Local Irreducibility of Tail-Biting Trellises

Heide Gluesing-Luerssen and G. David Forney, Jr.[*]

November 13, 2018

Abstract

This paper investigates tail-biting trellis realizations for linear block codes. Intrinsic trellis properties are used to characterize irreducibility on given intervals of the time axis. It proves beneficial to always consider the trellis and its dual simultaneously. A major role is played by trellis properties that amount to observability and controllability for fragments of the trellis of various lengths. For fragments of length less than the minimum span length of the code it is shown that fragment observability and fragment controllability are equivalent to irreducibility. For reducible trellises, a constructive reduction procedure is presented. The considerations also lead to a characterization for when the dual of a trellis allows a product factorization into elementary ("atomic") trellises.

1 Introduction

The powerful performance of iterative decoding algorithms for codes on graphs has made graphical models a major topic in coding theory. In particular, it has led to a vivid interest in optimal graphical representations of (linear) block codes. For cycle-free graphs, this realization theory is by now well understood. In this case, a realization is minimal if and only if it is trim and proper (i.e., every state occurs in some constraint codeword and no constraint codeword is supported by a single state variable), and minimal realizations are unique up to state space isomorphisms. Moreover, every non-minimal realization can be reduced to a minimal one by a process of trimming and merging. For details on all of this, see [7, 10], as well as the excellent survey [19] for the special case of conventional trellis realizations.

The focus of this paper is on linear tail-biting trellises, which form the simplest type of realizations on graphs with cycles. Tail-biting trellises gained a lot of attention after the appearance of [2], where it was shown that, for a given code, the complexity of a tail-biting trellis realization may be considerably lower than that of the best conventional (i.e., cycle-free) trellis. This resulted in increased study of (minimal) tail-biting trellises [15, 18, 11, 13, 17, 9, 8, 3] as well as for normal realization on general graphs [4, 16, 10, 1].

A systematic theory of tail-biting trellis realizations was initiated by Koetter/Vardy in their landmark paper [13]. Among other things, they highlighted the fundamental problem that all meaningful concepts of complexity measures for tail-biting trellises lead to different (pre-)orderings,

all of which are only partial. As a consequence, a given code does not have a unique minimal trellis realization. By extending factorization ideas from conventional trellises, Koetter/Vardy showed in [12] that every reduced trellis (i.e., all states and branches appear on valid trajectories) is a product trellis, that is, it can can be obtained as the product of elementary trellises. In [13], they showed that product trellises based on "shortest generators" of the code in a circular interval sense, form a reasonably small class of trellises which is guaranteed to contain all minimal trellises.

In [9, 8] it is shown that these trellises, called KV-trellises, enjoy nice properties; for instance, they are non-mergeable, and the dual of a KV-trellis (in the sense of normal realization dualization [4]) is a KV-trellis. As a consequence, the dual of a KV-trellis is a product trellis representing the dual code. The last property is remarkable because in general the dual of a product trellis (even if non-mergeable) is not a product trellis. Yet, this invariance under dualization does not characterize KV-trellises, and in fact, no intrinsic characterizations for being a KV-trellis or being minimal are known. As a consequence, no constructive method for reducing a given realization to a minimal one has been found yet.

In this paper, we study tail-biting trellises, and make some progress on these aforementioned topics. We derive constructive procedures to reduce the complexity of a trellis, and derive irreducibility criteria.

We began our investigation of irreducibility in [7], where we considered normal realizations on general graphs. In that paper, a local reduction was defined as the replacement of one state space by a smaller one, and an adjustment of the adjacent constraint codes, without changing the rest of the realization or the code that is realized. We showed that a realization that is not trim, proper, observable and controllable (TPOC) may be locally reduced by a trimming or merging operation on an appropriate state space. We note that in the prior literature, starting with [20], merging has been studied much more than trimming, no doubt because it seems obvious to trim unused states; however, because trimming and merging are dual operations [7], we weight them equally.

This paper continues our investigation for the special case of linear tail-biting trellises. In this case, the reductions considered in [7] may be regarded as reductions of trellis fragments of length 2 (two consecutive constraint codes and the state space involved in both of them), and it appears natural to extend such reductions to fragments of any length. As we will see, reducibility is then closely related to properties of the dual trellis realization. Indeed, if the given trellis or its dual lacks certain basic, easily detectable properties, then both can be reduced simultaneously. In this sense, the dual trellis may reveal defects that are not immediately apparent in the primal trellis. It is therefore natural to treat a trellis and its dual on an equal footing, so that all reductions come with an analogue for the dual trellis.

In this fashion we can show that the necessary properties for irreducibility on fragments of length 2, presented in [7], extend to necessary conditions for irreducibility on longer fragments. They amount to fragment observability and controllability, and are closely related to trimness of the trellis and its dual in a fragment sense. We also prove that these conditions are sufficient for irreducibility on fragments of length less than the minimum span length of the code, which is a measure of the lengths of zero runs in the codewords. We then discuss the remaining case of reducibility on longer fragments, and illustrate with an example how this problem may be approached; however, this case remains largely open.

Finally, we relate our results to the approach taken by Koetter and Vardy in [12, 13]. As mentioned earlier, they investigated the class of reduced trellises and narrowed it further down

to KV-trellises in their search for minimal trellises. In this paper, we do not assume reducedness because the dual of a reduced trellis is not necessarily reduced. In fact, our results will indeed provide an easy-to-check criterion for when the dual of an observable reduced trellis is a reduced trellis. We also present a summary of various trellis classes and their relationship.

2 Codes, Trellises, and Reducibility

In this section we introduce the basic notions for trellises as needed in this paper. Most of it is simply a specialization of the terminology used in [7] for general normal realizations. In addition, we also define local reductions of trellises; this will be more refined than the definition used in [7].

Throughout, a linear block code $\mathcal{C}$ over a finite field $\mathbb{F}$ is a subspace of a *symbol configuration space* $\mathcal{A} = \Pi_{i=0}^{m-1} \mathcal{A}_i$, where each *symbol alphabet* $\mathcal{A}_i$ is a finite-dimensional vector space over $\mathbb{F}$.

A *linear tail-biting trellis realization* $\mathcal{R}$ *of length* m is a normal linear realization, in the sense of [4] or [7], on a graph that is a single cycle of length m. Thus, it consists of a set of symbol spaces $\mathcal{A}_i$, a set of state spaces $\mathcal{S}_i$, and a set of constraint codes $\mathcal{C}_i \subseteq \mathcal{S}_i \times \mathcal{A}_i \times \mathcal{S}_{i+1}$, where all index sets are equal to $\mathbb{Z}_m$ and index arithmetic is modulo m. Every constraint code thus involves precisely two state variables. All variable alphabets are finite-dimensional vector spaces over $\mathbb{F}$. The elements of $\mathcal{C}_i$ (called *constraint codewords*, or *transitions*, or *branches*) will be written as (s_i, a_i, s_{i+1}).

As noted in [7], if any state space $\mathcal{S}_i$ is trivial, then we may simply delete it. The graph of $\mathcal{R}$ then becomes a finite path, and the realization becomes a conventional linear trellis realization of length m. Thus finite conventional trellis realizations may be regarded as special cases of tail-biting trellis realizations.

Henceforth, we will call a linear tail-biting trellis realization simply a *trellis*.

The space $\mathcal{S} = \prod_{i=0}^{m-1} \mathcal{S}_i$ is called the *state configuration space* of $\mathcal{R}$. The *behavior* is the set $\mathfrak{B}$ of all *trajectories* (or *configurations*) $(\mathbf{a}, \mathbf{s}) \in \mathcal{A} \times \mathcal{S}$ such that all constraints are satisfied, i.e., $(s_i, a_i, s_{i+1}) \in \mathcal{C}_i$ for all $i \in \mathbb{Z}_m$. The pairs $(\mathbf{a}, \mathbf{s}) \in \mathfrak{B}$ are called *valid* trajectories (or configurations). The code $\mathcal{C}$ *generated by the trellis* is the set of all symbol trajectories $\mathbf{a} \in \mathcal{A}$ that appear in some $(\mathbf{a}, \mathbf{s}) \in \mathfrak{B}$. A code generated by a linear trellis is linear.

The *dual trellis of a trellis* $\mathcal{R}$, denoted by $\mathcal{R}^\circ$, is defined as the trellis with the same index set in which the symbol and state spaces $\mathcal{A}_i$ and $\mathcal{S}_i$ are replaced by their linear algebra duals $\hat{\mathcal{A}}_i$, $\hat{\mathcal{S}}_i$ (which are unique up to isomorphism), the constraint codes $\mathcal{C}_i$ are replaced by their orthogonal codes $\mathcal{C}_i^\perp$ under the standard inner product, and the sign of each dual state variable is inverted in one of the two constraints in which it is involved. For trellises it is convenient to apply the sign inversion to $\hat{s}_{i+1}$; thus a dual trajectory $(\hat{\mathbf{a}}, \hat{\mathbf{s}})$ is valid if and only if $(\hat{s}_i, \hat{a}_i, -\hat{s}_{i+1}) \in \mathcal{C}_i^\perp$ for all $i \in \mathbb{Z}_m$. The behavior $\mathfrak{B}^\circ$ of $\mathcal{R}^\circ$ is the space of all such valid dual trajectories. The *Normal Realization Duality Theorem* [4] states that if $\mathcal{R}$ realizes the code $\mathcal{C}$, then its dual $\mathcal{R}^\circ$ realizes the orthogonal code $\mathcal{C}^\perp$. For other proofs see [16, 1, 5, 7].

Trellises $\mathcal{R}$ and $\tilde{\mathcal{R}}$ of length m with state spaces and constraint codes $\mathcal{S}_i$, $\mathcal{C}_i$, $\tilde{\mathcal{S}}_i$, $\tilde{\mathcal{C}}_i$ are called *isomorphic* if there exist state-space isomorphisms $\varphi_i : \mathcal{S}_i \to \tilde{\mathcal{S}}_i$ such that $\tilde{\mathcal{C}}_i = \{(\varphi_i(s_i), a_i, \varphi_{i+1}(s_{i+1})) \mid (s_i, a_i, s_{i+1}) \in \mathcal{C}_i\}$.[1] Evidently, isomorphic trellises realize the same code. Following [13, Def. 3.1], we say a trellis $\tilde{\mathcal{R}}$ is *smaller* than $\mathcal{R}$ if their state spaces satisfy $\dim \tilde{\mathcal{S}}_i \leq \dim \mathcal{S}_i$ for all i, and we

---

[1]It is interesting to note that for connected trellises this definition may be relaxed to requiring that the maps φ_i only be bijections rather than isomorphisms. Indeed, Conti proved recently [3, Thm. 3.28, Cor. 6.6] that the isomorphism classes of connected trellises coincide for these two notions.

call $\tilde{\mathcal{R}}$ *strictly smaller* than $\mathcal{R}$ if we have at least one strict inequality. A trellis is called *minimal* if there is no strictly smaller trellis realizing the same code.

In [7, Thm. 3] it has been shown that realizations on cycle-free graphs are minimal if and only if they are trim and proper (see the next section for the definitions), and that a nonminimal realization can be reduced in a constructive way. It is well known that trimness and properness are not sufficient for minimality of realizations on graphs with cycles. The goal of this paper is to study the particular case of single-cycle graphs, develop constructive methods of reducing the complexity of a given realization on such a graph, and present irreducibility criteria. Definition 2.1 below will be the central concept of this paper.

A major part of our approach will be the analysis of trellis fragments. We use the following notation. For $j, k \in \mathbb{Z}_m$ and $j \neq k$ let $[j, k) := \{j, j+1, \ldots, k-1\} \subseteq \mathbb{Z}_m$ denote a (possibly circular) subinterval of $\mathbb{Z}_m$; thus $\mathbb{Z}_m$ is the disjoint union of the two complementary subintervals $[j, k)$ and $[k, j)$. Correspondingly, a trellis $\mathcal{R}$ of length m may be divided into two cycle-free *complementary fragments*, denoted by $\mathcal{R}^{[j,k)}$ and $\mathcal{R}^{[k,j)}$, by cutting the edges associated with states $\mathcal{S}_j$ and $\mathcal{S}_k$. The fragment $\mathcal{R}^{[j,k)}$ includes all symbol spaces and constraint codes with indices in $[j, k)$, and $\mathcal{R}^{[k,j)}$ includes all with indices in $[k, j)$. The fragment $\mathcal{R}^{[j,k)}$ also contains the state spaces $\mathcal{S}_i$ with indices in (j, k) as *internal state spaces*, and similarly $\mathcal{R}^{[k,j)}$ contains the internal state spaces $\mathcal{S}_i$ with $i \in (k, j)$. The two boundary state spaces $\mathcal{S}_j$ and $\mathcal{S}_k$ may be regarded as *external state spaces* in both fragments. Note that the internal state spaces have degree 2 and correspond to normal edges, whereas the external state spaces have degree 1 and correspond to half-edges, like symbol spaces.

We extend this notation by defining $[j, j + m)$ to be the entire time axis $\mathbb{Z}_m$ "starting at j", and the complementary interval $[j + m, j)$ to be the empty interval "starting at j". Then $\mathcal{R}^{[j+m,j)}$ denotes the cycle-free fragment consisting of all of $\mathcal{R}$ except for the edge $\mathcal{S}_j$, while $\mathcal{R}^{[j+m,j)}$ denotes the complementary fragment consisting only of the edge $\mathcal{S}_j$. Both fragments have two external state variables with common alphabet $\mathcal{S}_j$, as we will discuss further in Section 5.

Definition 2.1 Let $[j, k)$ be a non-empty interval. A $[j, k)$-*reduction* $\tilde{\mathcal{R}}$ of a trellis $\mathcal{R}$ is a replacement of the state spaces $\mathcal{S}_{j+1}, \mathcal{S}_{j+2}, \ldots, \mathcal{S}_{k-1}$ by state spaces $\tilde{\mathcal{S}}_{j+1}, \tilde{\mathcal{S}}_{j+2}, \ldots, \tilde{\mathcal{S}}_{k-1}$ of at most the same size and the adjacent constraint codes $\mathcal{C}_j, \ldots, \mathcal{C}_{k-1}$ by suitable constraint codes $\tilde{\mathcal{C}}_j, \ldots, \tilde{\mathcal{C}}_{k-1}$ of any size, without changing the rest of the realization or the code $\mathcal{C}$ that it realizes. We also call this a t-*reduction*, where $t = (k - j) \bmod m \in \{1, \ldots, m\}$. The reduction will be called *strict* if at least one of the state space sizes decreases strictly, and *conservative* if none of the constraint code sizes increases. A trellis $\mathcal{R}$ is called t-*irreducible* if each t-reduction is isomorphic to $\mathcal{R}$.

Note that a $[j, k)$-reduction affects the constraint codes and the internal state spaces in the fragment $\mathcal{R}^{[j,k)}$, but not its external state spaces $\mathcal{S}_j$ and $\mathcal{S}_k$. Thus a t-reduction affects t constraint codes and $t - 1$ state spaces. An m-reduction affects all constraint codes, and all but one state space.

If a trellis $\mathcal{R}$ is (strictly) t-reducible, then it is (strictly) t'-reducible for all $t' \geq t$. Furthermore, it is immediate from the dualization of trellis realizations, that a trellis $\mathcal{R}$ is (strictly) t-reducible if and only if the dual trellis $\mathcal{R}^\circ$ is (strictly) t-reducible.

We note in passing that a minimal trellis may have a conservative t-reduction for some $t > 1$. This is due to the fact that a code may have non-isomorphic minimal trellises with the same state space and constraint code dimensions [9, Ex. III.16]. However, we will see later that minimal trellises are always 1-irreducible, which is to say that no single constraint code can be replaced by any other

4

constraint code without changing the code realized by the trellis.

The primary goal of our reduction procedures will be the reduction of state spaces, but we will also address the constraint code dimensions. While state space dimensions do not change under dualization, this is not the case for the constraint code dimensions. This causes the constraint code dimensions to be less predictable in general.

It will become clear later that non-strict t-reductions form indeed a useful concept — even though they may not immediately lead to a net decrease of the state space sizes and may even increase the constraint code sizes. Non-strict reductions will be used to produce strictly reducible trellises so that ultimately a net reduction in state space sizes is achieved. A particular instance of a non-strict reduction is a 1-reduction that consists of the replacement of a single constraint code (by one that may be bigger or smaller), and thus does not alter the state complexity profile of the realization.

The most important instances of strict and conservative 2-reductions are the mutually dual processes of trimming and merging. We briefly recall these concepts. Let $\mathcal{R}$ be a trellis with state spaces $\mathcal{S}_i$ and constraint codes $\mathcal{C}_i$. Fix $j \in \mathbb{Z}_m$ and let $\mathcal{Y}_j$ be a subspace of $\mathcal{S}_j$. We say that $\mathcal{R}$ is *trimmed to* $\mathcal{Y}_j$ if we restrict the state space $\mathcal{S}_j$ to $\mathcal{Y}_j$ and restrict the two adjacent constraint codes $\mathcal{C}_{j-1}$ and $\mathcal{C}_j$ accordingly. We say that $\mathcal{R}$ is *merged to the quotient space* $\mathcal{S}_j/\mathcal{Y}_j$ if we replace the state space $\mathcal{S}_j$ by the quotient space $\mathcal{S}_j/\mathcal{Y}_j$ and replace the states at time j in the two adjacent constraint codes $\mathcal{C}_{j-1}$ and $\mathcal{C}_j$ by their cosets modulo $\mathcal{Y}_j$. Projection/cross-section duality (given in (2.1) below) implies that the trellis $\mathcal{R}'$ is obtained from $\mathcal{R}$ by trimming $\mathcal{S}_j$ to the subspace $\mathcal{Y}_j$ if and only if $(\mathcal{R}')^\circ$ is obtained from $\mathcal{R}^\circ$ by merging $\hat{\mathcal{S}}_j$ to $\hat{\mathcal{S}}_j/\mathcal{Y}_j^\perp$. For a proof, further details and a graphical illustration of the duality of trimming and merging see [7, Sec. III.B]. In general, the trimmed/merged realization generates a different code than the original realization. We will, of course, be interested in the case where the code does not change after trimming/merging. In this case, trimming and merging obviously form simultaneous strict and conservative 2-reductions of the trellis and its dual. For the notions of non-mergeability and non-trimmability, see the next section.

We close this section by briefly recalling the projection/cross-section duality theorem. This identity is one of the most fundamental and useful duality relationships for linear codes and will be used frequently throughout this paper. Let $\mathcal{C}$ be a subspace contained in a vector space $\mathcal{T} = \mathcal{T}_1 \times \mathcal{T}_2$. The *projection* and *cross-section* of $\mathcal{C}$ on $\mathcal{T}_1$ are defined as $\mathcal{C}_{|\mathcal{T}_1} := \{t_1 \mid \exists\, t_2 \in \mathcal{T}_2 : (t_1, t_2) \in \mathcal{C}\}$ and $\mathcal{C}_{:\mathcal{T}_1} := \{t_1 \mid (t_1, 0) \in \mathcal{C}\}$, respectively. Suppose we have inner products between $\mathcal{T}_i$ and $\hat{\mathcal{T}}_i$ for each i which we extend in the natural way to $\mathcal{T}_1 \times \mathcal{T}_2$ and its dual $\hat{\mathcal{T}}_1 \times \hat{\mathcal{T}}_2$. Projection/cross-section duality [4] (see also [7, Sec. II.H]) states that

$$\left(\mathcal{C}_{:\mathcal{T}_i}\right)^\perp = (\mathcal{C}^\perp)_{|\hat{\mathcal{T}}_i} \tag{2.1}$$

for any subspace $\mathcal{C} \subseteq \mathcal{T}_1 \times \mathcal{T}_2$ and its orthogonal subspace $\mathcal{C}^\perp \subseteq \hat{\mathcal{T}}_1 \times \hat{\mathcal{T}}_2$.

3 Local and Global Trellis Properties

We recall some basic properties of trellises as they have been discussed in detail in [7] for general normal graphs, and discuss some subtleties related to these notions.

We begin with local trellis properties. A trellis is called *trim at state space* $\mathcal{S}_i$ if $(\mathcal{C}_{i-1})_{|\mathcal{S}_i} = \mathcal{S}_i = (\mathcal{C}_i)_{|\mathcal{S}_i}$, and *proper at state space* $\mathcal{S}_i$ if the cross-sections $(\mathcal{C}_{i-1})_{:\mathcal{S}_i}$ and $(\mathcal{C}_i)_{:\mathcal{S}_i}$ are both trivial.

The former means that each state in $\mathcal{S}_i$ has an incoming branch and an outgoing branch while the latter means that there are no nontrivial branches of the form $(0, 0, s_i)$ in $\mathcal{C}_{i-1}$ and none of the form $(s_i, 0, 0)$ in $\mathcal{C}_i$. As in [7] we call a trellis *trim* (resp. *proper*) if it is trim (resp. proper) at each state space. (In the prior literature, e.g. [13] and [19], "proper" is often called "biproper".)

Using projection/cross-section duality (2.1) one obtains immediately the following.

Theorem 3.1 ([8, 7]) *Let $j \in \{i-1, i\}$. The projection of $\mathcal{C}_j$ on $\mathcal{S}_i$ is surjective if and only if the cross-section of $\mathcal{C}_j^\perp$ on $\hat{\mathcal{S}}_i$ is trivial. Thus, $\mathcal{R}$ is trim at $\mathcal{S}_i$ if and only if $\mathcal{R}^\circ$ is proper at $\hat{\mathcal{S}}_i$.*

Let us now turn to global trellis properties. A trellis is said to be *state-trim* if each state appears on a valid trajectory, i.e., $\mathfrak{B}_{|\mathcal{S}_i} = \mathcal{S}_i$ for all i. It is clear that a state-trim trellis is trim. We call a trellis *branch-trim* if each branch appears on a valid trajectory, i.e., $\mathfrak{B}_{|\mathcal{S}_i \times \mathcal{A}_i \times \mathcal{S}_{i+1}} = \mathcal{C}_i$ for all i. It is well known that a state-trim trellis need not be branch-trim; see, e.g., Fig. 3(b). In [13], a trellis that is both state-trim and branch-trim is called *reduced*, and all linear trellises are assumed to be reduced. We do not adopt this stance here since the dual of a reduced trellis is not necessarily reduced (see, for example, Fig. 1).

If a trellis or its dual is not branch-trim, then both are 1-reducible, since in this case there exists a 1-reduction that is not isomorphic to the given trellis. The converse will be discussed in Theorem 7.1.

We define a trellis to be *nonmergeable* (in the linear sense) if no state space can be merged to a proper quotient space without changing the code generated by the trellis. This is the only kind of merging that preserves linearity of the trellis. Our notion of non-mergeability differs from the definition in [13], which simply requires that no two states can be merged without changing the code. For example, the non-state-trim trellis of Fig. 1(b) below is nonmergeable in our sense, but is mergeable in the sense of [13], since states 00 and 11 in $\mathcal{S}_2$ can be merged to produce a smaller nonlinear trellis that realizes the same code. However, it is easy to show that for state-trim trellises this situation cannot arise (since whenever $s, s' \in \mathcal{S}_i$ can be merged, then $\mathcal{S}_i$ can be merged to the quotient space $\mathcal{S}_i / \langle s - s' \rangle$ without changing the code); see also [3, Observ. 2.8], where "state-trim" is called "almost reduced".

Dually, a trellis is called *non-trimmable* if it does not allow a proper trimming (in the sense of Section 2) resulting in a trellis for the same code. Duality of trimming and merging, as described in Section 2, implies that a trellis $\mathcal{R}$ is non-mergeable if and only if its dual $\mathcal{R}^\circ$ is non-trimmable. It is worth noting that a state-trim trellis may be trimmable. This will be addressed in Theorem 4.4.

A trellis $\mathcal{R}$ that realizes a code $\mathcal{C}$ is called *observable* (or *one-to-one*) if for each $\mathbf{a} \in \mathcal{C}$ there is precisely one valid trajectory $(\mathbf{a}, \mathbf{s}) \in \mathfrak{B}$. The trellis is called *controllable* if its dual $\mathcal{R}^\circ$ is observable. These definitions are discussed in [7, Sec. IV-C], where it is shown that controllability in this sense is equivalent to having independent constraints.

We will use the following controllability test from [7, Thm. 6].

Theorem 3.2 *For every trellis $\mathcal{R}$ we have $\sum_i \dim \mathcal{C}_i \leq \dim \mathfrak{B} + \sum_i \dim \mathcal{S}_i$, with equality if and only if $\mathcal{R}$ is controllable.*

In other words, Theorem 3.2 says that the total constraint dimension $\sum_i \dim \mathcal{C}_i$ is maximized if and only if the trellis is controllable.

As noted in [7, Footnote 4], this theorem is related to Theorem 4.6 of [13] as follows. Koetter and Vardy show that if a reduced (i.e., state-trim and branch-trim) product trellis is observable (so

dim $\mathfrak{B} = \dim \mathcal{C}$), and no generator has a degenerate span equal to the entire time axis $\mathbb{Z}_m$, then (in our notation) $\sum_i \dim \mathcal{C}_i = \dim \mathfrak{B} + \sum_i \dim \mathcal{S}_i$, so the realization is controllable (in our terminology).

A key property of an uncontrollable trellis $\mathcal{R}$ is that, under weak conditions, its valid trajectories partition into disconnected cosets (for examples, see Figs. 2(b) and 4(b)). In Appendix A, we show that this property holds provided that $\mathcal{R}$ is state-trim, which is a weaker condition than the "reduced" condition of [13]. Moreover, we give examples showing that state-trimness is indeed a necessary condition.

4 Basic Results about Irreducibility and Motivating Examples

We start by presenting a list of necessary conditions for 2-irreducibility that has been derived in [7]. The rest of the section is devoted to examples illustrating that these properties are not sufficient and that the dual of a trellis may be helpful in revealing the reducibility of both the trellis and its dual.

Recall that a (strict) 2-reduction consists of reducing one state space and altering the two adjacent constraint codes. Obviously trimming and merging are 2-reductions.

Theorem 4.1 ([7, Thm. 2, Thm. 9]) *A 2-irreducible realization $\mathcal{R}$ must be trim, proper, observable, and controllable, else there exists a strict and conservative 2-reduction of $\mathcal{R}$, whose dual is a strict and conservative 2-reduction of $\mathcal{R}^\circ$.*

In view of Theorem 4.1, we will abbreviate "trim, proper, observable and controllable" by TPOC.

The following fact has been mentioned already in [11, Sec. 4].

Remark 4.2 Each 2-irreducible realization must be state-trim and nonmergeable, else there exists a 2-reduction in form of a state-trimming or state-merging.

Since 2-reductions of unobservable trellises will be crucial later, we present the process for further reference in the next remark.

Remark 4.3 In the proof of [7, Thm. 9] it has been shown that whenever a trellis $\mathcal{R}$ contains a nontrivial unobservable valid trajectory $(\mathbf{0}, \mathbf{s})$, one may pick any of its nonzero states, say $s_i \in \mathcal{S}_i$, and trim the state space $\mathcal{S}_i$ to any subspace $\mathcal{T}_i$ satisfying $\mathcal{T}_i \oplus \langle s_i \rangle = \mathcal{S}_i$. This results in a trellis realization of the same code (which may be not trim at time $i-1$ or $i+1$, in which case it can be further reduced). Being a trimming, this process forms a strict and conservative 2-reduction, and the dual merging process is a strict and conservative 2-reduction of $\mathcal{R}^\circ$. More precisely, in the trimming of $\mathcal{R}$ the branches $(s_{i-1}, 0, s_i)$ and $(s_i, 0, s_{i+1})$ in the constraint codes $\mathcal{C}_{i-1}$ and $\mathcal{C}_i$ are deleted, and thus the dimensions of these constraint codes decrease by 1. On the other hand, the constraint code dimensions of the dual merging stay the same, since $\dim \mathcal{C}_i^\perp = \dim \mathcal{S}_i + \dim \mathcal{S}_{i+1} + \dim \mathcal{A}_i - \dim \mathcal{C}_i$.

The following theorem is essentially due to Koetter [11, Thm. 9]. For the second statement recall that non-mergeability is defined in the linear sense (see Section 3) and is thus dual to non-trimmability.

Theorem 4.4 *A trellis is non-trimmable if and only if it is observable and state-trim. Dually, a trellis $\mathcal{R}$ is nonmergeable if and only if the dual trellis $\mathcal{R}^\circ$ is observable and state-trim.*

PROOF: As we have just seen, an unobservable trellis is trimmable. As noted by Koetter [11, Sec. 4], an observable trellis is trimmable if and only if it is not state-trim. The last part follows from the duality of state-merging and state-trimming discussed in Section 2. □

In order to derive further necessary conditions for strict irreducibility, we present two examples of mutually dual trellises. Both illustrate that the dual may reveal some shortcomings of the trellis and its dual that are not directly discernible from the primal trellis (an observation that has been made already by Koetter in [11]), and which cause the trellis and its dual to be strictly reducible. Motivated by these phenomena, we will study strict reducibility simultaneously for a trellis and its dual.

We assume that the reader is familiar with product trellises [12], [13, Sec. IV.C], [14, Sec. III]. Every "reduced" (i.e., state-trim and branch-trim) trellis is a product trellis in the sense that its behavior $\mathfrak{B}$ has a basis consisting of $\dim \mathfrak{B}$ one-dimensional "atomic" sub-behaviors, each characterized by a codeword and its (possibly circular) span.

Example 4.5 We present a trellis that is TPOC, and yet strictly 2-reducible. This shows that Theorem 4.1 provides necessary but not sufficient conditions for 2-irreducibility. We will see that the dual trellis, which is also TPOC, is not state-trim, and thus the primal trellis is mergeable due to Theorem 4.4. As a consequence, both trellises are strictly 2-reducible. It will become clear that the non-state-trimness of the dual trellis is easy to detect, whereas the mergeability of the primal trellis is much less obvious.

Fig. 1(a) shows a trellis realization of the code $\mathcal{C} = \{000, 110, 101, 011\} \subseteq \mathbb{F}_2^3$.[2] The symbol spaces are $\mathcal{A}_i = \mathbb{F}_2$ for $i \in \mathbb{Z}_3$. The realization is the product trellis obtained from the generators $\underline{1}01$, $\underline{1}\,\underline{1}0$ with the indicated circular spans. The trellis appeared first in [13, Fig. 5], and it has been used subsequently for various purposes in [17, Ex. 1], [9, Ex. IV.8], and [3, Sec. 7.2]; however, its dual (Fig. 1(b)) has not been discussed in any of these papers.
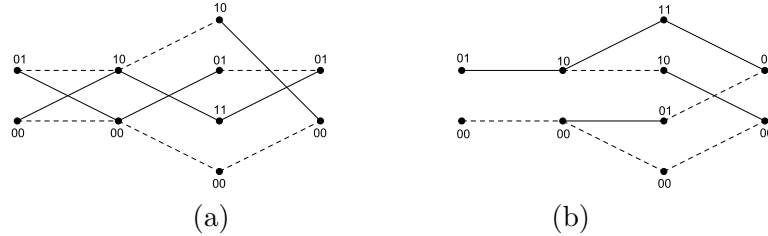

Figure 1: TPOC, state-trim and branch-trim trellis and its non-state-trim dual

Fig. 1(b) shows the dual trellis, which generates the dual code $\mathcal{C}^\perp = \{000, 111\}$. For the dualization we choose $\hat{\mathcal{S}}_i = \mathcal{S}_i$ and the standard inner product for all dual state spaces; no sign inverter is needed. Both trellises are TPOC. The trellis in (a) is state-trim, whereas the dual trellis in (b) is not state-trim: the states 10 and 01 at time 2 are not on any valid trajectory of this realization. Thus by Remark 4.2 both trellises are strictly 2-reducible, and in particular, the trellis in (a) is mergeable due to Theorem 4.4. Trimming the dual trellis to the state space $\{00, 11\}$ at time 2 and

---

[2] **Trellis Illustration Conventions.** Throughout the paper, we will use the following conventions for drawing trellises: dashed (resp. solid) lines denote branches with symbol variable 0 (resp. 1). Most of the time, especially for dual pairs of trellises, a choice of state labels will be shown as well. It should be kept in mind that the state labeling does not play any particular role and may be changed at each index using a state space isomorphism. The states in $\mathcal{S}_0$ will always appear in the same ordering at the beginning and end of the trellis.

dually merging the primal trellis leads to the pair of mutually dual trellises shown in Figure 2. They form $[1, 0)$-reductions of the trellises in Fig. 1.

(a)   (b)

Figure 2: State-merged trellis and state-trimmed dual trellis

The trellises in Fig. 2(a) and 2(b) are standard small examples of unobservable and uncontrollable trellises, respectively. As discussed in Remark 4.3, both can be further reduced; the reduced trellises in this case are conventional, trim and proper, and therefore minimal.   □

Example 4.6 We present a trellis that is TPOC, state-trim, branch-trim, nonmergeable, and yet strictly 2-reducible. The dual trellis is also TPOC, state-trim, and nonmergeable (due to the previous duality results), but is not branch-trim. This tells us that both trellises are reducible. As in the previous example, the reducibility of these trellises will be obvious from the dual trellis, but is not at all evident from the primal trellis.

The example appeared first in [9, 8]. Fig. 3(a) shows the product trellis obtained from the generators $01\underline{110}$, $\underline{100}1\underline{0}$, $\underline{01}\,1\underline{01}$ with the indicated circular spans. Its dual is shown in Fig. 3(b).
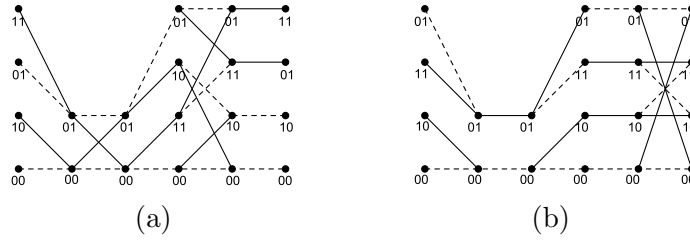

(a)   (b)

Figure 3: TPOC, state-trim, branch-trim, nonmergeable trellis and its dual

Both trellises are TPOC; they are easily seen to be state-trim, and thus both are nonmergeable due to Theorem 4.4. The trellis in (a) is also branch-trim, but the dual trellis is not: the diagonal branches of the last constraint code are not on any valid trajectory, and deleting them does not change the code generated by that trellis. This provides us with a 1-reduction where we replace the 3-dimensional constraint code $\mathcal{C}_4 = \langle 00|1|01,\ 10|1|10,\ 11|1|11 \rangle$ by the subspace $\tilde{\mathcal{C}}_4 = \langle 10|1|10,\ 11|1|11 \rangle$. It results in the trellis in Fig. 4(b). By duality, the trellis in Fig. 3(a) must also be 1-reducible. The dual process consists of replacing $\mathcal{C}_4^{\perp} = \langle 10|0|10,\ 11|1|01 \rangle$ by the supercode $\tilde{\mathcal{C}}_4^{\perp} = \langle 10|0|10,\ 11|1|01,\ 01|0|01 \rangle$. This results in the trellis in Fig. 4(a), which then is the dual of that in Fig. 4(b).

Evidently, the trellis in Fig. 4(a) is unobservable and by duality (or by Theorem A.1) the trellis in Fig. 4(b) is uncontrollable. Thus, we may apply the procedure from Remark 4.3 and trim the first trellis in a suitable way. We pick time 4 and trim to the subspace $\{00, 11\}$; dually, we merge the dual state space to $\mathbb{F}_2^2/\{00, 11\}$. This results in the mutually dual trellises shown in Fig. 5.[3]

---

[3]Observe that we could also have trimmed the state space $\mathcal{S}_4$ to $\{00, 10\}$ or the state space $\mathcal{S}_0$ suitably. Evidently these various options could lead to different results; for another one, see [7, Ex. 4].
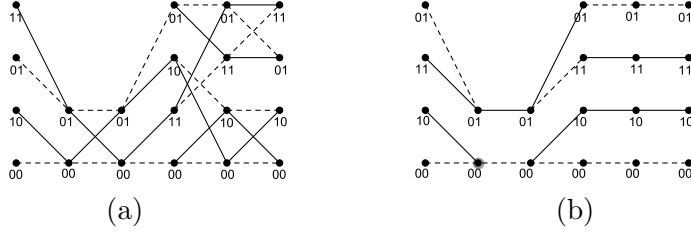
9

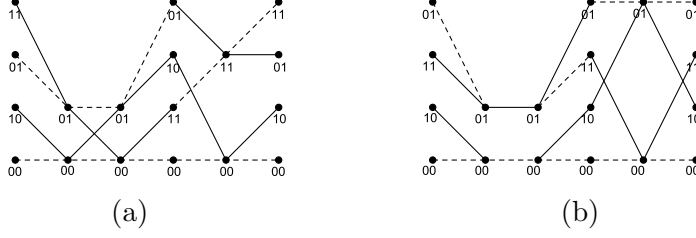Figure 4: Branch-expanded and branch-trimmed dual trellis

Figure 5: State-trimmed trellis and state-merged dual trellis

These trellises still generate the original code $\mathcal{C} = \langle 01110, 10010, 01101 \rangle$ and its dual $\mathcal{C}^{\perp}$, respectively. They form strict and conservative 2-reductions of the trellises in Fig. 3 because no state space has changed except $\mathcal{S}_4$, which is smaller, and only the constraint codes at times 3 and 4 have changed, and none is larger. $\square$

Summarizing, we observe that the trellis in Fig. 3(a) is strictly 2-reducible, even though it is TPOC, state-trim, branch-trim, and nonmergeable. We first had to perform an auxiliary branch-addition (a non-conservative 1-reduction) before a state-trimming resulted in a strict and conservative 2-reduction. In the next sections we will derive the appropriate concepts for a systematic study of these phenomena.

It is also worth noting that even though the trellises in Fig. 1(a) and Fig. 3(a) are both product trellises, their duals are not – simply because they are not branch-trim, a property all product trellises share. In Theorem 5.6 we will give an intrinsic characterization of those product trellises that have a product trellis dual.

5 Trellis Fragments and Global Trellis Properties

In this section we study trellis fragments and show the duality of controllability and observability of such fragments. This will allow us to present some relations between various trellis properties, which then in turn leads to a characterization of when the dual of a product trellis is a product trellis.

Let $\mathcal{R}$ be a trellis of length m, and thus with symbol spaces $\mathcal{A}_i$, state spaces $\mathcal{S}_i$, and constraint codes $\mathcal{C}_i \subseteq \mathcal{S}_i \times \mathcal{A}_i \times \mathcal{S}_{i+1}$, all with index sets equal to $\mathbb{Z}_m$.

Recall from Section 2 the definition of trellis fragments $\mathcal{R}^{[j,k)}$, where $[j,k)$ is any subinterval of $\mathbb{Z}_m$. In particular, the fragment $\mathcal{R}^{[j+m,j)}$ is defined as the edge $\mathcal{S}_j$, i.e., it consists of two external state variables, denoted by s_{j+m} and s_j, with a common alphabet $\mathcal{S}_j$ and an equality constraint

10

$\mathcal{C}_= = \{(s_{j+m}, s_j) \mid s_{j+m} = s_j\}$ between them. This fragment contains no symbol spaces or internal state spaces.

The complementary fragment to $\mathcal{R}^{[j+m,j)}$ is the fragment $\mathcal{R}^{[j,j+m)}$, which consists of $\mathcal{R}$ with the edge $\mathcal{S}_j$ cut out. It contains all constraint codes and symbol spaces of $\mathcal{R}$ and has internal state spaces $\mathcal{S}_i$, where $i \in (j, j+m)$, and two external state variables, with values $s_j \in \mathcal{S}_j$ and $s_{j+m} \in \mathcal{S}_j$.

The *internal behavior* $\mathfrak{B}^{[j,k]}$ of a fragment $\mathcal{R}^{[j,k]}$ is the set of all trajectories $(\mathbf{a}^{[j,k]}, \mathbf{s}^{[j,k]}) \in \mathcal{A}^{[j,k]} \times \mathcal{S}^{[j,k]} := \prod_{i \in [j,k)} \mathcal{A}_i \times \prod_{i \in [j,k]} \mathcal{S}_i$, that satisfy all constraints, hence $(s_i, a_i, s_{i+1}) \in \mathcal{C}_i$ for all $i \in [j, k)$. Such trajectories will be called *valid $[j, k)$-paths*, or simply *valid paths*. The *external behavior* $\mathcal{C}^{[j,k]}$ is the projection of $\mathfrak{B}^{[j,k]}$ on $\mathcal{A}^{[j,k]} \times \mathcal{S}_j \times \mathcal{S}_k$. A fragment $\mathcal{R}^{[j,k]}$ is thus a normal realization of its external behavior $\mathcal{C}^{[j,k]}$. Note that $\mathfrak{B}^{[j+m,j)} = \mathcal{C}^{[j+m,j)} = \mathcal{C}_=$, whereas $\mathfrak{B}^{[j,j+m)}$ is the set of all valid $[j, j+m)$-paths of $\mathcal{R}$.

We will say that a valid $[j, k)$-path $(\mathbf{a}^{[j,k]}, \mathbf{s}^{[j,k]}) \in \mathfrak{B}^{[j,k]}$ *lies on a valid trajectory* if there exists some valid trajectory $(\mathbf{a}, \mathbf{s}) \in \mathfrak{B}$ whose projection on $\mathcal{A}^{[j,k]} \times \mathcal{S}^{[j,k]}$ is $(\mathbf{a}^{[j,k]}, \mathbf{s}^{[j,k]})$. In particular, a valid $[j, j+m)$-path lies on a valid trajectory if and only if $s_j = s_{j+m}$. Notice that a traditional trellis diagram of a tail-biting trellis $\mathcal{R}$ of length m, such as any trellis diagram in this paper, actually illustrates the fragment $\mathcal{R}^{[0,m)}$, and the reader has to identify $\mathcal{S}_m$ with $\mathcal{S}_0$. The behavior $\mathfrak{B}^{[0,m)}$ is not in general the same as $\mathfrak{B}$, because, again, the latter consists only of the valid $[0, m)$-paths for which $s_0 = s_m$.

All trellis fragments are cycle-free. It follows from [7, Thm. 3] that a cycle-free fragment is a minimal realization of its external behavior if and only if all of its constraint codes are trim and proper.

The *dual fragment* $(\mathcal{R}^\circ)^{[j,k]}$ is the dual normal realization to $\mathcal{R}^{[j,k]}$, where the external state spaces (whose degree is one) are taken as symbol spaces. Thus $(\mathcal{R}^\circ)^{[j,k]}$ comprises the dual constraint codes $(\mathcal{C}_i)^\perp, i \in [j, k)$, the dual symbol spaces $\hat{\mathcal{A}}_i, i \in [j, k)$, the internal dual state spaces $\hat{\mathcal{S}}_i, i \in (j, k)$, and the external dual state spaces $\hat{\mathcal{S}}_j$ and $\hat{\mathcal{S}}_k$. In the dual realization $(\mathcal{R}^\circ)^{[j,k]}$, we recall that a sign inversion is applied to the value of $\hat{s}_k$ in $\mathcal{C}_{k-1}^\perp$, but not to the value of $\hat{s}_k$ in $\mathcal{C}_k^\perp$. Consequently, a sign inversion is applied to one of the two occurrences of each internal state space $\hat{\mathcal{S}}_i, i \in (j, k)$, and to the one occurrence of $\hat{\mathcal{S}}_k$, but not to the one occurrence of $\hat{\mathcal{S}}_j$.

The internal behavior $(\mathfrak{B}^\circ)^{[j,k]} \subseteq \hat{\mathcal{A}}^{[j,k]} \times \hat{\mathcal{S}}^{[j,k]}$ of $(\mathcal{R}^\circ)^{[j,k]}$ consists of all *valid $[j, k)$-paths* $(\hat{\mathbf{a}}^{[j,k]}, \hat{\mathbf{s}}^{[j,k]})$, i.e., $(\hat{s}_i, \hat{a}_i, -\hat{s}_{i+1}) \in (\mathcal{C}_i)^\perp$ for all $i \in [j, k)$. Its external behavior $(\mathcal{C}^\circ)^{[j,k]} \subseteq \hat{\mathcal{S}}_j \times \hat{\mathcal{A}}^{[j,k]} \times \hat{\mathcal{S}}_k$ is the set of all $(\hat{s}_j, \hat{\mathbf{a}}^{[j,k]}, -\hat{s}_k)$ such that $(\hat{\mathbf{a}}^{[j,k]}, \hat{s}_j, \hat{s}_k) \in \left((\mathfrak{B}^\circ)^{[j,k]}\right)_{|\hat{\mathcal{A}}^{[j,k]} \times \hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k}$. By normal realization duality, the external behavior of the dual fragment $(\mathcal{R}^\circ)^{[j,k]}$ satisfies $(\mathcal{C}^\circ)^{[j,k]} = (\mathcal{C}^{[j,k]})^\perp$.

For example, for the fragment $\mathcal{R}^{[j+m,j)}$, representing the single edge $\mathcal{S}_j$ and with behavior $\mathfrak{B}^{[j+m,j)} = \mathcal{C}^{[j+m,j)} = \mathcal{C}_=$, this reads as follows. Note that the dual code to $\mathcal{C}_=$ is the *sign inversion constraint code* $\mathcal{C}_\sim = \{(\hat{s}_{j+m}, \hat{s}_j) \in \hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_j \mid \hat{s}_{j+m} = -\hat{s}_j\}$. Thus, the dual fragment $(\mathcal{R}^\circ)^{[j+m,j)}$ has behavior $(\mathfrak{B}^\circ)^{[j+m,j)} = \{(\hat{s}_{j+m}, \hat{s}_j) \mid (\hat{s}_{j+m}, -\hat{s}_j) \in \mathcal{C}_\sim\} = \{(\hat{s}_{j+m}, \hat{s}_j) \mid \hat{s}_{j+m} = \hat{s}_j\}$, which is the equality constraint on $\hat{\mathcal{S}}_j$. In other words, the dual fragment to an edge corresponding to $\mathcal{S}_j$ is an edge corresponding to $\hat{\mathcal{S}}_j$. The external behavior is $(\mathcal{C}^\circ)^{[j+m,j)} = \mathcal{C}_\sim$, the sign inversion constraint.

For any interval $[j, k)$, we define the *transition space* $\mathcal{T}^{[j,k]}$ of the fragment $\mathcal{R}^{[j,k]}$ as the projection $\mathcal{C}^{[j,k]}_{|\mathcal{S}_j \times \mathcal{S}_k}$, and the *unobservable transition space* $\mathcal{U}^{[j,k]}$ as the cross-section $\mathcal{C}^{[j,k]}_{:\mathcal{S}_j \times \mathcal{S}_k}$, hence $\mathcal{U}^{[j,k]} = \{(s_j, s_k) \in \mathcal{S}_j \times \mathcal{S}_k \mid (\mathbf{0}, s_j, s_k) \in \mathcal{C}^{[j,k]}\}$. Thus $\mathcal{T}^{[j,k]}$ consists of all state pairs (s_j, s_k) such that there exists a valid $[j, k)$-path $(\mathbf{a}^{[j,k]}, s_j, \mathbf{s}^{(j,k)}, s_k)$, while $\mathcal{U}^{[j,k]}$ consists of all such pairs for which there is a valid $[j, k)$-path with $\mathbf{a}^{[j,k]} = \mathbf{0}^{[j,k]}$. If the fragment lacks symbol spaces, as with

the edge fragment $\mathcal{R}^{[j+m,j)}$, then the cross-section equals the projection, i.e., $\mathcal{U}^{[j,k)} = \mathcal{T}^{[j,k)}$.

We note immediately that if $\mathcal{C}^{[j,k)}$ is trim and proper (i.e., $\mathcal{C}^{[j,k)}_{|\mathcal{S}_\ell} = \mathcal{S}_\ell$ and $\mathcal{C}^{[j,k)}_{:\mathcal{S}_\ell} = \{0\}$ for $\ell = j, k$), then $\mathcal{T}^{[j,k)}$ is trim and $\mathcal{U}^{[j,k)}$ is proper.

A trellis $\mathcal{R}$ will be called $[j,k)$-*controllable* if $\mathcal{T}^{[j,k)} = \mathcal{S}_j \times \mathcal{S}_k$, and $[j,k)$-*observable* if $\mathcal{U}^{[j,k)} = \{(0,0)\}$. When $[j,k)$ is a subinterval of the conventional discrete time axis $\mathbb{Z}$, these definitions correspond to classical notions of controllability and observability in linear system theory.

For a dual fragment $(\mathcal{R}^\circ)^{[j,k)}$, we similarly define its transition spaces as the set of all state pairs in $\hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k$ for which there exists a valid $[j,k)$-path, and where in addition, for the unobservable transition space, the symbol sequence $\mathbf{a}^{[j,k)}$ is $\mathbf{0}^{[j,k)}$. Precisely, the *transition space* is defined as $(\mathcal{T}^\circ)^{[j,k)} = \{(\hat{s}_j, \hat{s}_k) \in \hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k \mid \exists\, (\mathbf{a}^{[j,k)}, \hat{s}_j, \mathbf{s}^{(j,k)}, \hat{s}_k) \in (\mathfrak{B}^\circ)^{[j,k)}\} = \{(\hat{s}_j, \hat{s}_k) \mid (\hat{s}_j, -\hat{s}_k) \in ((\mathcal{C}^\circ)^{[j,k)})_{|\hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k}\}$, and the *unobservable transition space* is $(\mathcal{U}^\circ)^{[j,k)} = \{(\hat{s}_j, \hat{s}_k) \mid (\hat{s}_j, -\hat{s}_k) \in ((\mathcal{C}^\circ)^{[j,k)})_{:\hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k}\}$.

A dual trellis $\mathcal{R}^\circ$ is $[j,k)$-*controllable* if $(\mathcal{T}^\circ)^{[j,k)} = \hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k$, and $[j,k)$-*observable* if $(\mathcal{U}^\circ)^{[j,k)} = \{(0,0)\}$. As a consequence, we have, just as for $\mathcal{R}$, that $\mathcal{R}^\circ$ is $[j,k)$-controllable if $((\mathcal{C}^\circ)^{[j,k)})_{|\hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k} = \hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k$, and $[j,k)$-observable if $((\mathcal{C}^\circ)^{[j,k)})_{:\hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k} = \{(0,0)\}$, since the sign inversion is evidently immaterial. Now we obtain

Theorem 5.1 *If $\mathcal{R}$ and $\mathcal{R}^\circ$ are dual trellises, then $\{(\hat{s}_j, \hat{s}_k) \mid (\hat{s}_j, -\hat{s}_k) \in (\mathcal{T}^\circ)^{[j,k)}\} = (\mathcal{U}^{[j,k)})^\perp$. In particular, $\mathcal{R}$ is $[j,k)$-observable if and only if its dual $\mathcal{R}^\circ$ is $[j,k)$-controllable.*

PROOF: By $(\mathcal{C}^\circ)^{[j,k)} = (\mathcal{C}^{[j,k)})^\perp$ and projection/cross-section duality (2.1), we have $((\mathcal{C}^\circ)^{[j,k)})_{|\hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k} = (\mathcal{C}^{[j,k)}_{:\mathcal{S}_j \times \mathcal{S}_k})^\perp = (\mathcal{U}^{[j,k)})^\perp$, and the definition of $(\mathcal{T}^\circ)^{[j,k)}$ yields the desired orthogonality. The second statement follows from $\{(0,0)\}^\perp = \hat{\mathcal{S}}_j \times \hat{\mathcal{S}}_k$. □

We note that $\mathcal{R}$ is $[j+m,j)$-observable if and only if $\mathcal{S}_j = \{0\}$. Since $\hat{\mathcal{S}}_j = \{0\}$ if and only if $\mathcal{S}_j = \{0\}$, Theorem 5.1 shows that $\mathcal{R}$ is also $[j+m,j)$-controllable if and only if $\mathcal{S}_j = \{0\}$.

It is worth stressing that all statements pertaining to valid $[j,k)$-paths, $[j,k)$-controllability and $[j,k)$-observability are equally valid for the primal trellis $\mathcal{R}$ and its dual $\mathcal{R}^\circ$. The only slight asymmetry, due to sign inversion, is contained in the external behavior (and thus in the transition spaces), and has been dealt with in the previous result. From this point on, no distinction needs to be made between a primal and a dual trellis.

We next discuss various global notions of trimness. A trellis $\mathcal{R}$ will be called $[j,k)$-*trim* if every valid $[j,k)$-path $(\mathbf{a}^{[j,k)}, \mathbf{s}^{[j,k]})$ lies on a valid trajectory $(\mathbf{a}, \mathbf{s}) \in \mathfrak{B}$. Evidently $\mathcal{R}$ is $[j,k)$-trim if and only if all state pairs (s_j, s_k) in the transition space $\mathcal{T}^{[j,k)}$ also occur in $\mathcal{T}^{[k,j)}$; i.e., $\mathcal{T}^{[j,k)} \subseteq \mathcal{T}^{[k,j)}$. Thus if $\mathcal{R}$ is $[k,j)$-controllable, so $\mathcal{T}^{[k,j)} = \mathcal{S}_j \times \mathcal{S}_k$, then $\mathcal{R}$ is $[j,k)$-trim. However, the converse is not true unless we require also that $\mathcal{R}$ be controllable.

Theorem 5.2 *(a) A $[k,j)$-controllable trellis is $[j,k)$-trim.*
(b) A controllable and $[j,k)$-trim trellis is $[k,j)$-controllable.

PROOF: We have already shown (a). To prove (b), we use the dual trellis $\mathcal{R}^\circ$. As we have seen above, $[j,k)$-trimness of $\mathcal{R}$ implies $\mathcal{T}^{[j,k)} \subseteq \mathcal{T}^{[k,j)}$. By Theorem 5.1 this yields $(\mathcal{U}^\circ)^{[k,j)} \subseteq (\mathcal{U}^\circ)^{[j,k)}$. But this means that for every $(\hat{s}_j, \hat{s}_k) \in (\mathcal{U}^\circ)^{[k,j)}$ there is a valid trajectory in $\mathfrak{B}^\circ$ with $\hat{\mathbf{a}} = \mathbf{0}$. Since $\mathcal{R}^\circ$ is observable, $(\mathcal{U}^\circ)^{[k,j)}$ must be trivial; i.e., $\mathcal{R}^\circ$ must be $[k,j)$-observable. By Theorem 5.1, $\mathcal{R}$ must be $[k,j)$-controllable. □

We now introduce notions of controller and observer memory similar to those of classical linear system theory. For $1 \leq t \leq m$, we will say that $\mathcal{R}$ is *t-controllable* (resp. *t-observable*) if $\mathcal{R}$ is $[j, j + t)$-controllable (resp. $[j, j + t)$-observable) for all length-t intervals $[j, j + t)$.

In particular, a trellis $\mathcal{R}$ is m-controllable if and only if for all j there is a valid $[j, j+m)$-path of length m from any state $s_j \in \mathcal{S}_j$ to any state $s_{j+m} \in \mathcal{S}_j$. Thus m-controllable trellises are not only trim, but also state-trim. Dually, m-observable trellises are proper. Moreover, m-observable trellises are evidently observable, which implies that m-controllable trellises are controllable. However, the converses of these statements are not necessarily true, as we now proceed to show.

We first note that a trellis is state-trim at $\mathcal{S}_j$ if and only if every valid $[j+m, j)$-path $(s_{j+m}, s_j) \in \mathcal{C}^{[j+m,j)}$ lies on a valid trajectory in $\mathfrak{B}$. Then we obtain the following corollary:

Corollary 5.3 *A controllable trellis is m-controllable if and only if it is state-trim. Dually, an observable trellis $\mathcal{R}$ is m-observable if and only if its dual $\mathcal{R}^\circ$ is state-trim. Consequently, a non-mergeable observable trellis is m-observable.*

PROOF: By Theorem 5.2, a controllable trellis $\mathcal{R}$ is $[j, j + m)$-controllable for all j if and only if it is $[j + m, j)$-trim for all j, which is to say if it is state-trim at $\mathcal{S}_j$ for all j. The second statement follows from Theorem 5.1. The last statement follows from Theorem 4.4, which shows that if $\mathcal{R}$ is non-mergeable, then $\mathcal{R}^\circ$ must be state-trim. □

For example, Fig. 1(a) shows an observable trellis that is not m-observable (note the all-zero path between states $s_2 = 01$ and $s_2' = 10$); its dual in Fig. 1(b) is a controllable trellis that is not m-controllable, and not state-trim at $\mathcal{S}_2$.

The last observation of Corollary 5.3 also appears in [3, Thm. 7.8], where an m-observable trellis is called "totally one-to-one."

Similarly, we note that a trellis is branch-trim at constraint code $\mathcal{C}_i$ if and only if every valid $[i, i+1)$-path $(s_i, a_i, s_{i+1}) \in \mathcal{C}_i$ lies on a valid trajectory in $\mathfrak{B}$. Thus we obtain the following corollary:

Corollary 5.4 *A controllable trellis is $(m-1)$-controllable if and only if it is branch-trim. Dually, an observable trellis $\mathcal{R}$ is $(m-1)$-observable if and only if its dual $\mathcal{R}^\circ$ is branch-trim.*

PROOF: By Theorem 5.2, a controllable trellis $\mathcal{R}$ is $[i + 1, i)$-controllable for all i if and only if it is $[i, i + 1)$-trim for all i, which is to say if it is branch-trim at $\mathcal{C}_i$ for all i. The second statement follows from Theorem 5.1. □

For example, Fig. 3(a) shows an observable trellis that is not $(m-1)$-observable (note the all-zero path between states $s_5 = 01$ and $s_4 = 01$); its dual in Fig. 3(b) is a controllable trellis that is not $(m-1)$-controllable, and not branch-trim at $\mathcal{C}_4$.

We remark that in the first statement of Corollary 5.4, controllability is necessary; for example, Figs. 2(b) and 4(b) show uncontrollable trellises that are branch-trim, but not t-controllable for any $t \leq m$.

Combining these corollaries, we obtain the following results. Part (b) can also be found in [3, Thm. 7.8].

Corollary 5.5 *(a) An $(m-1)$-controllable trellis is branch-trim, state-trim and trim.*
(b) A controllable m-observable trellis is nonmergeable and proper.

PROOF: (a) If $\mathcal{R}$ is $(m-1)$-controllable, then it is also m-controllable, so it is not only branch-trim by Corollary 5.4, but also state-trim by Corollary 5.3, and thus trim. (b) If $\mathcal{R}$ is m-observable, then its dual $\mathcal{R}^\circ$ is m-controllable, and thus state-trim by Corollary 5.3. Theorem 4.4 implies that $\mathcal{R}$ is nonmergeable, which in turn implies properness. □

Controllability is necessary in (b) because an uncontrollable trellis is mergeable; for example, the trellis of Fig. 2(b) is 3-observable but uncontrollable, hence mergeable.

Finally, we address the question of when the dual of a product trellis is a product trellis. Koetter and Vardy [10] showed that a trellis is a product trellis if and only if it is "reduced" (state-trim and branch-trim), which gave them a powerful tool in their search for minimal trellises [13]. But Examples 4.5 and 4.6 show that the dual of a product trellis is not necessarily reduced. For observable product trellises, Corollaries 5.4 and 5.5 give us a nice characterization of when the dual is a product trellis.

Theorem 5.6 *(a) If $\mathcal{R}$ is an $(m-1)$-observable trellis, then its dual $\mathcal{R}^\circ$ is a product trellis. (b) If $\mathcal{R}$ is observable but not $(m-1)$-observable, then its dual $\mathcal{R}^\circ$ is not a product trellis.*

PROOF: (a) If $\mathcal{R}$ is $(m-1)$-observable, then $\mathcal{R}^\circ$ is $(m-1)$-controllable, thus reduced by Corollary 5.5, thus a product trellis. (b) If $\mathcal{R}$ is observable but not $(m-1)$-observable, then $\mathcal{R}^\circ$ is controllable but not $(m-1)$-controllable, thus not branch-trim by Corollary 5.4, thus not a product trellis. □

For example, the observable trellis of Fig. 1(a) is not t-observable for any $t \leq m$; its dual in Fig. 1(b) is neither state-trim nor branch-trim. For another example, the observable trellis of Fig. 3(a) is m-observable but not $(m-1)$-observable; its dual in Fig. 3(b) is not branch-trim.

We remark that Theorem 5.6(b) may be extended to unobservable proper trellises as follows. (By Theorem 3.1, the dual of an improper trellis is not trim, hence not reduced.) We have to redefine $[j, k]$-observability as follows. Given a trellis $\mathcal{R}$ with behavior $\mathfrak{B}$, the *unobservable state configuration space* is defined as $\mathcal{S}^{\mathrm{u}} = \mathfrak{B}_{:\mathcal{S}} = \{\mathbf{s} \in \mathcal{S} : (\mathbf{0}, \mathbf{s}) \in \mathfrak{B}\}$ [7]. Then $\mathcal{R}$ is called $[j, k]$-observable if the unobservable transition space $\mathcal{U}^{[j,k)}$ equals the projection $(\mathcal{S}^{\mathrm{u}})_{|\mathcal{S}_j \times \mathcal{S}_k}$; i.e., if $\mathcal{R}$ does not contain any valid $[j, k]$-paths with $\mathbf{a}^{[j,k)} = \mathbf{0}^{[j,k)}$ other than those that lie on valid unobservable trajectories $(\mathbf{0}, \mathbf{s}) \in \mathfrak{B}$. Then we can show that if $\mathcal{R}$ is proper but not $(m-1)$-observable in this sense, then $\mathcal{R}^\circ$ is not a product trellis.

6 Constructing Reductions

This section begins to establish the main results of this paper concerning whether a trellis $\mathcal{R}$ is $[k, j]$-reducible. First, we shall give sufficient conditions for the $[k, j]$-irreducibility of $\mathcal{R}$, i.e., for $\mathcal{R}$ not having any $[k, j]$-reduction other than itself, up to isomorphism. Second, when these conditions are not met, but another auxiliary condition is met, we will construct a strict and conservative $[k, j]$-reduction.

Due to Theorem 4.1 we may assume without loss of generality that $\mathcal{R}$ is TPOC.

Theorem 6.1 *Let $j \neq k \mod m$. Suppose that both $\mathcal{R}$ and $\mathcal{R}^\circ$ are TPOC and $[j, k]$-observable. Then both $\mathcal{R}$ and $\mathcal{R}^\circ$ are $[k, j]$-irreducible.*

PROOF: It suffices to consider $\mathcal{R}$. Let $\tilde{\mathcal{R}}$ be a $[k,j)$-reduction of $\mathcal{R}$. Without loss of generality we may assume that $\tilde{\mathcal{R}}$ is trim and proper at time $k+1,\ldots,j-1$ since otherwise we may reduce further. We must show that $\mathcal{R}$ and $\tilde{\mathcal{R}}$ are isomorphic.

Since $\tilde{\mathcal{R}}^{[j,k)} = \mathcal{R}^{[j,k)}$, the trellis $\tilde{\mathcal{R}}$ is $[j,k)$-observable and thus observable. Furthermore, $\tilde{\mathcal{R}}$ is $[j,k)$-controllable, and thus $[k,j)$-trim by Theorem 5.2. Using trimness and $[k,j)$-trimness of $\mathcal{R}$, we conclude that $\mathcal{R}$ is state-trim at times k and j. But then $\tilde{\mathcal{R}}^{[j,k)} = \mathcal{R}^{[j,k)}$ and $[j,k)$-observability imply that $\tilde{\mathcal{R}}$ is also state-trim, thus trim, at times k and j.

Thus it remains to show that the fragments $\mathcal{R}^{[k,j)}$ and $\tilde{\mathcal{R}}^{[k,j)}$ are isomorphic. In order to do so we show first $\mathcal{C}^{[k,j)} = \tilde{\mathcal{C}}^{[k,j)}$. Let $(s_k, \mathbf{a}^{[k,j)}, s_j) \in \mathcal{C}^{[k,j)}$. Then there is a path $(\mathbf{a}^{[k,j)}, \mathbf{s}^{[k,j)}) \in \mathfrak{B}^{[k,j)}$, and by $[k,j)$-trimness this path lies on a trajectory in $\mathcal{R}$, say $(\mathbf{a}, \mathbf{s}) \in \mathfrak{B}$. Observability of $\tilde{\mathcal{R}}$ implies that there is a unique trajectory $(\mathbf{a}, \tilde{\mathbf{s}})$ in the behavior $\tilde{\mathfrak{B}}$ of $\tilde{\mathcal{R}}$, and $\tilde{\mathcal{R}}^{[j,k)} = \mathcal{R}^{[j,k)}$ along with $[j,k)$-observability yields $(\mathbf{a}^{[j,k)}, \mathbf{s}^{[j,k]}) = (\mathbf{a}^{[j,k)}, \tilde{\mathbf{s}}^{[j,k]})$. Hence $(s_j, s_k) = (\tilde{s}_j, \tilde{s}_k)$, and this proves that $(s_k, \mathbf{a}^{[k,j)}, s_j) \in \tilde{\mathcal{C}}^{[k,j)}$. In the same way one concludes that $\tilde{\mathcal{C}}^{[k,j)} \subseteq \mathcal{C}^{[k,j)}$.

All of this shows that $\mathcal{R}^{[k,j)}$ and $\tilde{\mathcal{R}}^{[k,j)}$ are both trim and proper cycle-free trellis fragments that realize the same external behavior. By [7, Thm. 3] they are both minimal, and must be isomorphic. But then $\mathcal{R}$ and $\tilde{\mathcal{R}}$ are isomorphic, and this concludes the proof. □

We note in passing that Theorem 6.1 is also true in the case where $j = k \mod m$, in which it reproduces earlier results. On the one hand, if $\mathcal{R}$ and $\mathcal{R}^\circ$ are $[j, j+m)$-observable, hence $[j, j+m)$-controllable, then the proof of Corollary 5.3 implies that they are both state-trim at time j, and this may be regarded as $[j+m, j)$-irreducibility (recall the equality constraint $\mathcal{C}_=$). On the other hand, we have seen already that $\mathcal{R}$ and $\mathcal{R}^\circ$ are $[j+m, j)$-observable if and only if $\mathcal{S}_j = \{0\}$; in this case they are conventional TPOC trellises, hence minimal, and thus irreducible on any interval.

Next, we recall that, as we have already seen in Example 4.6 and Corollary 5.4, a controllable trellis that is not $(m-1)$-controllable is not branch-trim, and thus has a conservative 1-reduction consisting of branch-trimming some constraint code, as in Fig. 4(b). Dually, an observable trellis that is not $(m-1)$-observable has a non-conservative 1-reduction consisting of branch-expanding some constraint code, as in Fig. 4(a). But this "reduced" trellis must be unobservable, and hence may be state-trimmed so as to make the reduction both strict and conservative, as in Fig. 5(a). We record these observations as a lemma:

Lemma 6.2 *If a trellis $\mathcal{R}$ of length m is observable but not $(m-1)$-observable, then it has a strict and conservative 2-reduction. Its dual is a strict and conservative 2-reduction of $\mathcal{R}^\circ$.*

PROOF: The dual trellis $\mathcal{R}^\circ$ is controllable but not $(m-1)$-controllable, and thus by Corollary 5.4 is not branch-trim at some constraint code $(\mathcal{C}_i)^\perp$. Replacing $(\mathcal{C}_i)^\perp$ by a suitably branch-trimmed $(\tilde{\mathcal{C}}_i)^\perp$ reduces the dimension of $(\mathcal{C}_i)^\perp$ by one, without changing the realized code $\mathcal{C}^\perp$. Dually, replacing $\mathcal{C}_i$ by $\tilde{\mathcal{C}}_i$ expands $\mathcal{C}_i$ by one dimension, without changing the realized code $\mathcal{C}$. Denote the resulting realizations by $\tilde{\mathcal{R}}^\circ$ and $\tilde{\mathcal{R}}$, respectively.

By Theorem 3.2, $\tilde{\mathcal{R}}^\circ$ must be uncontrollable, since we have reduced the dimension of a constraint code without changing $\mathfrak{B}$ or $\mathcal{S}$. Hence $\tilde{\mathcal{R}}$ is unobservable. As shown in Remark 4.3, we can trim $\tilde{\mathcal{R}}$ at any state space without changing the code $\mathcal{C}$. Trimming the state space $\mathcal{S}_i$ reduces the dimensions of $\mathcal{C}_{i-1}$, $\mathcal{S}_i$, and $\tilde{\mathcal{C}}_i$ by one, thus achieving a strict and conservative 2-reduction. By construction, the dual reduction is also strict and conservative. □

We now generalize Lemma 6.2 to trellises that are not $(m-t)$-observable, provided that the

trellis satisfies a certain technical condition. Without loss of generality, we may assume that the trellis is not $[0, m-t)$-observable.

Theorem 6.3 *Let $2 \leq t \leq m - 1$. Let $\mathcal{R}$ be a TPOC trellis of length m that is not $[0, m - t)$-observable; i.e., $\mathcal{U}^{[0,m-t)}$ is nontrivial, so (by properness) there is some $s_0 \neq 0$, $s_{m-t} \neq 0$ such that there is an unobservable path from s_0 to s_{m-t} in $\mathcal{R}^{[0,m-t)}$. Suppose $\mathcal{R}$ satisfies one of the following two conditions:*

Condition A: In the fragment $\mathcal{R}^{[m-t,m-1)}$, there is no valid path from s_{m-t} to $0 \in \mathcal{S}_{m-1}$.

Condition A': In the fragment $\mathcal{R}^{[m-t+1,0)}$, there is no valid path from $0 \in \mathcal{S}_{m-t+1}$ to s_0.

Then $\mathcal{R}$ has a conservative t-reduction and a strict and conservative $(t + 1)$-reduction. For each of these reductions the dual is a reduction of the same type of the dual trellis $\mathcal{R}^\circ$.

The t-reduction stated in the theorem is the main step of the reduction process and leads immediately to the strict $(t + 1)$-reduction. Both parts will be used in Theorem 7.4.

We give a sketch of the proof and outline the reduction procedure. The details are carried out in Appendix B.

SKETCH OF PROOF: By assumption there exists an unobservable valid path from $s_0 \neq 0$ to $s_{m-t} \neq 0$ in $\mathcal{R}^{[0,m-t)}$. However, since $\mathcal{R}$ is observable, there can be no unobservable valid path from s_{m-t} to s_0 in the complementary fragment $\mathcal{R}^{[m-t,0)}$.

Step 1: We expand $\mathcal{R}^{[m-t,0)}$ so that it contains an unobservable path $(\mathbf{0}^{[m-t,0)}, \tilde{\mathbf{s}}^{[m-t,0]})$ from $\tilde{s}_{m-t} := s_{m-t}$ to $\tilde{s}_0 := s_0$ via a sequence of new states $\tilde{s}_i \notin \mathcal{S}_i$ for $i \in (m - t, 0)$. The internal state spaces and the constraint codes of $\mathcal{R}^{[m-t,0)}$ are expanded to

$$\mathcal{S}_i^+ = \mathcal{S}_i \oplus \langle \tilde{s}_i \rangle \text{ for } i \in (m - t, 0) \quad \text{and} \quad \mathcal{C}_i^+ = \mathcal{C}_i \oplus \langle (\tilde{s}_i, 0, \tilde{s}_{i+1}) \rangle \text{ for } i \in [m - t, 0).$$

It is straightforward to show that the internal behavior of the fragment $\mathcal{R}^{[m-t,0)}$ consequently expands to $(\mathfrak{B}^{[m-t,0)})^+ = \mathfrak{B}^{[m-t,0)} \oplus \langle (\mathbf{0}^{[m-t,0)}, \tilde{\mathbf{s}}^{[m-t,0]}) \rangle$.

By construction, the expanded trellis $\mathcal{R}^+$ is unobservable and has a valid trajectory $(\mathbf{0}, \tilde{\mathbf{s}})$ that passes through s_0 and s_{m-t}. It is straightforward to show that its behavior is $\mathfrak{B}^+ = \mathfrak{B} \oplus \langle (\mathbf{0}, \tilde{\mathbf{s}}) \rangle$, so it continues to realize the same code.

Step 2: Assume that $\mathcal{R}$ satisfies Condition A. Then this guarantees (see Appendix B) that we can find a strict subspace $\tilde{\mathcal{S}}_{m-1} \subset \mathcal{S}_{m-1}^+$ such that there is no valid path from s_{m-t} to any state in $\tilde{\mathcal{S}}_{m-1}$. We trim $\mathcal{S}_{m-1}^+$ to $\tilde{\mathcal{S}}_{m-1}$.

Step 3: By Condition A, the resulting trellis will not be trim at times $m - 2, m - 3, \ldots, m - t$, and so one can successively trim the subspaces $\mathcal{S}_{m-2}^+, \ldots, \mathcal{S}_{m-t+1}^+, \mathcal{S}_{m-t}$ by at least one dimension. This will also reduce the adjacent constraint codes $\mathcal{C}_{m-1}^+, \mathcal{C}_{m-2}^+, \ldots, \mathcal{C}_{m-t}^+, \mathcal{C}_{m-t-1}$ by at least one dimension. All of this results in a strict and conservative $[m - t - 1, 0)$-reduction whose dual is of the same form. Trimming only the state spaces $\mathcal{S}_{m-2}^+, \ldots, \mathcal{S}_{m-t+1}^+$ and the adjacent constraint codes $\mathcal{C}_{m-1}^+, \mathcal{C}_{m-2}^+, \ldots, \mathcal{C}_{m-t}^+$ results in the stated conservative t-reduction. □

We illustrate this procedure with an example.

Example 4.6' The trellis of Fig. 5(a) is TPOC but not $[0, 3)$-observable. It satisfies Condition A, in that there is no transition in $\mathcal{C}_4$ from $01 \in \mathcal{S}_3$ to $00 \in \mathcal{S}_4$. Figure 6 illustrates the initial expansion of $\mathcal{S}_4$ to $\mathcal{S}_4^+ = \langle 1, \tilde{s} \rangle$ and the corresponding expansion of the adjacent constraint codes, $\mathcal{C}_3$ and $\mathcal{C}_4$.
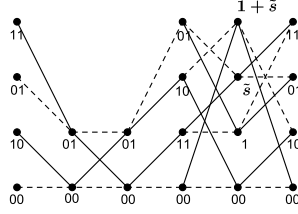
Figure 6: Expanding the trellis in Fig. 5(a)

Next, we identify $\tilde{\mathcal{S}}_4 = \langle 1 + \tilde{s} \rangle$ as a strict subspace of $\mathcal{S}_4^+$ such that there is no valid path from $01 \in \mathcal{S}_3$ to $\tilde{\mathcal{S}}_4$. Then $\mathcal{S}_4^+$ is trimmed to $\tilde{\mathcal{S}}_4$, as in Fig. 7, which also reduces $\mathcal{C}_3^+$ and $\mathcal{C}_4^+$.
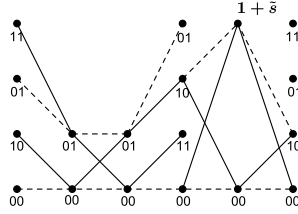

Figure 7: Trimming the trellis in Fig. 6

The resulting trellis is guaranteed not to be trim at $\mathcal{S}_3$, so $\mathcal{S}_3$ and $\mathcal{C}_2$ may be trimmed to achieve a strict and conservative 3-reduction. Moreover, in this example, if we trim all states and branches that are not on any valid trajectories, then we reach the conventional (hence minimal) trellis shown in Fig. 8.
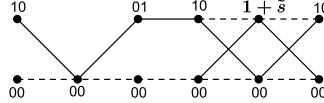

Figure 8: Trimming the trellis in Fig. 7 □

7 t-Irreducibility

Now we are in a position to discuss t-irreducibility for all values of t not larger than a certain invariant of the code. We will see that it is characterized by $(m - t)$-observability and $(m - t)$-controllability. Thereafter we will illustrate the open problems that arise for larger values of t. Due to Theorem 4.1 we may restrict attention to TPOC trellises.

We begin with 1-reductions, that is, the replacement of one constraint code. The following intrinsic characterization of 1-irreducibility shows in particular that every 1-reduction is either a branch-trimming (replacement of a constraint code by a strict subcode) or a branch-addition (replacement of a constraint code by a strict supercode).

Theorem 7.1 *Let $\mathcal{R}$ be TPOC. Then the following are equivalent.*

(i) $\mathcal{R}$ and $\mathcal{R}^\circ$ are 1-irreducible,

(ii) $\mathcal{R}$ and $\mathcal{R}^\circ$ are branch-trim,

(iii) $\mathcal{R}$ and $\mathcal{R}^\circ$ are $(m-1)$-observable and $(m-1)$-controllable.

Furthermore, if $\mathcal{R}$ is 1-reducible, then it has a strict and conservative 2-reduction $\tilde{\mathcal{R}}$ whose dual $\tilde{\mathcal{R}}^\circ$ is a strict and conservative 2-reduction of $\mathcal{R}^\circ$.

PROOF: (i) $\Rightarrow$ (ii) is obvious. (ii) $\Rightarrow$ (iii) follows from Corollary 5.4, and (iii) $\Rightarrow$ (i) is a special case of Theorem 6.1. For the last statement, let $\mathcal{R}$ be 1-reducible. Then either the trellis or its dual is $(m-1)$-unobservable and the result follows from Lemma 6.2. $\qquad\square$

An example for the reduction of an $(m-1)$-unobservable trellis has been shown already in Example 4.6.

Recall that minimality of a trellis is defined on the basis of its state space dimensions. From the above we obtain the following corollary, which tells us that the constraint codes of a minimal trellis cannot be replaced by smaller (or larger) constraint codes. In particular, minimal trellises are branch-trim, which has also recently been established by different arguments in [3, Thm. 7.1]. One may recall that in [13] Koetter/Vardy restrict themselves to state-trim and branch-trim trellises, so that in their terminology minimal trellises are branch-trim by definition.

Corollary 7.2 *A minimal trellis is 1-irreducible.*

We now consider t-irreducibility for $t > 1$. We first introduce the following parameter of a code $\mathcal{C}$, which derives from the "characteristic spans" of Koetter/Vardy [13] (see also [9, 8]).

Definition 7.3 *The minimum span length $\chi(\mathcal{C})$ of a code $\mathcal{C} \subseteq \mathcal{A} = \prod_i \mathcal{A}_i$ is the minimum length of all possible spans of the nonzero codewords $\mathbf{a} \in \mathcal{C}$, where a span of $\mathbf{a} \neq \mathbf{0}$ is any interval, possibly circular, that covers the support of $\mathbf{a}$.*

For example, the code $\mathcal{C} = \{00000, 10110, 11001, 01111\} \subset \mathbb{F}_2^5$ has minimum span length $\chi(\mathcal{C}) = 3$.

Now we can formulate the following characterization of t-irreducibility. The proof provides us with a constructive reduction method for t-reducible trellises.

Theorem 7.4 *Let $\min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\} > t > 1$ and let $\mathcal{R}$ be a TPOC trellis of $\mathcal{C}$. Then the following are equivalent.*

(i) $\mathcal{R}$ (and thus $\mathcal{R}^\circ$) is $(m - t)$-observable and $(m - t)$-controllable.

(ii) $\mathcal{R}$ (and thus $\mathcal{R}^\circ$) is t-irreducible.

For a t-reducible trellis $\mathcal{R}$ we have the following cases:
if $\mathcal{R}$ is $(m-t+1)$-unobservable or $(m-t+1)$-uncontrollable then $\mathcal{R}$ allows a strict and conservative t-reduction. If $\mathcal{R}$ is $(m-t+1)$-observable and $(m-t+1)$-controllable, then $\mathcal{R}$ allows non-strict and conservative t-reduction which gives rise to a subsequent strict and conservative $(t + 1)$-reduction. In either case, the dual process is a reduction of the same type for $\mathcal{R}^\circ$.

PROOF: (i) $\Rightarrow$ (ii) is Theorem 6.1. As for the converse assume without loss of generality that $\mathcal{R}^{[0,m-t)}$ is unobservable, and let s_0 and s_{m-t} be the end states of a nontrivial unobservable valid path of this

fragment. Since $\mathcal{R}$ is t-irreducible, Theorem 6.3 tells us that there must exist valid paths from s_{m-t} to $0 \in \mathcal{S}_{m-1}$ in the fragment $\mathcal{R}^{[m-t,m-1)}$ and from $0 \in \mathcal{S}_{m-t+1}$ to s_0 in $\mathcal{R}^{[m-t+1,0)}$. Adding those two paths (suitably appended by zero branches) results in a valid path from s_{m-t} to s_0. But this means that the given unobservable path lies on a valid trajectory in $\mathcal{R}$. Hence it represents a codeword in $\mathcal{C}$ that has a span of length at most t, and this contradicts our assumption.

From Theorem 6.3 it is clear that every t-reducible trellis allows a non-strict conservative t-reduction that is non-trim and thus gives rise to a subsequent strict and conservative $(t+1)$-reduction. Suppose now that $\mathcal{R}$ is $(m-t+1)$-unobservable. Hence there exists an unobservable valid path of length $m-t+1$, say in $\mathcal{R}^{[0,m-t+1)}$. With the aid of $\chi(\mathcal{C}) > t$, we conclude that this path cannot lie on a valid trajectory, which in turn means that Condition A or A′ of Theorem 6.3 must be satisfied. Consequently, by that theorem there exists a conservative $(t-1)$-reduction and a strict and conservative t-reduction, whose duals are reductions of the same type. □

We remark that Theorem 7.4 applies to conventional trellises as follows. A trim and proper conventional trellis $\mathcal{R}$ and its dual $\mathcal{R}^\circ$ are minimal and therefore t-irreducible for all t. Hence by Theorem 7.4, $\mathcal{R}$ and $\mathcal{R}^\circ$ must be $(m-t)$-observable for all $t < \min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\}$. But this is obvious, since in a minimal conventional trellis an unobservable path of length $m-t$ would imply a nonzero codeword with circular span length t or less, since all valid paths lie on valid trajectories. In this sense, the limit $t < \min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\}$ in Theorem 7.4 is the best possible.

Let us now return to the general situation. Theorem 7.4 characterizes t-irreducibility for small values of t. For $t \geq \min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\}$ the implication (i) $\Rightarrow$ (ii) remains valid (being solely based on Theorem 6.1), whereas the converse is not true in general. For instance, the trellis in Fig. 7 is conventional and minimal, thus t-irreducible for each t, but it is not 2-observable.

The next example illustrates that in some cases, t-reducibility for $t \geq \min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\}$ may follow directly from Theorem 6.3.

Example 7.5 The trellis $\mathcal{R}$ in Fig. 9 is the product trellis obtained from the generators $\underline{110110}000$, $0\underline{10100}000$, $000\underline{11010}$, $\underline{1}00000\underline{011}$, $0\underline{11}00000\underline{1}$, $\underline{11}0001\underline{101}$ with the indicated spans. The state labels are suppressed, and it suffices to keep in mind that the states at time zero appear in the same ordering at the beginning and end of the trellis.
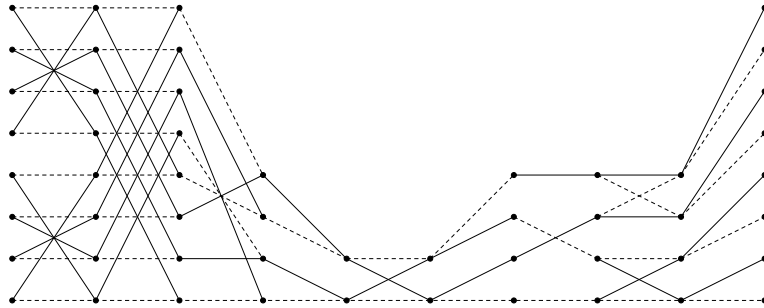

Figure 9: 2-irreducible and $(m-3)$-unobservable trellis

The code $\mathcal{C} \subseteq \mathbb{F}_2^9$ generated by these vectors satisfies $\chi(\mathcal{C}) = 3$ and $\chi(\mathcal{C}^\perp) = 6$. One can see directly that the trellis is $(m-2)$-observable, and it is also not hard to check that it is $(m-2)$-controllable.

Thus, by Theorem 7.4 the trellis is 2-irreducible. But the trellis is $(m-3)$-unobservable. Even more, the unobservable valid path in the fragment $\mathcal{R}^{[0,6)}$ satisfies Condition A of Theorem 6.3: there is no valid path from the ending state at time 6 to the zero state at time 8. As a consequence, the trellis is reducible on the interval $[6,0)$. Performing the reduction procedure as in the proof of Theorem 6.3, we obtain the trellis shown in Figure 10 (where we added suitable state labels for further referencing).
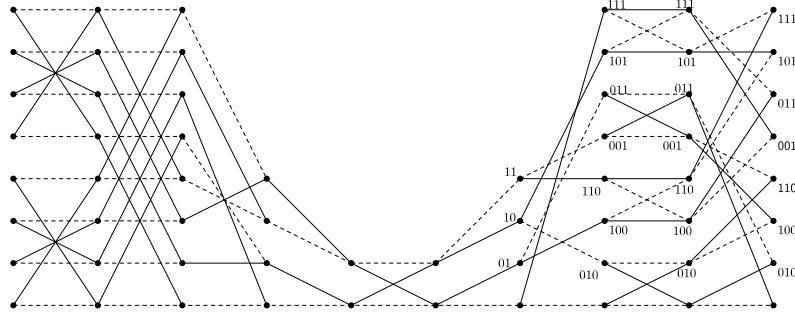

Figure 10: Step 1 of the reduction

The (unique) subspace $\mathcal{X}$ of $\mathcal{S}_8^+$ as in Step 2 of the reduction process is $\mathcal{X} = \langle 010, 111 \rangle \subseteq \mathcal{S}_8^+$. Trimming $\mathcal{S}_8^+$ to $\mathcal{X}$ leads to a trellis that is not trim at time 6 because there is no valid path from $11 \in \mathcal{S}_6$ to any state in $\mathcal{X}$. Thus, as stated in Steps 3 and 4, we can subsequently trim the state spaces $\mathcal{S}_7^+$ and $\mathcal{S}_6$. In this case, trimming $\mathcal{S}_7^+$ to $\langle 010, 111 \rangle$ and trimming $\mathcal{S}_6$ to $\langle 10 \rangle$ results in the trellis shown in Figure 11. This trellis, denoted by $\tilde{\mathcal{R}}$, coincides with $\mathcal{R}$ on $[0,5)$ and is thus
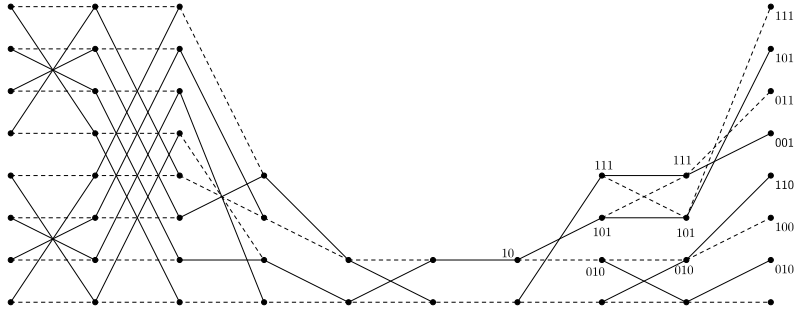

Figure 11: Steps 3 and 4 of the reduction leading to a strict 4-reduction

a strict and conservative 4-reduction. The trellis is $(m-2)$-observable and $(m-2)$-controllable. Furthermore, it can easily be checked that $\tilde{\mathcal{R}}$ and $\tilde{\mathcal{R}}^\circ$ do not contain any unobservable valid paths satisfying Conditions A or A′ of Theorem 6.3, and thus we do not have any other reduction method at our disposal. This is not surprising because it can be shown (with the aid of the class of KV-trellises as introduced by Koetter and Vardy in [13]; see also Section 8), that $\tilde{\mathcal{R}}$ is a minimal trellis and every trellis for the same code with state spaces of at most the same sizes is isomorphic to $\tilde{\mathcal{R}}$. Thus $\tilde{\mathcal{R}}$ is t-irreducible for every t. □

Example 7.6 The trellis in Fig. 12(a) is the product trellis obtained from the generators $\underline{101100}$, $00\underline{1101}$, $\underline{011011}$. Its dual is shown in (b).
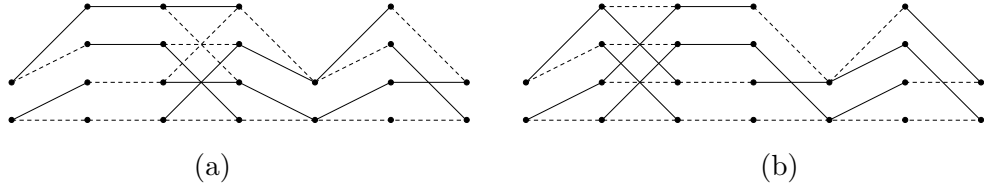
20

Figure 12: Dual pair of nonminimal trellises

The code $\mathcal{C}$ generated by these vectors contains the word 100001 and thus satisfies $\chi(\mathcal{C}) = 2$, whereas $\chi(\mathcal{C}^\perp) = 3$. Hence Theorem 7.4 is not applicable. Furthermore, it is easy to see that the trellis and its dual do not contain any unobservable valid paths satisfying Condition A or A' from Theorem 6.3, and thus none of our reduction methods applies. Yet, in this case the trellises are reducible. Indeed, Fig. 13 shows mutually dual strict reductions on the interval $[4, 3)$.
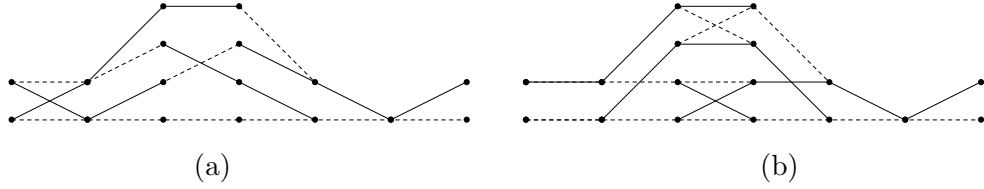

Figure 13: Dual pair of $(m-1)$-reductions

These trellises are conventional, and thus minimal and t-irreducible for every t. Note that the trellis in Fig. 13(a) is a conservative reduction of Fig. 12(a), whereas the dual reduction is a non-conservative reduction of Fig. 12(b).

The trellises in Fig. 13 can be obtained constructively from those in Fig. 12: applying twice and in a suitable way the reduction from the proof of Theorem 6.3 to Fig. 12(a) will result in Fig. 13(a). Even though Theorem 6.3 does not apply, this will eventually lead to the reduction on the interval $[4, 3)$. With the aid of the class of KV-trellises, one can show that the trellis in Fig. 13(a) is the only trellis for $\mathcal{C}$ that is strictly smaller than that in Fig. 12(a). Therefore, it is also the only possible reduction. By duality, the same applies to the dual trellises in Fig. 13(b) and Fig. 12(b). Hence we conclude that there exist strictly t-reducible trellises that do not allow a conservative reduction. □

As the last two examples have illustrated, it remains an open problem how to characterize t-irreducibility for $t \geq \ell := \min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\}$. By Theorem 6.1 we can expect t-reducibility only if there exists an unobservable fragment of length $m-t$ in the trellis or its dual. If such an unobservable valid path satisfies Condition A or A' from Theorem 6.3, then the reduction procedure from the proof of this theorem is applicable and the trellis is reducible. As a consequence, the only remaining case is where the unobservable valid path does not satisfy either of the conditions. As shown in the proof of Theorem 7.4 this means that the path lies on a valid trajectory (which, for instance, is the case in any conventional trellis). Obviously, the support of the associated codeword is contained in an interval of length t. Summarizing, one needs to study unobservable valid paths of length $\leq m - \ell$ that lie on valid trajectories. Supported by many examples, we formulate the following conjecture.

Conjecture 7.7 *Any nonminimal trellis can be reduced constructively using a finite number of steps as in the reduction procedure from the proof of Theorem 6.3 along with suitable state trimmings applied to the trellis or its dual.*

8   Comparison to KV-Trellises

In this section we relate our results to previous work on trellises and their complexity. To this end, we restrict ourselves to trellises with symbol spaces $\mathcal{A}_i = \mathbb{F}$ for all i and to codes $\mathcal{C}$ in $\mathbb{F}^m$ such that both $\mathcal{C}$ and $\mathcal{C}^\perp$ have full support, that is, the codewords do not all vanish at a fixed coordinate; in other words, $\min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\} > 1$.

Koetter and Vardy [13] showed that the search for possibly minimal trellises can be narrowed to a certain canonical class, which we call *KV-trellises*. A KV-trellis is a product trellis based on $\dim \mathcal{C}$ linearly independent generators with shortest spans that all start and end at different positions. A span $[\alpha, \alpha + r]$ (in the circular sense) is called a *shortest span of the code* if r is the smallest length of the spans of all codewords (in the sense of Definition 7.3) that are nonzero at α; see also [9, 8]. KV-trellises may be regarded as the tail-biting version of the realizations resulting from the "shortest basis" approach in [6].

Being product trellises, KV-trellises are state-trim and branch-trim, and from the choice of the generators it follows that they are proper, observable and controllable. Moreover, in [13, Thm. 5.5] Koetter and Vardy have shown that each (reduced) minimal trellis is a KV-trellis. However, the converse is not true: not all KV-trellises are minimal.

In [8, Thm. IV.3] it is proven that the dual of a KV-trellis is a KV-trellis of the dual code. As a consequence, Theorem 5.6 implies that KV-trellises are $(m-1)$-observable and $(m-1)$-controllable, and thus 1-irreducible due to Theorem 7.1. With the machinery developed in [8] (more precisely, by a generalization of Theorem II.13 in [8]), one can show that if a code $\mathcal{C}$ satisfies $\min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\} > t$, then all its KV-trellises (and their duals) are $(m-t)$-observable and hence t-irreducible by Theorem 7.4. However, this property does not characterize KV-trellises. Indeed, the trellis in Fig. 9 (and its dual) is not a KV-trellis[4], but 2-irreducible and represents a code satisfying $\min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\} = 3$.

Summarizing, for a given t and a code $\mathcal{C}$ such that $\min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\} > t > 1$ we have the following proper containments of trellis classes

$$\underbrace{\Big\{\text{minimal}\Big\} \subsetneq \Big\{\text{KV}\Big\} \subsetneq \left\{\begin{matrix} t\text{-irreducible,} \\ \text{T}_{\text{sb}}\text{POC} \end{matrix}\right\} = \left\{\begin{matrix} (m-t)\text{-obs./contr,} \\ \text{T}_{\text{sb}}\text{POC} \end{matrix}\right\}}_{\text{invariant under dualization}} \subsetneq \underbrace{\Big\{\text{NT}_{\text{sb}}\text{POC}\Big\} \subsetneq \Big\{\text{T}_{\text{sb}}\text{POC}\Big\}}_{\text{not invariant under dualization}},$$

where N stands for nonmergeable, T_{sb} for state- and branch-trim (as opposed to the weaker trimness), and, as before, P,O,C stand for proper, observable, and controllable, respectively.

As indicated, the last two containments are strict as well: for instance, the trellis in Figure 1(a) is in the rightmost class, but not in $\{\text{NT}_{\text{sb}}\text{POC}\}$; moreover, it is straightforward to verify that the product trellis generated by $00\underline{10100}$, $000\underline{1011}$, $1000\underline{100}$, $\underline{0110001}$ realizes a code $\mathcal{C}$ that satisfies $\min\{\chi(\mathcal{C}), \chi(\mathcal{C}^\perp)\} = 3$ and is in $\{\text{NT}_{\text{sb}}\text{POC}\}$, but 2-reducible.

---

[4]The sum of the first four generators given in Example 7.5 results in a codeword with span $[5, 8]$. Since this span is shorter than the span $[5, 1]$ of the last generator, the latter is not a shortest span of the code.

The three leftmost classes are invariant under taking duals. This is clear for minimal trellises and for t-irreducible trellises and follows from [8, Thm. IV.3] for KV-trellises. Fig. 3 and Fig. 1 show that the two rightmost classes are not invariant under taking duals.

9   Conclusion

We have presented constructive procedures for reducing a given tail-biting trellis realization and its dual and have provided criteria for when a trellis is irreducible on an interval of length t. The criteria are sufficient for all codes, and necessary and sufficient for codes of minimum span length bigger than t. We have also discussed the remaining case of reducibility on intervals of length at least the minimum span length of the code. While we believe that all nonminimal trellises can be reduced, finding a constructive procedure remains a largely open problem.

As a main tool of our approach, we have used trellis fragments, i.e., realizations obtained by cutting two edges in the normal graph of the tail-biting trellis. We have introduced the notions of fragment controllability and observability, which are, naturally, the same as those in classical linear systems theory, and have shown that they are mutually dual.

With the aid of fragment trimness, which implies that every valid path in the fragment is the restriction of a valid trajectory in the entire trellis, we have presented criteria for state-trimness and branch-trimness of a tail-biting trellis. Using the well-known fact that a tail-biting trellis is a product trellis if and only if it is state-trim and branch-trim, our results have also led to a characterization of when the dual of a product trellis is a product trellis.

Finally, we have discussed the relation of our results to the prior tail-biting trellis literature that relies on product representations.

Beyond trellises, we believe that many of our results can be generalized to normal realizations on general graphs.

A   Controllability and Connectedness

In this appendix we discuss the relationship between controllability of a trellis realization and connectedness of the trajectories in its trellis diagram. In [7, Thm. 10] it was shown that if a trellis is trim but not controllable, then its valid trajectories partition into disconnected subsets. Moreover, it was noted in [7] that, using the product representation of [13], a state-trim and branch-trim trellis is uncontrollable if and only if its trajectories are disconnected. We will now show that this statement holds if the trellis is merely state-trim.

Theorem A.1 *A state-trim trellis is controllable if and only if it is connected.*

PROOF: The if-part has been proven (for trim trellises) in [7, Thm. 10]. For state-trim and branch-trim trellises a proof of the converse is sketched in [7, Sec. IV.F] by using the fact that every such trellis is a product realization of one-dimensional trellises. For non-branch-trim trellises a proof of the converse is as follows. Let $\mathcal{R}$ be a controllable, state-trim trellis and suppose $\mathcal{R}$ is not connected. Consider the connected component of $\mathcal{R}$ containing the zero trajectory. Denote the state and constraint sets of this subtrellis by $\mathcal{S}_{i,0}$ and $\mathcal{C}_{i,0}$, respectively. With the aid of state-trimness one easily verifies that $\mathcal{C}_{i,0}$ is a linear subspace of $\mathcal{C}_i$ for each i. Notice that $\mathcal{S}_{i,0}$ is the

union of the projections of $\mathcal{C}_{i,0}$ and $\mathcal{C}_{i-1,0}$ on $\mathcal{S}_i$. Again using state-trimness one can see that $\mathcal{S}_{i,0}$ is a linear subspace of $\mathcal{S}_i$ for all i. Thus, the connected component forms a linear subtrellis $\mathcal{R}_0$ with state spaces $\mathcal{S}_{i,0}$ and constraint codes $\mathcal{C}_{i,0}$. Denote its behavior by $\mathfrak{B}_0$. Then the behavior $\mathfrak{B}$ of $\mathcal{R}$ is the union of q^ℓ disconnected subbehaviors, where $\ell := \dim \mathfrak{B} - \dim \mathfrak{B}_0$, and by linearity and state-trimness each state space $\mathcal{S}_i$ and constraint code $\mathcal{C}_i$ of $\mathcal{R}$ is the union of q^ℓ cosets of $\mathcal{S}_{i,0}$ and $\mathcal{C}_{i,0}$, respectively. Using controllability of $\mathcal{R}$ we obtain $\dim \mathfrak{B}_0 < \dim \mathfrak{B} = \sum_i(\dim \mathcal{C}_i - \dim \mathcal{S}_i) = \sum_i(\dim \mathcal{C}_{i,0} - \dim \mathcal{S}_{i,0})$, and this contradicts Theorem 3.2 for the trellis $\mathcal{R}_0$. $\qquad\square$

We wish to point out that state-trimness is indeed necessary for the only-if part to be true. The two linear, non-state-trim trellises shown in Fig. 14 are disconnected, yet form controllable realizations of the code $\mathcal{C} = \{00\}$. It is also worth observing that in the trellis in Fig. 14(b) the connected component containing the zero trajectory is not a linear subtrellis.

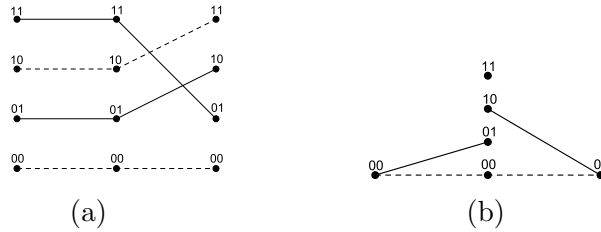(a)   (b)

Figure 14: Disconnected and controllable trellises

Finally, we note that [7, Sec. IV.G] shows that Theorem A.1 does not generalize to normal linear realizations on general graphs with cycles.

B   Proof of Theorem 6.3

Without loss of generality we assume that $\mathcal{R}$ satisfies Condition A. In a first step we expand the trellis $\mathcal{R}$ to an unobservable trellis by increasing the state spaces at times $m - t + 1, \ldots, m - 1$ by one dimension. Thus, pick new states $\tilde{s}_i \notin \mathcal{S}_i$ for $i = m - t + 1, \ldots, m - 1$. For ease of notation, define $\tilde{s}_{m-t} := s_{m-t}$, $\tilde{s}_0 := s_0$, and expand the fragment $\mathcal{R}^{[m-t,0)}$ via

$$\mathcal{S}_i^+ = \mathcal{S}_i \oplus \langle \tilde{s}_i \rangle \text{ for } i \in (m - t, 0) \ \text{ and } \ \mathcal{C}_i^+ = \mathcal{C}_i \oplus \langle(\tilde{s}_i, 0, \tilde{s}_{i+1})\rangle \text{ for } i \in [m - t, 0). \qquad (\text{B.1})$$

By construction, the expanded trellis $\mathcal{R}^+$ has a nontrivial unobservable valid trajectory $(\mathbf{0}, \tilde{\mathbf{s}})$ that passes through the states $\tilde{s}_i$, $i \in [m - t, 0]$.

Using the fact that $\tilde{s}_i \notin \mathcal{S}_i$ for $i \in (m - t, 0)$, we obtain immediately for $i \in \{m - t + 1, \ldots, m - 2\}$

$$(v_i + \alpha \tilde{s}_i, a_i, v_{i+1} + \beta \tilde{s}_{i+1}) \in \mathcal{C}_i^+ \text{ for some } (v_i, a_i, v_{i+1}) \in \mathcal{S}_i \times \mathcal{A}_i \times \mathcal{S}_{i+1}, \ \alpha, \beta \in \mathbb{F} \implies \alpha = \beta. \ (\text{B.2})$$

The trellis $\mathcal{R}^+$ has the following properties.

1) The behavior of $\mathcal{R}^+$ is given by $\mathfrak{B}^+ = \mathfrak{B} \oplus \langle(\mathbf{0}, \tilde{\mathbf{s}})\rangle$, and thus $\mathcal{R}^+$ represents the same code $\mathcal{C}$. To see this, consider a valid trajectory in $\mathfrak{B}^+$. Due to (B.2), the state sequence must be of the form $(v_0, \ldots, v_{m-t}, v_{m-t+1} + \alpha \tilde{s}_{m-t+1}, \ldots, v_{m-1} + \alpha \tilde{s}_{m-1}, v_0)$, where $v_i \in \mathcal{S}_i$ for all i. Subtracting the unobservable trajectory $\alpha(\mathbf{0}, \tilde{\mathbf{s}})$ yields a valid trajectory that is entirely in the subtrellis $\mathcal{R}$, hence it is an element of $\mathfrak{B}$.

24

2) There is no valid path from $\tilde{s}_{m-t} \in \mathcal{S}_{m-t}$ to $0 \in \mathcal{S}_{m-1}^+$ in the fragment $(\mathcal{R}^+)^{[m-t,m-1)}$. To show this, suppose we have such a path. By (B.2) its state sequence is of the form $(\tilde{s}_{m-t}, v_{m-t+1} + \alpha \tilde{s}_{m-t+1}, \ldots, v_{m-2} + \alpha \tilde{s}_{m-2}, 0)$, where $v_i \in \mathcal{S}_i$, and once more by (B.2) we conclude that $\alpha = 0$. But then the given path is a valid path in $\mathcal{R}^{[m-t,m-1)}$, and this contradicts Condition A.

3) There exists a subspace $\mathcal{X}$ of $\mathcal{S}_{m-1}^+$ satisfying $\mathcal{X} \oplus \langle \tilde{s}_{m-1} \rangle = \mathcal{S}_{m-1}^+$ and such that there is no valid path in $(\mathcal{R}^+)^{[m-t,m-1)}$ from $\tilde{s}_{m-t}$ to any $x \in \mathcal{X}$. This can be seen as follows. Put

$$\mathcal{Y} := \{s \in \mathcal{S}_{m-1}^+ \mid \text{there is a valid path from } 0 \in \mathcal{S}_{m-t} \text{ to } s \text{ in the fragment } (\mathcal{R}^+)^{[m-t,m-1)}\}.$$

Then $\tilde{s}_{m-1} \notin \mathcal{Y}$ because if there was a valid path from 0 to $\tilde{s}_{m-1}$, then the existence of a path with state sequence $(\tilde{s}_{m-t}, \tilde{s}_{m-t+1}, \tilde{s}_{m-t+2}, \ldots, \tilde{s}_{m-1})$ leads to a valid path in $(\mathcal{R}^+)^{[m-t,m-1)}$ from $\tilde{s}_{m-t}$ to $0 \in \mathcal{S}_{m-1}^+$, and this contradicts our observation in 2). Using again the path from $\tilde{s}_{m-t}$ to $\tilde{s}_{m-1}$ we observe that for each $\alpha \in \mathbb{F}$ the coset $\mathcal{Y} + \alpha \tilde{s}_{m-1}$ is exactly the set of all states in $\mathcal{S}_{m-1}^+$ that can be reached by a valid path from $\alpha \tilde{s}_{m-t}$. Now, we may choose any subspace $\mathcal{Z} \subset \mathcal{S}_{m-1}^+$ such that $\mathcal{Z} \oplus \mathcal{Y} \oplus \langle \tilde{s}_{m-1} \rangle = \mathcal{S}_{m-1}^+$ and put $\mathcal{X} = \mathcal{Z} \oplus \mathcal{Y}$.

Having established these properties we can perform the reduction:

a) Trim $\mathcal{S}_{m-1}^+$ to the subspace $\tilde{\mathcal{S}}_{m-1} := \mathcal{X}$, where $\mathcal{X}$ is as in 3). By Remark 4.3 this results in a trellis that still represents $\mathcal{C}$. Denote its constraint code at time $m-2$ by $\tilde{\mathcal{C}}_{m-2}$.

b) Let $\tilde{\mathcal{S}}_{m-2}$ be the projection of $\tilde{\mathcal{C}}_{m-2}$ on the state space $\mathcal{S}_{m-2}^+$. Then $\tilde{\mathcal{S}}_{m-2}$ is contained in the set $\{s \in \mathcal{S}_{m-2}^+ \mid \text{there is no valid path from } s_{m-t} \text{ to } s\}$, and the latter is a proper subset of $\mathcal{S}_{m-2}^+$. Thus, $\dim \tilde{\mathcal{S}}_{m-2} < \dim \mathcal{S}_{m-2}^+$. Obviously, the states not in $\tilde{\mathcal{S}}_{m-2}$ are not on any valid trajectory, and thus we may trim $\mathcal{S}_{m-2}^+$ to $\tilde{\mathcal{S}}_{m-2}$. After this trimming denote the constraint code at time $m-3$ by $\tilde{\mathcal{C}}_{m-3}$ and continue in the same manner.

c) All this shows that we can trim all state spaces $\mathcal{S}_{m-1}^+, \mathcal{S}_{m-2}^+, \ldots, \mathcal{S}_{m-t+1}^+$ by one dimension. Since the branches that have been added in (B.1) will be trimmed, this also reduces the constraint codes $\mathcal{C}_i^+$, $i = m-t, \ldots, m-1$, by one dimension. Thus the resulting trellis, denoted by $\tilde{\mathcal{R}}$, is an $[m-t, 0)$-reduction with the same state space and constraint code dimensions as $\mathcal{R}$. Consequently, the same is true for the dual reduction $\tilde{\mathcal{R}}^\circ$ of $\mathcal{R}^\circ$. Finally, by construction, $\tilde{s}_{m-t} \in \mathcal{S}_{m-t}$ is not on any branch in the constraint code $\tilde{\mathcal{C}}_{m-t}$, and thus $\tilde{\mathcal{R}}$ is not trim at time $m-t$. Trimming results in the desired strict and conservative $[m-t-1, 0)$-reduction of $\mathcal{R}$, whose dual is a reduction of $\mathcal{R}^\circ$ of the same type. $\qquad \square$

References

[1] A. Al-Bashabsheh and Y. Mao. Normal factor graphs and holographic transformations. *IEEE Trans. Inform. Theory*, IT-57:752–763, 2011.

[2] A. R. Calderbank, G. D. Forney, Jr., and A. Vardy. Minimal tail-biting trellises: The Golay code and more. *IEEE Trans. Inform. Theory*, IT-45:1435–1455, 1999.

[3] D. Conti. *An Algebraic Development of Trellis Theory*. PhD thesis, University College Dublin, 2012.

[4] G. D. Forney, Jr. Codes on graphs: Normal realizations. *IEEE Trans. Inform. Theory*, IT-47:520–548, 2001.

[5] G. D. Forney, Jr. Codes on graphs: Duality and MacWilliams identities. *IEEE Trans. Inform. Theory*, IT-57:1382–1397, 2011.

[6] G. D. Forney, Jr. Minimal realizations of linear systems: The "shortest basis" approach. *IEEE Trans. Inform. Theory*, IT-57:726–737, 2011.

[7] G. D. Forney, Jr. and H. Gluesing-Luerssen. Codes on graphs: Observability, controllability and local reducibility. Preprint 2012. Accepted for publication in *IEEE Trans. Inform. Theory*. ArXiv: cs.IT/1203.3115v2.

[8] H. Gluesing-Luerssen and E. Weaver. Characteristic generators and dualization for tail-biting trellises. *IEEE Trans. Inform. Theory*, IT-57:7418–7430, 2011.

[9] H. Gluesing-Luerssen and E. Weaver. Linear tail-biting trellises: Characteristic generators and the BCJR-construction. *IEEE Trans. Inform. Theory*, IT-57:738–751, 2011.

[10] N. Kashyap. On minimal tree realizations of linear codes. *IEEE Trans. Inform. Theory*, IT-55:3501–3519, 2009.

[11] R. Koetter. On the representation of codes in Forney graphs. In *Codes, Graphs, and Systems: A celebration of the life and career of G. David Forney, Jr.* (R. E. Blahut and R. Koetter, eds.), pages 425–450. Kluwer Academic Publishers, 2002.

[12] R. Koetter and A. Vardy. On the theory of linear trellises. In *Information, Coding and Mathematics* (M. Blaum and P. G. Farrell and H. C. A. van Tilborg, eds.), pages 323–354. Kluwer, Boston, 2002.

[13] R. Koetter and A. Vardy. The structure of tail-biting trellises: Minimality and basic principles. *IEEE Trans. Inform. Theory*, IT-49:2081–2105, 2003.

[14] F. R. Kschischang and V. Sorokine. On the trellis structure of block codes. *IEEE Trans. Inform. Theory*, IT-41:1924–1937, 1995.

[15] S. Lin and R. Y. Shao. General structure and construction of tail-biting trellises for linear block codes. In *Proceedings of the 2000 IEEE International Symposium on Information Theory*, page 117, 2000.

[16] Y. Mao and F. R. Kschischang. On factor graphs and the Fourier transform. *IEEE Trans. Inform. Theory*, IT-51:1635–1649, 2005.

[17] A. V. Nori and P. Shankar. Unifying views of tail-biting trellis constructions for linear block codes. *IEEE Trans. Inform. Theory*, IT-52:4431–4443, 2006.

[18] Y. Shany and Y. Be'ery. Linear tail-biting trellises, the square root bound, and applications for Reed-Muller codes. *IEEE Trans. Inform. Theory*, IT-46:1514–1523, 2000.

[19] A. Vardy. Trellis structure of codes. In *Handbook of Coding Theory, Vol. 2* (V. S. Pless and W. C. Huffman, eds.), pages 1989–2117. Elsevier, Amsterdam, 1998.

[20] A. Vardy and F. R. Kschischang. Proof of a conjecture of McEliece regrading the expansion index of the minimal trellis. *IEEE Trans. Inform. Theory*, IT-42:2027–2034, 1996.