compatible with the induced aggregation that merges the states $(0,1,1)$ and $(2,0,0)$. ◁

3. Aggregating Weighted Binary Tournaments

In the previous section we argued why we are interested only in aggregations of the infinite model T. Here, we seek to characterise these aggregations for the case when T is modelling a Binary Tournament of the form (1).

We start by orienting the complete graph K_n over $G = [n]$ and define a map $T\colon \mathbb{R}^n \to \mathbb{R}^n$, $T = (T_1, \dots, T_n)$ as in (1) by setting

$$T_i(p) = T_i(p_1, \dots, p_n) = p_i \left(p_i + 2 \sum\nolimits_{k \to i} p_k \right), \text{ for any } 1 \leq i \leq n.$$

Recall that an equivalence relation $\equiv$ on G is *contiguous* with respect to the orientation $\to$ if for all $i, j, k \in G$ we have

$$i \equiv j \equiv k \quad \text{whenever} \quad i \equiv k \text{ and } k \to j \to i. \tag{3}$$

Note that i and k in the definition above are distinct as we cannot have $j \to i$ and at the same time $i \to j$. Also, in general, the relation $\to$ need not be transitive. An example of a non-transitive rule is the game 'Rock, paper, scissors' and extensions thereof. We can now state our characterisation of compatible aggregations for these systems. Its proof is omitted as it follows from more general Theorem 3.2 we prove below.

Theorem 3.1. *Let T be a heuristic as in (1). An equivalence relation on G is compatible (i.e. gives a coarse graining) with T if and only if it is contiguous with respect to $\to$.*

Suppose that it is not an orientation that is given on K_n but let $P\colon G^2 \to [0,1]$ be a function such that $P(i,j) = 1 - P(j,i)$ for all $i, j \in G$. We call such a P a *selection map*. The interpretation is that in a clash between i and j, gene i dominates with probability $P(i,j)$ and hence j dominates with the probability $1 - P(i,j) = P(j,i)$. Note that this forces $P(i,i) = 1/2$ for all $i \in G$.

If $P(i,j) \in \{0,1\}$ whenever $i \neq j$ then this reduces to the previous case as $P(i,j) = 1$ can be interpreted as orienting the edge connecting i and j by choosing $j \to i$. It turns out that there is a characterisation of admissible aggregations that is the same as the one given in Theorem 3.1. Of course, the definition of a contiguous partition requires some amendments.

An equivalence relation $\equiv$ on G is *contiguous* with respect to a selection map P if for all $i, j, k \in G$ we have

$$i \equiv j \equiv k \quad \text{whenever} \quad i \equiv k \text{ and } P(i,j) \neq P(k,j). \tag{4}$$

We remark that this reduces to (3) when P maps off-diagonal pairs to $\{0,1\}$.

If we denote by $\pi\colon [n] \to [m]$ the partition induced by the classes of $\equiv$ then the condition above amounts to requiring that for any two blocks $\pi^{-1}(s)$ and $\pi^{-1}(t)$, with $s \neq t$, and for any $i, k \in \pi^{-1}(s)$ and $j \in \pi^{-1}(t)$ we have $P(i,j) = P(k,j)$.

If we let l be any other element in $\pi^{-1}(t)$, the same condition now applied with the roles of $\pi^{-1}(s)$ and $\pi^{-1}(t)$ swapped gives $P(j,k) = P(l,k)$. But since P is a selection this implies $P(k,j) = P(k,l)$ and hence $P(i,j) = P(k,l)$.

This shows that for a contiguous partition there is a well defined selection map $P\colon [m]^2 \to [0,1]$ such that $P(s,t) + P(t,s) = 1$ for all $s, t \in [m]$, and for every

$i \in \pi^{-1}(s)$ and $j \in \pi^{-1}(t)$ we have $P(s,t) = P(i,j)$. Although these are two different maps, we shall keep the same notation.

How does the heuristic map look in this more general setting? Let us compute the probability that the winning gene in a random pairing is of type i.

$$T_i(p_1, \ldots, p_n) = p_i^2 + \sum_{k \in [n] \setminus \{i\}} 2p_i p_k P(i,k) = 2p_i \sum_{k=1}^{n} P(i,k) p_k, \tag{5}$$

where we noted that $P(i,i) = 1/2$. This again reduces to (1) if the range of P is $\{0,1\}$ on the off-diagonal pairs. We are now ready to prove our characterisation.

Theorem 3.2. *Let P be a selection map on G, and let the heuristic T be as in (5). An equivalence relation on G is compatible (i.e. gives a coarse graining) with T if and only if it is contiguous with respect to P.*

Proof. Let $\pi\colon [n] \to [m]$ be the map associated to the partition of $G = [n]$ induced by a contiguous equivalence relation, and let Ξ_π be the associated aggregation. Take $s \in [m]$ and calculate

$$\begin{aligned}((\Xi_\pi \circ T)(p))_s &= \sum_{\pi(i)=s} T_i(p) = \sum_{\pi(i)=s} 2p_i \left(\sum_{k=1}^{n} P(i,k) p_k \right) = \\ &= 2 \sum_{\pi(i)=s} p_i \left(\sum_{t=1}^{m} \sum_{\pi(k)=t} P(i,k) p_k \right) = 2 \sum_{\pi(i)=s} p_i \left(\sum_{t=1}^{m} P(s,t) \sum_{\pi(k)=t} p_k \right) \\ &= 2(\Xi_\pi(p))_s \sum_{t=1}^{m} P(s,t) (\Xi_\pi(p))_t = \tilde{T}_s(\Xi_\pi(p)) = (\tilde{T}_s \circ \Xi_\pi)(p).\end{aligned}$$

Thus $\Xi_\pi \circ T = \tilde{T} \circ \Xi_\pi$, where $\tilde{T} = (\tilde{T}_1, \ldots, \tilde{T}_m)$. This shows that Ξ is a coarse graining of the system and, moreover, we see that the coarse grained map $\tilde{T}$ is in the same form as the original one. It expresses the rule of transformation for meta-genes that are given by the blocks of the partition π.

Conversely, suppose that $\pi\colon [n] \to [m]$ is a partition of $[n]$ whose associated aggregation Ξ_π coarse grains T. We need to prove that the equivalence relation that it induces is contiguous with respect to P. To that end, take $s, t \in [m]$, $s \neq t$. Following the discussion after (4) above, it suffices to prove that for any $i, k \in \pi^{-1}(s)$ and $j \in \pi^{-1}(t)$ we have $P(i,j) = P(k,j)$. For the sake of getting a contradiction assume that i, k and j are chosen such that $P(i,j) < P(k,j)$.

To simplify the notation we can, without loss of generality, assume that $i = 1$, $k = 2$, $j = 3$. Take vectors $v = (1/2, 0, 1/2, 0, \ldots, 0)$ and $w = (0, 1/2, 1/2, 0, \ldots, 0)$ and note that

$$\Xi_\pi(v) = \Xi_\pi(w)$$

and since Ξ_π is a coarse graining we must also have

$$(\Xi_\pi \circ T)(v) = (\Xi_\pi \circ T)(w)$$

and hence also

$$((\Xi_\pi \circ T)(v))_{\pi(1)} = \sum_{\pi(k)=\pi(1)} T_k(v) = T_1(v) = \frac{1}{2}(P(1,1) + P(1,3))$$

is equal to

$$((\Xi_\pi \circ T)(w))_{\pi(1)} = \sum_{\pi(k)=\pi(1)} T_k(w) = T_2(w) = \frac{1}{2}(P(2,2) + P(2,3)).$$

As $P(1,1) = P(2,2) = 1/2$, we get $P(1,3) = P(2,3)$, which contradicts the initial assumption $P(1,3) < P(2,3)$. This finishes the proof of the other implication. □

4. Intermezzo: On complexity

In this section we present a result showing that even for linear maps, i.e. Markov Chains, finding aggregations is a difficult task in general. More precisely we show

Theorem 4.1. *Deciding if there exists a non-trivial aggregation for a Markov chain is* NP*-complete.*

Recall that a Markov chain defined by a stochastic $n \times n$ matrix M possesses a non-trivial aggregation if there exist a (non-trivial) surjective partitioning function $\pi\colon [n] \to [m]$ where $1 < m < n$, and an $m \times m$ matrix R such that

$$\Xi M = R\Xi$$

where $\Xi = \Xi_\pi$ is the change of variable transformation associated to π.

In order for this decision problem to be well defined we clearly need to restrict the entries of transition matrices to be in a countable domain, say in $\mathbb{Q}$. Note however that by multiplying each entry of the matrix M by a sufficiently large number we can obtain a matrix with integer entries and constant column sums which clearly has a non-trivially aggregation if and only if M does.

Further note that this problem is indeed in NP as given M, Ξ, and R verifying if this is a solution amounts to multiplying and comparing matrices. It will thus suffice to show that some well-know NP-complete problem, say the subset sum problem (SSP), is *polynomially reducible* to the problem of deciding if an integer matrix with constant column sums has a non-trivial aggregation. For more details on these notions we refer the reader to any textbook dealing with complexity, e.g. [18].

Recall that SSP (originally KNAPSACK problem in Karp's list [8]) is an NP-complete decision problem that given a set of positive integers $\{a_1, \dots, a_n\}$ and an integer $0 < K < \sum_{i=1}^n a_i$ asks if there exists a subset $I \subseteq [n]$ such that $\sum_{i\in I} a_i = K$.

Proof of Theorem 4.1. Let an input for SSP $(\{a_1, \dots, a_n\}, K)$ be given. Set $L = \sum_{i=1}^n a_i - K$ and let M be an $(n+2) \times (n+2)$ matrix as below

$$M = \begin{bmatrix} a_1 & 2a_1 & 3a_1 & \dots & (n+2)a_1 \\ a_2 & 2a_2 & 3a_1 & \dots & (n+2)a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & 2a_n & 3a_1 & \dots & (n+2)a_n \\ (n+2)K & (n+1)K & na_1 & \dots & K \\ (n+2)L & (n+1)L & na_1 & \dots & L \end{bmatrix}.$$

Note that $0 < L < \sum_{i=1}^n a_i$. We claim that M has a non-trivial aggregation if and only if the given SSP instance has a solution. This will be enough to finish the proof as this reduction is clearly polynomial in the size of the input.

First assume that $I \subseteq [n]$ solves the SSP, i.e. $\sum_{i\in I} a_i = K$. Let I^c denote the complement of I in $[n]$. Thus $\sum_{i\in I^c} a_i = L$.

Define $\pi\colon [n+2] \to [2]$ by

$$\pi(i) = \begin{cases} 1, & \text{if } i \in I \text{ or } i = n+1, \\ 2, & \text{otherwise,} \end{cases}$$

and let $\Xi = \Xi_\pi$ be the associated change of variables transformation. One readily checks that

$$\Xi M = R\Xi,$$

where

$$R = \begin{bmatrix} (n+3)K & (n+3)K \\ (n+3)L & (n+3)L \end{bmatrix}$$

and thus M has a non-trivial aggregation.

Conversely, assume that $\pi\colon [n+2] \to [m]$ is a non-trivial aggregation compatible with M, where $1 < m < n+2$. Let $\Xi = \Xi_\pi$ and R be such that

$$\Xi M = R\Xi \tag{6}$$

holds. By the pigeon-hole principle and because of non-triviality there must exist $k \in [m]$ such that $1 < |\pi^{-1}(k)| < n+2$. Choose two different integers $s, t \in \pi^{-1}(k)$. Inspecting the elements on positions (k, s) and (k, t) in matrices ΞM and $R\Xi$ and using (6) we conclude that they are the same and they equal

$$s\sum_{i\in I} a_i + (n+3-s)(\kappa K + \lambda L) = t\sum_{i\in I} a_i + (n+3-t)(\kappa K + \lambda L) = R_{(k,k)},$$

where $I = \pi^{-1}(k) \cap [n]$ and κ and λ are 0-1 indicators depending on whether $(n+1)$ and $(n+2)$ respectively are in $\pi^{-1}(k)$. From here we get

$$(s-t)\sum_{i\in I} a_i = (s-t)(\kappa K + \lambda L)$$

and as $s \neq t$

$$\sum_{i\in I} a_i = \kappa K + \lambda L.$$

The only possibilities are now:

- $\kappa = \lambda = 1$ and $I = [n]$,
- $\kappa = 1$, $\lambda = 0$ and $I \subsetneq [n]$,
- $\kappa = 0$, $\lambda = 1$ and $I \subsetneq [n]$,
- $\kappa = \lambda = 0$ and $I = \emptyset$.

The first can be discarded as it would imply that $\pi^{-1}(k) = [n+2]$ and the aggregation given by π is trivial. Similarly, the last is impossible as it would imply that $\pi^{-1}(k) = \emptyset$ contradictory to our choice of $k \in [m]$. If the second holds then I is the subset solving our SSP $\sum_{i\in I} a_i = K$, and in case the third one holds, the complement I^c of I in $[n]$ solves the SSP $\sum_{i\in I^c} a_i = K$, as we know that $\sum_{i\in I} a_i = L$. This completes the proof of the theorem. □

To aid the understanding, we illustrate the proof above with an example.

Example 4.2. Take two instances of SSP $(\{5,6,1,8\},12)$ and $(\{9,4,12,1\},15)$. The construction from the proof gives matrices

$$M_1 = \begin{bmatrix} 5 & 10 & 15 & 20 & 25 & 30 \\ 6 & 12 & 18 & 24 & 30 & 36 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 8 & 16 & 24 & 32 & 40 & 48 \\ 72 & 60 & 48 & 36 & 24 & 12 \\ 48 & 40 & 32 & 24 & 16 & 8 \end{bmatrix}, \qquad M_2 = \begin{bmatrix} 9 & 18 & 27 & 36 & 45 & 54 \\ 4 & 8 & 12 & 16 & 20 & 24 \\ 12 & 24 & 36 & 48 & 60 & 72 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 90 & 75 & 60 & 45 & 30 & 15 \\ 66 & 55 & 44 & 33 & 22 & 11 \end{bmatrix}.$$

The first system has a valid aggregation $\{\{1,2,3,5\},\{4,6\}\}$ and one can check that $\Xi_1 M_1 = R_1 \Xi_1$ where

$$\Xi_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \quad R_1 = \begin{bmatrix} 84 & 84 \\ 56 & 56 \end{bmatrix}.$$

This agrees with the fact that the first instance of SSP has a solution $a_1 + a_2 + a_3 = 5 + 6 + 1 = 12$. Here $n = 4$ and the set of indices giving the solution is $I = \{1,2,3\} = \{1,2,3,5\} \cap [4]$.

The other instance of SSP does not have a solution and the consequence is, as we have proved, that the Markov chain defined via transition probabilities in the normalised matrix $\frac{1}{182} M_2$ has no non-trivial aggregations. ◁

It is worth noting that the problem of finding a non-trivial aggregation can be efficiently solved for certain classes of systems. For example, weighted binary tournaments discussed in Section 3 form one such a class. The polynomial algorithm that checks for non-trivial aggregations in that case is given below.

The algorithm relies heavily on the fact that we have an efficient way of checking whether a partition gives a compatible aggregations by means of the contiguity test (4), see also Theorem 3.2. In particular this means that given any non-trivial compatible partition $\{S_1, \ldots, S_k\}$ of $[n]$, and assuming that $|S_1| > 1$, the refined partition $\{S_1\} \cup \{\{i\} \mid i \in [n] \setminus S_1\}$ is also a non-trivial aggregation for the same weighted binary tournament. This reduces the number of aggregations we need to check greatly, but still leaves exponentially many of them to be considered. The final trick making this work in polynomial time is that by virtue of (4) at each step we can either verify that the current lump S is a block of a compatible aggregation, or we can increase it by at least one element $S \longleftarrow S \cup Q$ and be certain we are not omitting any solutions by doing so.

Let us explain in plain words what this algorithm does. It starts by considering all possible pairs $s, t \in [n]$, $s \neq t$ and then attempts to prove that there is a compatible aggregation lumping those two variables together. S is the current candidate for a block of a compatible aggregation. Using criterion (4) the algorithm picks all the elements $Q \subseteq [n] \setminus S$ that falsify the contiguity property. If $Q = \emptyset$ then we have a certificate of S being a lump of a non-trivial compatible aggregation, otherwise $S \cup Q$ is taken to be a new candidate. If in the end this results with S being everything (S=$[n]$) we conclude that no non-trivial aggregation lumps s and t together. Should this be the case for all the pairs s and t, we have a proof that no non-trivial aggregations exist.

This discussion thus proves

Theorem 4.3. *Determining whether a non-trivial aggregation for a weighted binary tournament exists is in* P.

input : Rational $n \times n$ matrix P satisfying $\mathsf{P}_{ij} + \mathsf{P}_{ji} = 1$ and $\mathsf{P}_{ii} = 1/2$
output: Yes if the tournament given by P has a non-trivial aggregation, No otherwise

```
foreach s, t ∈ [n], s ≠ t do
    S ⟵ {s, t}
    repeat
        Q ⟵ ∅
        foreach j ∈ [n] \ S do
            j_is_not_compatible ⟵ No
            foreach i, k ∈ S, i ≠ k do
                if P_ij ≠ P_kj then
                    j_is_not_compatible ⟵ Yes
                end
            end
            if j_is_not_compatible then
                Q ⟵ Q ∪ {j}
            end
        end
        S ⟵ S ∪ Q
    until Q = ∅
    if S ≠ [n] then
        return Yes
    end
end
return No
```

5. Artificial Chemistries

We shall now show that the *selection map* model considered in Section 3 can further be generalised to include even larger class of quadratic maps. A natural way to interpret these is through artificial chemistries.

Example 5.1 (Chemical reactions of degree 2)**.** Let each number in $[n]$ represent a different chemical. Assume that for each choice of two chemicals c_1 and c_2 (not necessarily different) we are given a distribution $\tau_v = (\tau_{1,v}, \tau_{2,v}, \ldots, \tau_{n,v})$ over $[n]$ whose entries are interpreted as the proportions of each of the chemicals produced by a chemical reaction involving the particles c_1 and c_2 as reactants. The vector $v \in X_2^n$ is, as before, used to represent the chosen pair by setting $v = e_{c_1} + e_{c_2}$. We can also write this as a set of $\binom{n+1}{2}$ equations of the form

$$c_1 + c_2 \to \tau_{1,v} \cdot 1 + \tau_{2,v} \cdot 2 + \cdots + \tau_{n,v} \cdot n.$$

The evolution of such a system (assuming the chemical solution contains a large number of particles and is well-mixed) is given by $T = (T_1, \ldots, T_n) \colon \mathbb{R}^n \to \mathbb{R}^n$ where

$$T_i(p) = \sum_{v \in X_2^n} \frac{2}{v_1! \cdots v_n!} \tau_{i,v} p_1^{v_1} \cdots p_n^{v_n} = \sum_{v \in X_2^n} \frac{2}{v!} \tau_{i,v} p^v.$$

The factor $2/v!$ accounts for the number of different ways to choose the reactants represented by v. Here it is either 2, if the reactants are different, or 1, if they are not.

One can show that for an aggregation Ξ to be compatible with this dynamics $\Xi(\tau_{1,v},\dots,\tau_{n,v}) = \Xi(\tau_{1,w},\dots,\tau_{n,w})$ must hold true whenever $\Xi(v) = \Xi(w)$. In other words, taking the blocks of the partition associated with Ξ as meta-chemicals, the distributions over these must be well-defined for all the pairings of meta-chemicals as reactants. This is a direct consequence of Theorem 6.1 we prove later on. ◁

We could interpret any system given by a selection map as in (5), as a chemistry of degree 2. But the chemistries are still more general since they also incorporate systems as in the following example.

Example 5.2. Consider a second order chemistry on $\mathbb{Z}_n = \{0,1,2,\dots n-1\}$ with the reaction rules

$$i + j \to 1 \cdot (i + j \bmod n), \qquad \text{for all } i,j \in \mathbb{Z}_n.$$

Following Example 5.1, a partition $\pi\colon \mathbb{Z}_n \to M$ will be a valid aggregation if and only for any two of its blocks, it is well defined which block they produce. Let us fix one of the blocks $\pi^{-1}(m)$ for $m \in M$. Then for any element $d \in \mathbb{Z}_n$ the set $(\pi^{-1}(m) + d \bmod n)$ must be contained within one block. From here it is not hard to see that all the blocks are translates (or cosets) of the set $\{0, l, 2l, \dots, (n-l)\}$, where l is some positive divisor of n. Thus the range of π in M can be given a group structure that makes it isomorphic to $\mathbb{Z}_l$.

This establishes a correspondence between the compatible aggregations and the divisors of n. In fact, both the compatible aggregations and the set of divisors of n come equipped with a natural partial order making them into lattices. It is not hard to see that these lattices are isomorphic and the refinement relation in the former corresponds to the relation 'is divisible by' in the latter. ◁

6. Aggregating analytic maps

In this section we give a general framework incorporating, amongst others, models from Sections 3 and 5. Our results in this section apply to maps that are given by an absolutely convergent series $T = (T_1,\dots,T_n)\colon \mathbb{R}^n \to \mathbb{R}^n$

$$T_i(p) = \sum_{v_1,\dots,v_n \in \mathbb{N}_0} \frac{\alpha_{i,v}}{v_1!\cdots v_n!} p_1^{v_1}\cdots p_n^{v_n} = \sum_{v \in \mathbb{N}_0^n} \frac{\alpha_{i,v}}{v!} p^v \tag{7}$$

where $\alpha_{i,v} \in \mathbb{R}$ for $1 \le i \le n$, $v \in \mathbb{N}_0^n$ is a family of parameters. Our goal is to derive a sufficient and necessary condition for an aggregations $\pi\colon [n] \to [m]$ to be compatible with T solely in terms of these parameters. Indeed, we prove

Theorem 6.1. *Let $T = (T_1,\dots,T_n)\colon \mathbb{R}^n \to \mathbb{R}^n$ be an absolutely convergent series as in (7) defining an analytic function on $\mathbb{R}^n$. An aggregation $\Xi = \Xi_\pi\colon \mathbb{R}^n \to \mathbb{R}^m$ (associated to $\pi\colon [n] \to [m]$) is a valid coarse graining if and only if*

$$\Xi(\alpha_{1,v},\dots,\alpha_{n,v}) = \Xi(\alpha_{1,w},\dots,\alpha_{n,w}) \text{ whenever } v,w \in \mathbb{N}_0^n \text{ with } \Xi(v) = \Xi(w). \tag{8}$$

In particular, for all $u \in \mathbb{N}_0^m$ the coefficients $(\beta_{1,u},\dots,\beta_{m,u}) = \Xi(\alpha_{1,v},\dots,\alpha_{n,v})$, where $v \in \mathbb{N}_0^n$ is chosen such that $\Xi(v) = u$, are well-defined and the coarse grained system is a convergent series on $\mathbb{R}^m$ given by $Q = (Q_1,\dots,Q_m)\colon \mathbb{R}^m \to \mathbb{R}^m$ where

$$Q_i(q) = \sum_{u_1,\dots,u_m \in \mathbb{N}_0} \frac{\beta_{i,u}}{u_1!\cdots u_m!} q_1^{u_1}\cdots q_m^{u_m} = \sum_{u \in \mathbb{N}_0^m} \frac{\beta_{i,u}}{u!} q^u.$$

Note the similarity between condition (8) above and (4). Indeed, it is not hard to see that the former generalises the latter when T is a heuristic induced by a two point selection, and hence the theorem above generalises Theorem 3.2.

Our proof of this result is elementary and relies on the well-known multi-variable multinomial theorem and the fact that aggregations compatible with T must also be compatible with any of the partial derivatives of T. We also use a criterion for coarse grainings from [17] specialised for the case of linear maps, Proposition 6.3 below.

Recall that, as before $p = (p_1, \dots, p_n) \in \mathbb{R}^n$, and e_i is i-th vector of the canonical basis in $\mathbb{R}^n$. Variables v, w, etc. are reserved for labelling vectors in $\mathbb{N}_0^n$ and $|v| := v_1 + v_2 + \cdots + v_n$. Following the multi index notation due to Schwartz, the factorial function naturally extends over integral vectors via $v! = v_1! \cdots v_n!$ and so do integral powers $p^v := p_1^{v_1} p_2^{v_2} \dots p_n^{v_n}$. It is also convenient to introduce a shorthand $C(n, r) := \binom{r+n-1}{r}$, and for the space of (real) $m \times n$ matrices $M_{m\times n}$. We also use vectors to denote partial derivatives as in $\partial_v T(p) = \frac{d^{|v|}}{dp_1^{v_1} \dots dp_n^{v_n}} T(p_1, \dots, p_n)$.

Multinomial theorem is extremely easy to state using the introduced notation. The proof is classical and is left to the reader.

Proposition 6.2 (The vector multinomial theorem)**.** *Let $\Xi \in M_{m\times n}$ be a stochastic 0-1 matrix ($m \leq n$) and $\pi\colon [n] \to [m]$ the associated aggregation, and let $u \in \mathbb{N}_0^m$. For any $p \in \mathbb{R}^n$ we have*

$$(\Xi p)^u = \prod_{j=1}^{m} \left(\sum_{\pi(i)=j} p_i \right)^{u_j} = \sum_{y \in \mathbb{N}_0^n \cap \Xi^{-1}(u)} \frac{u!}{y!} p^y.$$

Proposition 6.3 ([17, Theorem 1])**.** *Let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable map and let $\Xi\colon \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, be a linear transformation. Then Ξ is a coarse graining of the system T if and only if the kernel of Ξ is invariant under the differential of T at any point of the domain $\mathbb{R}^n$, formally*

$$(DT)_x \cdot \ker \Xi \subseteq \ker \Xi \quad \text{for all } x \in \mathbb{R}^n. \tag{9}$$

Lemma 6.4. *Let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be a smooth function of class C^d and let $\Xi\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. If Ξ is a coarse graining of the system T then it coarse grains $\partial_v T\colon \mathbb{R}^n \to \mathbb{R}^n$ for any $v \in \mathbb{N}_0^n$ for which $|v| \leq d$.*

Proof. It is enough to prove that $\partial_i T = \partial_{e_i} T = \frac{d}{dp_i} T$ is coarse grained via Ξ for any $1 \leq i \leq n$ as the general result will then follow by induction. We denote the coarse grained system by $\tilde{T}$ and calculate

$$\Xi \frac{d}{dp_i} T(p) = \Xi \lim_{h\to 0} \frac{1}{h} \left(T(p + he_i) - T(p)\right) = \lim_{h\to 0} \frac{1}{h} \left(\Xi T(p + he_i) - \Xi T(p)\right) =$$
$$= \lim_{h\to 0} \frac{1}{h} \left(\tilde{T}(\Xi p + h\Xi e_i) - \tilde{T}(\Xi p)\right) = (D\tilde{T})_{\Xi p} \Xi e_i$$

which is a function of Ξp as can be seen from the formula. □

Note that the linearity of both the derivative and Ξ is crucial here as it has allowed us to swap the operators ∂_{e_i} and Ξ. We also note in passing that the calculation above proves that a surjective linear coarse graining of a smooth map produces a system with the same degree of smoothness. In fact, as Tóth et al. prove

in [19, Theorem 2.1], the Lipschitz property is also preserved in the coarse grained system, and with the same order. We now proceed to prove the main result.

Proof of Theorem 6.1. To further simplify notation we write α_v for the vector of parameters $(\alpha_{1,v}, \dots, \alpha_{n,v})$ and similarly for α_w. This allows us to full exploit the vector notation and write (7) as

$$T(p) = \sum_{v \in \mathbb{N}_0^n} \frac{\alpha_v}{v!} p^v.$$

Assume first that $\Xi(v) = \Xi(w)$ implies $\Xi(\alpha_v) = \Xi(\alpha_w)$. We need to show that $\Xi(T(p))$ is a function of $\Xi(p)$. We calculate

$$\begin{aligned}\Xi(T(p)) = \Xi\left(\sum_{v \in \mathbb{N}_0^n} \frac{1}{v!}\alpha_v p^v\right) = \sum_{v \in \mathbb{N}_0^n} \frac{1}{v!}\Xi(\alpha_v) p^v = \sum_{u \in \mathbb{N}_0^m} \sum_{v : \Xi(v) = u} \frac{1}{v!}\Xi(\alpha_v) p^v = \\ = \sum_{u \in \mathbb{N}_0^m} \frac{1}{u!}\beta_u \sum_{v\colon\, \Xi(v) = u} \frac{u!}{v!} p^v = \sum_{u \in \mathbb{N}_0^m} \frac{1}{u!}\beta_u (\Xi p)^u = Q(\Xi p)\end{aligned}$$

where we used the multinomial theorem.

Conversely, assume that the aggregation Ξ is a coarse graining of T. It suffices to show that $\Xi(v) = \Xi(w)$ implies $\Xi(\alpha_v) = \Xi(\alpha_w)$ when $v - w = e_i - e_j$, i.e. when the vectors v and w differ by 1 at exactly two coordinates. For any other choice of an equivalent pair v and w we can construct a chain of intermediate equivalent vectors transforming v to w by changing one coordinate at the time. Applying the base case to each link in the chain will produce a chain of equalities giving the general case. For example, if $n = 4$ and the aggregation corresponds to the partition $\{\{1, 2, 3\}, \{4\}\}$ then

$$(3, 0, 0, 1) \rightsquigarrow (2, 1, 0, 1) \rightsquigarrow (1, 2, 0, 1) \rightsquigarrow (0, 2, 1, 1)$$

is a valid sequence of intermediate transformations.

Assume, therefore, that for some $y, v, w \in \mathbb{N}_0^n$ we have $y = v - e_i = w - e_j$, and $\Xi(v) = \Xi(w)$. Using Lemma 6.4 we get that Ξ coarse grains $\partial_y T$. Note that $e_i - e_j = v - w \in \ker \Xi$, and the necessary condition for coarse graining (9) implies that

$$\begin{aligned}(D(\partial_y T))_{(0,\dots,0)}(e_i - e_j) = \partial_{e_i}\partial_y T(0, \dots, 0) - \partial_{e_j}\partial_y T(0, \dots, 0) = \\ = \partial_v T(0, \dots, 0) - \partial_w T(0, \dots, 0) \in \ker \Xi.\end{aligned}$$

Hence,

$$\Xi(\partial_v T(0, \dots, 0)) = \Xi(\partial_w T(0, \dots, 0)).$$

But the series for T is given as a Taylor series about $(0, \dots, 0)$ and therefore $\partial_v T(0, \dots, 0) = \alpha_v$ for all $v \in \mathbb{N}_0^n$. Thus $\Xi(\alpha_v) = \Xi(\alpha_w)$, as required. □

Observation. Note that the proof works even if the series expansion for T is given about a point other than 0. This allows to extend the result to any analytic function T just by requiring that the condition (8) holds locally for a series expansion of T about any point in the domain.

Corollary 6.5. *Let $T = (T_1, \ldots, T_n) \colon \mathbb{R}^n \to \mathbb{R}^n$ be an analytic function as in (7). An aggregation $\Xi \colon \mathbb{R}^n \to \mathbb{R}^m$ is a valid coarse graining if and only if it coarse grains each of the homogeneous components of T which are defined for each $d \in \mathbb{N}_0$ as the projection $T^{(d)} = (T_1^{(d)}, \ldots, T_n^{(d)}) \colon \mathbb{R}^n \to \mathbb{R}^n$ on the space of homogeneous polynomials of degree d where*

$$T_i^{(d)}(p) = \sum_{\substack{v_1,\ldots,v_n \in \mathbb{N}_0 \\ v_1+\cdots+v_n=d}} \frac{\alpha_{i,v}}{v_1! \cdots v_n!} p_1^{v_1} \cdots p_n^{v_n} = \sum_{v \in \mathbb{N}_0^n : |v|=d} \frac{\alpha_{i,v}}{v!} p^v.$$

Proof. It suffices to note that $|v| = |\Xi(v)|$ and similarly for w, and hence $|v| = |w|$ whenever $\Xi(v) = \Xi(w)$. Applying Theorem 6.1 twice will now give the result. □

7. Concluding remarks

In the present paper we gave a criterion for an aggregation of variables to be compatible with a map that is given by its series expansion; and showed how it can be applied to artificial chemistries. Despite the attractive and deceptively simple looking characterisations presented here, we also showed that the existence of a non-trivial compatible aggregation is a difficult problem in general. This means that for some systems, an algorithm searching for compatible aggregations cannot essentially be better then the one that exhaustively tests all the partitions of the set of variables.

For certain classes of problems, as we saw in the case of binary tournaments, it is possible to give an efficient algorithm searching for aggregations. The ideas we used here are somewhat similar in flavour to those used in other model reduction algorithms for DTMCs, one of such being 'partition minimisation' introduced by Paige and Tarjan in [13], and adapted to Markov chains in [2].

We also looked into different ways of modelling finite number particles systems. Given an infinite population model, one possible way of simulating the system was using Random Heuristic Search framework. The main characteristic of the RHS approach is that it is generational. Each new generation, although dependent on the previous one, consists of an entirely new collection of entities. Sometimes, other approaches may be more appropriate. Indeed, if our algorithm is supposed to simulate a real biological system, it is natural to allow the members of different generations to coexist. Similarly, if we are to simulate a chemical process, we would like to keep most of the particles for the new generation, and only those few that bump into another particle with which they can react, will be replaced with a product of the reaction in the next time step.

Artificial chemistries are a fruitful area and it is no surprise that different authors (Fontana et al. [4], Tóth et al. [19], Dittrich et al. [3] to name a few) model chemical reactions in different ways. What seems promising is that all of them incorporate a polynomial model similar to the one we had above. This should mean that the results on compatible aggregations can be easily transferred across the modelling paradigms, and this is something we would like to explore in future.

Acknowledgements

The authors acknowledge support from the European Union through funding under FP7-ICT-2011-8 project HIERATIC (316705).

References

[1] Keki Burjorjee, *Sufficient conditions for coarse-graining evolutionary dynamics*, Foundations of genetic algorithms, 2007, pp. 35–53.

[2] S. Derisavi, H. Hermanns, and W. Sanders, *Optimal state-space lumping in Markov chains*, Information Processing Letters **87** (2003September), no. 6, 309–315.

[3] Peter Dittrich and Pietro Speroni Di Fenizio, *Chemical organisation theory*, Bulletin of mathematical biology **69** (2007), no. 4, 1199–1231.

[4] Walter Fontana and Leo W Buss, *The arrival of the fittest: Toward a theory of biological organization*, Bulletin of Mathematical Biology **56** (1994), no. 1, 1–64.

[5] Aleksandr Nikolaevich Gorban, Nikolas K Kazantzis, Ioannis G Kevrekidis, Hans Christian Öttinger, and Constantinos Theodoropoulos, *Model reduction and coarse-graining approaches for multiscale phenomena*, Springer, 2006.

[6] Martin N. Jacobi, *Hierarchical organization in smooth dynamical systems*, Artif. Life **11** (December 2005), no. 4, 493–512.

[7] Martin Nilsson Jacobi and Olof Görnerup, *A spectral method for aggregating variables in linear dynamical systems with application to cellular automata renormalization*, Advances in Complex Systems **12** (2009), no. 02, 131–155.

[8] Richard M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972), 1972, pp. 85–103. MR0378476

[9] John G. Kemeny and J. Laurie Snell, *Finite Markov chains*, Springer-Verlag, New York-Heidelberg, 1976. Reprinting of the 1960 original, Undergraduate Texts in Mathematics. MR0410929 (53 #14670)

[10] Genyuan Li and Herschel Rabitz, *A general analysis of exact lumping in chemical kinetics*, Chemical engineering science **44** (1989), no. 6, 1413–1430.

[11] ———, *A general analysis of approximate lumping in chemical kinetics*, Chemical engineering science **45** (1990), no. 4, 977–1002.

[12] Genyuan Li, Herschel Rabitz, and János Tóth, *A general analysis of exact nonlinear lumping in chemical kinetics*, Chemical Engineering Science **49** (1994), no. 3, 343–361.

[13] Robert Paige and Robert E. Tarjan, *Three partition refinement algorithms*, SIAM J. Comput. **16** (1987), no. 6, 973–989. MR917035 (89h:68069)

[14] Richard Richard Henze, Chunyan Mu, Mate Puljiz, Nishanthan Kamaleson, Jan Huwald, John Haslegrave, Bashar Ibrahim, David Parker, Chris Good, Jonathan E. Rowe, and Peter Dittrich, *Multi-scale stochastic organization-oriented coarse-graining exemplified on the human mitotic checkpoint*. preprint.

[15] Jonathan E. Rowe, Michael D. Vose, and Alden H. Wright, *Coarse graining selection and mutation*, Foundations of genetic algorithms, 2005, pp. 176–191. MR2238251

[16] ———, *State aggregation and population dynamics in linear systems*, Artif. Life **11** (December 2005), no. 4, 473–492.

[17] ———, *Differentiable coarse graining*, Theoret. Comput. Sci. **361** (2006), no. 1, 111–129. MR2254227 (2007h:68171)

[18] John Talbot and Dominic Welsh, *Complexity and cryptography*, Cambridge University Press, Cambridge, 2006. An introduction. MR2221458

[19] Alison S Tomlin, Genyuan Li, Herschel Rabitz, and János Tóth, *The effect of lumping and expanding on kinetic differential equations*, SIAM Journal on Applied Mathematics **57** (1997), no. 6, 1531–1556.

[20] Michael D. Vose, *The simple genetic algorithm*, Complex Adaptive Systems, MIT Press, Cambridge, MA, 1999. Foundations and theory, A Bradford Book. MR1713436 (2000h:65024)