

# On the support of extremal martingale measures with given marginals: the countable case\*

Luciano Campi<sup>†</sup> Claude Martini<sup>‡</sup>

## Abstract

We investigate the supports of extremal martingale measures with pre-specified marginals in a two-period setting. First, we establish in full generality the equivalence between the extremality of a given measure  $Q$  and the denseness in  $L^1(Q)$  of a suitable linear subspace, which can be seen as the set of all semi-static trading strategies. Moreover, when the supports of both marginals are countable, we focus on the slightly stronger notion of weak exact predictable representation property (henceforth, WEP) and provide two combinatorial sufficient conditions, called “2-link property” and “full erasability”, on how the points in the supports are linked to each other for granting extremality. When the support of the first marginal is a finite set, we give a necessary and sufficient condition for the WEP to hold in terms of the new concept of 2-net. Finally, we study the relation between cycles and extremality.

*Keywords and phrases:* model-free pricing, extremal measures, martingale optimal transport, weak predictable representation property, cycles.

*MSC Classification 2010:* 60G42, 91G80.

## 1 Introduction

In this paper we study the supports of extremal martingale measures, defined on the product space  $\mathbb{R}_+^2 = (0, \infty)^2$  equipped with its Borel  $\sigma$ -field, under the constraints of having given marginals  $\mu$  and  $\nu$ . The set of all such measures, which is nonempty if and only if  $\mu$  is smaller than  $\nu$  in the convex order, is at the core of *martingale optimal transport*, a new field of research that has been introduced by [3] in the discrete-time case and by [15] in the continuous-time case. The martingale optimal transport problem is a variant of the classical Monge-Kantorovich optimal transport problem (see [32]), and it consists in optimizing a given functional over the set  $\mathcal{M}(\mu, \nu)$  of all probability measures with pre-specified marginals  $\mu$  and  $\nu$  and satisfying the martingale property. The latter property is what makes the difference with the classical optimal transport problem and it is motivated

---

\*This work is partially supported by the ANR project ISOTACE (ANR-12-MONU-0013). We are also grateful to Beatrice Acciaio, Peter Allen, Graham Brightwell, Alex Cox and Frédéric Patras for their helpful remarks.

<sup>†</sup>London School of Economics and Political Science, Department of Statistics, United Kingdom.

<sup>‡</sup>Zeliade Systems, France.

by financial applications. An important growing literature originated from the seminal paper [18], which started the model-free approach to derivative pricing using techniques based on the Skorokhod embedding problem. Within this approach, only very weak assumptions are made, namely the price process of the underlying is a martingale (to rule-out arbitrage opportunities) and its marginals are given by the observation of European Call prices (via the so-called Breeden-Litzenberger formula<sup>1</sup>). Hence, computing for instance the super-replication price of some derivative boils down to maximizing the expected value of its pay-off, say  $f$ , over the set  $\mathcal{M}(\mu, \nu)$ , yielding the following martingale transport problem:

$$\sup_{Q \in \mathcal{M}(\mu, \nu)} Q(f), \tag{1.1}$$

where  $Q(f)$  denotes the expectation of  $f$  under  $Q$ . This problem has been studied in great depth in [4] for a large class of payoffs. The results therein have been further generalized in [16]. In the papers [20, 19] the (martingale) optimal transport has been found for  $f(x, y) = \pm|x - y|$ .

Our interest in the extremal elements of the set  $\mathcal{M}(\mu, \nu)$  is fundamentally motivated by the model-free approach. Indeed, the notion of extremal measures is intimately related to that of optimizers (see, for instance, Bauer Maximum Principle, [2, Sec. 7.69]) and hence to model-free derivatives pricing. Therefore, understanding the support of extremal measures can give insights on the solutions of martingale optimal transport problems such as (1.1). Another motivation for this study comes as a consequence of our first result (cf. Theorem 3.3), which roughly states the following equivalence:

*A martingale measure  $Q \in \mathcal{M}(\mu, \nu)$  is extremal if and only if every derivative can be approximately replicated on the support of  $Q$  by semi-static strategies.*

This can be seen as an extension, to the model-free setting, of the well-known equivalence in the classical setting between “market completeness” and extremality of  $Q$  in the set of all martingale measures without constraints on the marginals (see, e.g. [24] for the discrete-time case), which is in turn the financial translation of one of the most important results in martingale theory, namely that extremality is equivalent to the predictable representation property (see [11] in discrete-time and, e.g., [30, Theorem 4.7, Ch. V] in continuous-time).

Therefore, in the model-free setting, knowing the support of extremal measures in  $\mathcal{M}(\mu, \nu)$  gives a way to generate models where any derivative can be (approximately) replicated by semi-static strategies.

Besides the financial motivation, the problem of characterizing the supports of extremal measures is mathematically interesting on its own and it has quite a long history. Indeed, there is a rich literature on the support of extremal probability measures with given marginals (without the martingale property), which goes back to a paper by Birkhoff [7], where a complete description of extremal measures is given in the finite case, i.e. both

---

<sup>1</sup>We recall that the Breeden-Litzenberger formula (cf. [8]) states that from the prices of European Call options  $C(K, T) = \mathbb{E}_Q[(S_T - K)_+]$  with a fixed maturity  $T$  and for all strike prices  $K > 0$ , one can deduce the law of the underlying  $S_T$  under  $Q$ . Indeed taking the right-derivative in  $K$  gets  $\frac{\partial}{\partial K^+} C(T, K) = -Q(S_T > K)$ , for all  $K > 0$ .

marginals have finite supports. The main result therein establishes that a probability measure with given marginals is extremal if and only if its support does not contain any cycle. Many papers followed, e.g. [5, 6, 12, 14, 17, 22, 23, 25, 27, 28] among others, giving different kinds of characterizations in the finite or countable case and going from functional analysis to combinatorics. In particular [12] extends to the countable case Birkhoff’s result about absence of cycles in the support of extremal measures. In the general case, the problem of giving a complete description of extremal measures with given marginals is still open.

Inspired by this literature, the present paper provides, in full generality, a characterization of extremality in the martingale case in terms of a weaker form of the predictable representation property. In the more specific case of marginals with countable supports, we define a slightly stronger property (called WEP) which allows us to focus on the combinatorial properties of the support of a given measure in  $\mathcal{M}(\mu, \nu)$ . Therefore, we propose two sufficient conditions, called “2-link property” and “full erasability”, having a strong combinatorial flavour. Three important examples satisfy those criteria and hence they are extremal measures: the binomial tree, Hobson and Klimmek’s trinomial tree (cf. [19]) and the left curtain introduced in [4] (at least in the case when one of the two marginals has finite support). Those criteria are very easy to implement for generating many other examples of extremal supports. Moreover, we introduce the new notion of 2-net, which allows to formulate a necessary and sufficient condition for the WEP when the support of the first marginal is finite. Finally, we also investigate to which extent a characterization of extremality in terms of absence of cycles (compare [25, 28]) is possible in the martingale setting.

The paper is organized as follows: Section 2 sets the framework and the main notation, while in Section 3 we give a characterization of extremality in terms of a weak predictable representation property. In Section 4 we introduce the weak exact predictable representation property (WEP). Section 5 contains the two sufficient conditions with examples. Moreover, in Section 6 we study the relation between 2-nets and extremality. Finally, Section 7 focuses on the relation between cycles, extremality and WEP, and Section 8 concludes the paper.

## 2 Setting and notation

Let  $\mu$  and  $\nu$  be two probability laws on  $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ , where  $\mathbb{R}_+ := (0, \infty)$  denotes the set of all positive real numbers. Let  $\mathcal{P}(\mu, \nu)$  denote the set of all probability measures on  $(\mathbb{R}_+^2, \mathcal{B}(\mathbb{R}_+^2))$  with marginals  $\mu$  and  $\nu$ , i.e.

$$Q(A \times \mathbb{R}_+) = \mu(A), \quad Q(\mathbb{R}_+ \times A) = \nu(A), \quad \text{for all } A \in \mathcal{B}(\mathbb{R}_+).$$

For any  $Q \in \mathcal{P}(\mu, \nu)$ , the following decomposition holds:

$$Q(dx, dy) = q(x, dy)\mu(dx),$$

where  $q(x, dy)$  is a probability kernel. We will always work under the following assumption:

**Assumption 2.1.** Let  $\int x\mu(dx) = \int y\nu(dy) = 1$  and  $\mu \preceq \nu$  in the convex order, i.e.

$$\int c(x)\mu(dx) \leq \int c(y)\nu(dy),$$

for all convex functions  $c : \mathbb{R}_+ \rightarrow \mathbb{R}$ .

Let  $\mathcal{M}(\mu, \nu)$  denote the set of all probability measures in  $\mathcal{P}(\mu, \nu)$  with the martingale property, i.e.

$$\int_{\mathbb{R}_+} yq(x, dy) = x, \quad \mu - a.e.$$

Assumption 2.1 implies that  $\mathcal{M}(\mu, \nu)$  is nonempty (cf. [21, 31]). Moreover the set  $\mathcal{M}(\mu, \nu)$  is compact for the weak convergence of measures (cf. Proposition 2.4 in [3]).

The central notion of this paper is the one of extremal point of  $\mathcal{M}(\mu, \nu)$ , which is any probability measure  $Q \in \mathcal{M}(\mu, \nu)$  such that if  $Q = \alpha Q_1 + (1 - \alpha)Q_2$  for some  $\alpha \in (0, 1)$  and  $Q_i \in \mathcal{M}(\mu, \nu)$ ,  $i = 1, 2$ , then  $Q = Q_1 = Q_2$ . The fact that  $\mathcal{M}(\mu, \nu)$  is weakly compact yields that the set of its extremal points is nonempty (cf. Corollary 7.66 in [2]). When there is no ambiguity, “ $Q$  extremal” will mean “ $Q$  extremal in  $\mathcal{M}(\mu, \nu)$ ”.

Finally, for any real-valued measurable function  $f$  defined on some probability space  $(\Omega, \mathcal{F}, Q)$  we will use indifferently the notations  $\mathbb{E}_Q(f) = Q(f) = \int f dQ = \int f(x, y)dQ(x, y)$  for the expectation of  $f$  under  $Q$ .

**Remark 2.2.** The setting admits the usual model-free finance interpretation as follows: let  $(x, y)$  be a generic element of the sample space  $\mathbb{R}_+^2$ , and let  $X$  (resp.  $Y$ ) denote the application  $X(x, y) = x$  (resp.  $Y(x, y) = y$ ). Hence  $(X, Y)$  is a two-dimensional random vector defined on the measurable space  $(\mathbb{R}_+^2, \mathcal{B}(\mathbb{R}_+^2))$ . Under any measure  $Q \in \mathcal{M}(\mu, \nu)$ ,  $X$  and  $Y$  have respective laws  $\mu$  and  $\nu$ . Moreover, under Assumption 2.1,  $(1, X, Y)$  is martingale under  $Q$ . Hence it can be viewed as the (discounted) price process of some risky asset. Moreover any measure  $Q \in \mathcal{M}(\mu, \nu)$  specifies the full law of the price process, in other terms gives a price model for the risky asset, which is compatible with the knowledge of the marginals  $\mu, \nu$  as well with the absence of arbitrage opportunities (due to the martingale property).

### 3 The Douglas-Lindenstrauss-Naimark Theorem and its consequences

In this section, we give a functional analytical characterisation of extremality in  $\mathcal{M}(\mu, \nu)$  with a natural financial interpretation. We stress that the results in this part of the paper hold in full generality. More restrictive assumptions on the supports of the marginals  $\mu$  and  $\nu$  will be needed in the next sections. We start with recalling the following classical result relating extremality and denseness of subspaces of  $L^1(Q)$ .

**Theorem 3.1** (Douglas [14], Lindenstrauss [27], Naimark [29]). *Let  $(\Omega, \mathcal{F}, Q)$  be a probability space and let  $F$  be a linear subspace of  $L^1(Q)$  such that  $1 \in F$ . The following are equivalent:*

(i)  $Q$  is an extremal point of the set of all probability measures  $R$  on  $(\Omega, \mathcal{F})$  (not necessarily equivalent to  $Q$ ) such that  $\mathbb{E}_R(f) = \mathbb{E}_Q(f)$  for all  $f \in F \cap L^1(R)$ ;

(ii)  $F$  is dense in  $L^1(Q)$  with the strong topology.

**Remark 3.2.** The Douglas-Lindenstrauss-Naimark theorem was used in [30, Ch. V] to prove the predictable representation property (PRP) for continuous martingales in the Brownian filtration. Other applications of this theorem in relation with various notions of market completeness can be found in [9, 10].

A direct application of this theorem to our setting gives the following equivalence:

**Theorem 3.3.** *Let  $Q \in \mathcal{M}(\mu, \nu)$ . The following two properties are equivalent:*

(i)  $Q$  is extremal in  $\mathcal{M}(\mu, \nu)$ ;

(ii) the weak Predictable Representation Property (PRP) holds in the following sense: the set of all functions  $f \in L^1(Q)$  that can be represented as

$$f(x, y) = \varphi(x) + h(x)(y - x) - \psi(y), \quad Q - a.s. \quad (3.1)$$

for some functions  $\varphi \in L^1(\mu), \psi \in L^1(\nu)$  and  $h \in L^0(\mu)$ , is dense in  $L^1(Q)$ .

*Proof.* We prove first that weak PRP under  $Q \in \mathcal{M}(\mu, \nu)$  implies that  $Q$  is extremal. Assume that  $Q = \alpha Q_1 + (1 - \alpha)Q_2$  for some  $\alpha \in (0, 1)$  and  $Q_i \in \mathcal{M}(\mu, \nu)$  for  $i = 1, 2$ . Therefore,  $Q_i \ll Q$  for  $i = 1, 2$ . Consider any measurable bounded functions  $f$  such that  $f(x, y) = \varphi(x) + h(x)(y - x) - \psi(y)$  for suitable functions  $h, \varphi, \psi$  and  $Q$ -a.s., hence  $Q_i$ -a.s. as well. Taking the expectation under those measures, we get  $Q(f) = \mu(\varphi) + \nu(-\psi) = Q_i(f)$  for  $i = 1, 2$ , for all bounded functions  $f$  satisfying (3.1). By denseness, we obtain that  $Q = Q_1 = Q_2$ , i.e.  $Q$  is extremal in  $\mathcal{M}(\mu, \nu)$ . It remains to prove the converse. To do so, it suffices to apply Theorem 3.1 to the set

$$F = \{f \in L^1(Q) : f(x, y) = \varphi(x) + h(x)(y - x) - \psi(y), \\ \text{for some } \varphi \in L^1(\mu), \psi \in L^1(\nu), h \in L^0(\mu)\} \quad (3.2)$$

noticing that with such a choice of the subspace  $F$ , the condition  $\int f dR = \int f dQ$  for all  $f \in F \cap L^1(R)$  implies that  $R \in \mathcal{M}(\mu, \nu)$ .  $\square$

Theorem 3.3 has a natural financial interpretation. Any measure  $Q \in \mathcal{M}(\mu, \nu)$  can be seen as some price model consistent with the marginals  $\mu, \nu$ . Hence, the theorem above gives that the extremal models are the ones where any contingent claim can be (approximately) replicated by trading dynamically in the underlying in a self-financing way and statically in some European options with payoffs  $\varphi$  (at time 1) and  $\psi$  (at time 2).

**Proposition 3.4.**  $Q \in \mathcal{M}(\mu, \nu)$  is extremal if and only if for any  $R \in \mathcal{M}(\mu, \nu)$  with  $R \ll Q$  we have  $R = Q$ .

*Proof.* Let  $Q \in \mathcal{M}(\mu, \nu)$ . Assume that  $Q$  is not extremal, i.e.  $Q = \alpha Q_1 + (1 - \alpha)Q_2$  for some  $Q_i \in \mathcal{M}(\mu, \nu)$ ,  $i = 1, 2$ , with  $Q_1 \neq Q$  and some  $\alpha \in (0, 1)$ . Hence we have  $Q_1 \ll Q$ . Now let  $Q$  be extremal and  $R \in \mathcal{M}(\mu, \nu)$  such that  $R \ll Q$ . Denote  $\ell = \frac{dR}{dQ}$  the corresponding density given by the Radon-Nikodym Theorem. For any  $f \in L^1(Q) \cap L^1(R)$  with representation triple  $(h, \varphi, \psi)$  as in (3.1) one has

$$\mu(\varphi) - \nu(\psi) = \int f dQ = \int f dR = \int \ell f dQ,$$

Since the set  $F$  of such functions is dense in  $L^1(Q)$  it follows that the linear functional  $f \mapsto \int \ell f dQ$  matches the continuous linear functional  $f \mapsto \int f dQ$  on a dense subset of  $L^1(Q)$  and so it is continuous on  $L^1(Q)$ . Therefore we can choose a version of  $\ell$  in  $L^\infty(Q)$ . It follows that  $\beta = \ell - 1$  is orthogonal to  $F$  which entails  $\ell = 1$  so that  $R = Q$ .  $\square$

A consequence of Proposition 3.4 is that the extremal points of  $\mathcal{M}(\mu, \nu)$  are fully characterized by their supports.

## 4 The Weak Exact Predictable Representation Property

From now on we will work under the following standing assumption:

**Assumption 4.1.**  $\mu$  and  $\nu$  are supported on countable subsets  $X$  and  $Y$  of  $\mathbb{R}_+$ .

We introduce some notations in this discrete support context. Let  $S$  be any subset of  $X \times Y$ . For  $(x, y) \in S$  we let

$$Y_S(x) = \{z \in Y : (x, z) \in S\}, \quad X_S(y) = \{t \in X : (t, y) \in S\},$$

and

$$S_X = \{x \in X : \exists y \in Y, (x, y) \in S\}, \quad S_Y = \{y \in Y : \exists x \in X, (x, y) \in S\}.$$

We call *mesh* any set  $S \subset X \times Y$  such that  $|S_X| = 1$ . If in addition  $S$  satisfies  $|S_Y| = 2$ , it will be called *2-mesh* (or binomial mesh). For any measure  $Q \in \mathcal{P}(\mu, \nu)$  we define its support as the set

$$\text{supp}(Q) := \{(x, y) \in X \times Y : Q(x, y) > 0\}.$$

Whenever  $S$  is the support of a probability  $Q$ , i.e.  $S = \text{supp}(Q)$ , and when there is no ambiguity we will drop  $S$  from the notation  $X_S(y), Y_S(x)$  and simply write  $X(y), Y(x)$ .

Finally, the notation  $|A|$  denotes the cardinality of any set  $A$  and by numbering (or ordering) of any countable set  $A$  we mean any possible representation of the set as a sequence.

We introduce now the following *exact* strengthening of the weak PRP, which is motivated by Mukerjee purely geometrical characterization of extremality in the non-martingale case (see the introduction in [28] and his Theorem 2.7). Such a property is easier to handle than the weak PRP (cf. Theorem 3.3) since it avoids the issues of integrability of the functions appearing in the representation, as well as the passage to the limit in  $L^1(Q)$ . Yet the main

difference with respect to the weak PRP is subtler: we require the replication property only for functions defined on a subset  $S$  of the product space  $X \times Y$ . In most instances,  $S$  will be the support of some measure  $Q$ .

**Definition 4.2** (Weak Exact PRP). *Let  $S$  be a subset of  $X \times Y$ . We say that the Weak Exact PRP (henceforth, the WEP) holds for  $S$  if for every function  $f : S \rightarrow \mathbb{R}$ , there exist functions  $h, \varphi : S_X \rightarrow \mathbb{R}$  and  $\psi : S_Y \rightarrow \mathbb{R}$  such that*

$$f(x, y) = \varphi(x) + h(x)(y - x) - \psi(y), \quad (x, y) \in S. \quad (4.1)$$

Moreover, let  $f : S \rightarrow \mathbb{R}$  be an arbitrary function. We say that  $WEP(f)$  holds in  $S$  if (4.1) is satisfied for suitable functions  $h, \varphi, \psi$ .

We say that the WEP holds for  $Q \in \mathcal{M}(\mu, \nu)$  if the WEP holds for the support  $S$  of  $Q$ .

**Example 4.3** (Binomial mesh). *Let  $x, y_1, y_2$  be some given positive numbers with  $y_1 \neq y_2$  and let  $S := \{(x, y_i) : i = 1, 2\}$  be the corresponding 2-mesh. Then the WEP holds for  $S$ . Indeed for any given  $f$ , set  $\psi = 0$ ,  $h(x) = \frac{f(x, y_1) - f(x, y_2)}{y_1 - y_2}$ , and  $\varphi(x) = \frac{x - y_2}{y_1 - y_2} f(x, y_1) + \frac{y_1 - x}{y_1 - y_2} f(x, y_2)$ .*

**Remark 4.4.** Notice that by defining the WEP as a property of a given set of paths  $S$ , we possibly include sets that cannot be supports of martingale probability measures, as seen on the previous example when, for instance,  $y_1 > y_2 > x > 0$ . More generally, if the WEP holds for some set  $S \subset X \times Y$  and if  $m : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is an application such that  $id + m$  is invertible, where  $id$  stands for the identity map, then the WEP holds for the set  $S_m = \{(x + m(x), y) : (x, y) \in S\}$  as well. Indeed, consider a function  $f : S_m \rightarrow \mathbb{R}$  and let  $(\varphi, h, \psi)$  be the decomposition of the function  $g(x, y) = f(x + m(x), y)$ . We have  $f(x + m(x), y) = \varphi(x) + h(x)(y - x) - \psi(y) = \varphi(x) + h(x)m(x) + h(x)(y - (x + m(x))) - \psi(y)$  so that by setting

$$\varphi_f(x + m(x)) = \varphi(x) + h(x)m(x), \quad h_f(x + m(x)) = h(x),$$

the WEP for  $S_m$  follows since  $id + m$  is invertible. This shows in particular that the WEP is a property of purely geometric and combinatorial nature.

We conclude this section with showing that the WEP is a sufficient condition for the extremality of a measure  $Q$  in  $\mathcal{M}(\mu, \nu)$  when either the support of  $\mu$  or the one of  $\nu$  are finite. This is not surprising since the WEP is *exact* on the support of  $Q$ .

**Proposition 4.5.** *Let  $Q \in \mathcal{M}(\mu, \nu)$  and let  $S$  be its support. Assume that either  $S_X$  or  $S_Y$  is finite. If the WEP holds for  $Q$ , then  $Q$  is extremal.*

*Proof.* Take some function  $g \in L^1(Q)$ . By assumption there exist measurable functions  $\varphi, h : S_X \rightarrow \mathbb{R}$  and  $\psi : S_Y \rightarrow \mathbb{R}$  such that  $g(x, y) = \varphi(x) - \psi(y) + h(x)(y - x)$  on the support

of  $Q$ . If  $S_X$  is finite,  $(x, y) \mapsto \varphi(x)$  belongs to  $L^1(Q)$  as well as  $(x, y) \mapsto h(x)(y - x)$  since

$$\begin{aligned} \int |h(x)||y - x|dQ(x, y) &\leq \|h\|_\infty \int |y - x|dQ(x, y) \\ &\leq \|h\|_\infty \int (x + y)dQ(x, y) \\ &= 2\|h\|_\infty \int xd\mu(x) < \infty, \end{aligned}$$

where  $\|\cdot\|_\infty$  denotes the sup-norm and the last equality is due to the martingale property of  $Q$ . Therefore  $(x, y) \mapsto \psi(y) = -g(x, y) + \varphi(x) + h(x)(y - x)$  belongs also to  $L^1(Q)$ , so that  $Q$  is extremal by Douglas-Lindenstrauss-Naimark Theorem 3.1.

Assume now  $S_Y$  finite. Then the function  $(x, y) \mapsto g(x, y) + \psi(y)$  belongs to  $L^1(Q)$ , and by taking conditional expectations we get  $\int g(x, y)q(x, dy) = \varphi(x) - \int \psi(y)q(x, dy)$ . Now  $\int |\psi(y)|q(x, dy) \leq \|\psi\|_\infty$  and

$$\int \left| \int g(x, y)q(x, dy) \right| d\mu(x) \leq \int \int |g(x, y)|q(x, dy)d\mu(x) = \|g\|_{L^1(Q)} < \infty,$$

so that  $(x, y) \mapsto \varphi(x)$  belongs to  $L^1(Q)$ . On the other hand  $(x, y) \mapsto h(x)(y - x) = g(x, y) - \varphi(x) + \psi(y)$  belongs to  $L^1(Q)$  and we can conclude in the same way as before.  $\square$

**Remark 4.6.** When both  $\mu$  and  $\nu$  have finite support, WEP and extremality are actually equivalent. To see this, just notice that  $L^1(Q)$  can be identified with the set of all functions  $f : \text{supp}(Q) \rightarrow \mathbb{R}$ , which is a finite dimensional vector space. Hence any dense subspace of  $L^1(Q)$  equals  $L^1(Q)$  so that the WEP and the weak PRP coincide. In particular, every extremal measure  $Q$  satisfies the WEP.

**Remark 4.7.** In general, the set of semi-static trading strategies is not closed in  $L^1(Q)$ , as it is showed in the recent preprint [1]. More precisely, in their Theorem 1.1, the authors construct a discrete-time model, defined on a countable sample space, and a sequence of semi-static strategies converging in  $L^p$ , for every  $p \geq 1$ , to some limit which cannot even be dominated by the final outcome of a semi-static strategy. The problem of whether this would hold even if  $Q$  is an extremal measure in  $\mathcal{M}(\mu, \nu)$  is still open.

## 5 Some sufficient conditions for the WEP

In this section we provide some easily verifiable sufficient conditions for the WEP to hold.

### 5.1 Intersection lemma

We start with some preliminary results showing that the WEP (as well as extremality) entails quite a strong constraint on the intersection of the image sets  $Y(x)$ , for  $x \in X$ .



**Lemma 5.1** (Intersection Lemma under the WEP). *Assume that the WEP holds for  $S$  and let  $x_1, x_2$  be two distinct points in  $S_X$ . Then*

$$|Y(x_1) \cap Y(x_2)| \leq 2. \quad (5.1)$$

*Proof.* Assume that there exist two distinct points  $x_1, x_2 \in S_X$  such that  $Y(x_1) \cap Y(x_2) \supset \{y_1, y_2, y_3\}$  with  $y_1 < y_2 < y_3$ . Hence, one can choose a function  $f : S \rightarrow \mathbb{R}$  such that  $f(x_1, \cdot)$  and  $f(x_2, \cdot)$  have, respectively, strictly increasing and strictly decreasing increment ratios over the set  $\{y_i, i = 1, 2, 3\}$ , i.e.

$$\frac{f(x_1, y_2) - f(x_1, y_1)}{y_2 - y_1} < \frac{f(x_1, y_3) - f(x_1, y_2)}{y_3 - y_2} \quad (5.2)$$

and

$$\frac{f(x_2, y_2) - f(x_2, y_1)}{y_2 - y_1} > \frac{f(x_2, y_3) - f(x_2, y_2)}{y_3 - y_2}. \quad (5.3)$$

Since the WEP holds for  $S$ ,  $f$  can be represented as

$$f(x, y) = \varphi(x) + h(x)(y - x) - \psi(y), \quad x \in S_X, \quad y \in Y(x),$$

for some functions  $h, \varphi : S_X \rightarrow \mathbb{R}$  and  $\psi : S_Y \rightarrow \mathbb{R}$ . Hence (5.2) and (5.3) become

$$\frac{\psi(y_2) - \psi(y_1)}{y_2 - y_1} < \frac{\psi(y_3) - \psi(y_2)}{y_3 - y_2}, \quad \frac{\psi(y_2) - \psi(y_1)}{y_2 - y_1} > \frac{\psi(y_3) - \psi(y_2)}{y_3 - y_2}.$$

leading to a contradiction. □

Even though we do not know in general the relationship between WEP and extremality, we can show that in the intersection lemma above the WEP can be replaced by the extremality property while keeping the same conclusion.

**Lemma 5.2** (Intersection Lemma under extremality). *Assume that  $Q$  is extremal. Let  $S$  be the support of  $Q$  and  $x_1, x_2$  two distinct points in  $S_X$ . Then*

$$|Y(x_1) \cap Y(x_2)| \leq 2. \quad (5.4)$$

*Proof.* Let us proceed by contradiction and assume that there are at least three distinct points  $\{y_1, y_2, y_3\}$  in  $Y(x_1) \cap Y(x_2)$ . We are going to build a perturbation of  $Q$  in  $\mathcal{M}(\mu, \nu)$ , showing that  $Q$  cannot be extremal. Consider the positive number

$$c := \inf\{Q(x_i, y_j) : i = 1, 2, j = 1, 2, 3\} > 0.$$

Let us start with the path  $(x_1, y_1)$  and perturbate its probability by some number  $\alpha$ , so getting a new probability weight  $Q_1(x_1, y_1) = Q(x_1, y_1) + \alpha$ . Consider now the path  $(x_2, y_1)$  and perturbate it, in order to preserve the total mass at  $y_1$ , by  $-\alpha$ , i.e. define  $Q_1(x_2, y_1) = Q(x_2, y_1) - \alpha$ . In the same way we associate to the path  $(x_2, y_2)$  a perturbation  $\beta$ , to the path  $(x_1, y_2)$  the opposite perturbation  $-\beta$ , and lastly  $\gamma$  to the path  $(x_1, y_3)$  and  $-\gamma$  to the

path  $(x_2, y_3)$ . This procedure leads to a new probability measure  $Q_1$  on  $S$ . By construction the total mass at each point  $y_j$ , for  $j = 1, 2, 3$ , is preserved.

On the other hand, the mass  $\mu(\{x_1\})$  at the point  $x_1$  is preserved if and only if  $\alpha - \beta + \gamma = 0$ , and the conservation of the mass at  $x_2$ , i.e.  $\mu(\{x_2\})$ , gives the same condition. It remains to check the martingale property, giving the same condition for both points  $x_1$  and  $x_2$ , which is  $\alpha y_1 - \beta y_2 + \gamma y_3 = 0$ . Solving this system gives  $\alpha = \beta - \gamma$  and  $\beta(y_1 - y_2) = \gamma(y_1 - y_3)$ . We can choose  $\gamma$  sufficiently small so that  $\max\{|\alpha|, |\beta|, |\gamma|\} < c$ , which guarantees that the perturbation  $Q_1$  constructed above belongs to  $\mathcal{M}(\mu, \nu)$ .

Finally, performing an analogue perturbation whose weights have the same absolute value but opposite signs than those in  $Q_1$ , we obtain another measure  $Q_2$  in  $\mathcal{M}(\mu, \nu)$  such that  $Q = (Q_1 + Q_2)/2$ , which contradicts the extremality of  $Q$ .  $\square$

## 5.2 The 2-link property

In this section we introduce the notion of “2-link property”, which gives a sufficient condition for the WEP to hold on a given subset  $S$  of  $X \times Y$ . This property can be viewed as a strengthening of the necessary condition given in the intersection Lemma 5.1. It is simple to formulate and at the same time it gives an easy method to generate quite a rich family of supports of extremal measures.

**Definition 5.3** (2-link property). *We say that  $S$  has the 2-link property if there exists a numbering  $S_X = (x_n)_{n \geq 1}$  such that for all  $n \geq 1$  we have*

$$|Y(x_n) \cap \bigcup_{i=1}^{n-1} Y(x_i)| \leq 2, \quad (2LP)$$

with the convention  $\bigcup_{i=1}^0 = \emptyset$ .

With a slight abuse of terminology we will sometimes use (2LP) with the meaning of “2-link property”. It will be clear from the context.

**Proposition 5.4.** *If  $S$  has the 2-link property, then the WEP holds for  $S$ .*

*Proof.* Let  $f : S \rightarrow \mathbb{R}$  be any measurable real-valued function, we want to find functions  $\varphi, h : S_X \rightarrow \mathbb{R}$  and  $\psi : S_Y \rightarrow \mathbb{R}$  such that  $f(x, y) + \psi(y) = \varphi(x) + h(x)(y - x)$  on  $S$ . We construct such functions by induction using the condition (2LP), which we assume to be satisfied for at least one numbering  $S_X = (x_n)_{n \geq 1}$ . Take the first element  $x_1$  and consider all the points  $y \in Y(x_1)$ . Pick arbitrarily two such points, say  $y_1, y_2 \in Y(x_1)$ , take any two real numbers  $\psi_1, \psi_2$  and set  $\psi(y_1) := \psi_1$  and  $\psi(y_2) := \psi_2$ . We want to have

$$f(x_1, y_i) + \psi(y_i) = \varphi(x_1) + h(x_1)(y_i - x_1), \quad i = 1, 2,$$

so that the parameters  $\varphi(x_1), h(x_1)$  of the affine function in  $y \in Y$ ,  $y \mapsto \varphi(x_1) + h(x_1)(y - x_1)$ , are determined by the two points on the LHS in the equality above. As a consequence, the other values of the function  $\psi(y)$  for  $y \in Y(x_1) \setminus \{y_1, y_2\}$  are also determined via the equality

$$\psi(y) = f(x_1, y) - \varphi(x_1) - h(x_1)(y - x_1).$$

Now, assume that we have constructed the functions  $h, \varphi : (x_i)_{i=1}^{n-1} \rightarrow \mathbb{R}$  and  $\psi : \cup_{i=1}^{n-1} Y(x_i) \rightarrow \mathbb{R}$  such that

$$f(x_i, y) = \varphi(x_i) + h(x_i)(y - x_i) - \psi(y), \quad y \in Y(x_i), \quad 1 \leq i \leq n-1.$$

Consider the next point  $x_n$  in the given numbering of  $S_X$  satisfying the condition (2LP). The latter implies that there exist at most two distinct points, say,  $y_1, y_2 \in \cup_{i=1}^{n-1} Y(x_i)$  such that  $y_i \in Y(x_n)$  for  $i = 1, 2$ . Thus, the two values  $\psi(y_i)$ ,  $i = 1, 2$ , have already been determined and so are  $f(x_n, y_i) + \psi(y_i)$ ,  $i = 1, 2$ . This identifies completely and without any ambiguity the parameters  $\varphi(x_n)$ ,  $h(x_n)$  in the  $y$ -affine part of the following WEP representation

$$f(x_n, y) + \psi(y) = \varphi(x_n) + h(x_n)(y - x_n), \quad y \in Y(x_n).$$

Indeed we have

$$\begin{aligned} h(x_n) &= \frac{f(x_n, y_2) - f(x_n, y_1) + \psi(y_2) - \psi(y_1)}{y_2 - y_1}, \\ \varphi(x_n) &= f(x_n, y_1) - h(x_n)(y_1 - x_n) + \psi(y_1), \\ \psi(y) &= f(x_n, y) - \varphi(x_n) - h(x_n)(y - x_n), \quad y \in Y(x_n) \setminus \{y_1, y_2\}. \end{aligned}$$

Doing so, we have extended the functions  $h, \varphi$  to the finite set  $(x_i)_{i=1}^n$  and the function  $\psi$  to the set  $\cup_{i=1}^n Y(x_i)$ . By the induction principle, we can conclude that there exist functions  $h, \varphi : S_X \rightarrow \mathbb{R}$  and  $\psi : S_Y = \cup_{n \geq 1} Y(x_n) \rightarrow \mathbb{R}$  such that WEP( $f$ ) holds for any arbitrary function  $f$ .  $\square$

**Remark 5.5.** It is easy to construct conditional supports with infinitely many points satisfying (2LP). This is in contrast to what happens for martingale measures without constraints on the marginals, where the only extremal points have a two-point conditional support (compare to Lemme A in Dellacherie [11] and to Theorem 6 in Jacod and Shiryaev [24]). From a financial perspective, this is clearly due to the fact the extremal points in  $\mathcal{M}(\mu, \nu)$  correspond to semi-statically complete models as in Theorem 3.3, where in particular one is allowed to trade statically in infinitely many European options.

**Example 5.6** (“Binomial tree”). Any probability  $Q \in \mathcal{M}(\mu, \nu)$  whose conditional supports have two points, i.e.  $|Y(x)| = 2$  for all  $x \in X$ , is extremal. Indeed, property (2LP) is trivially satisfied.

**Example 5.7** (Hobson and Klimmek [19] “Trinomial tree”). Hobson and Klimmek [19] constructed an optimal martingale optimal transport whose conditional support is fully characterized as follows: there exist  $a < b$  and two decreasing maps  $p$  and  $q$  such that for those  $x$  with  $x \in Y(x)$  we have

- $Y(x) = \{x\}$  if  $x < a$  or  $x > b$ ,
- $Y(x) = \{p(x), x, q(x)\}$  otherwise, with  $p(x) < a$  and  $q(x) > b$ .

Moreover for those  $x$  which do not belong to  $Y(x)$  we have  $Y(x) = \{p(x), q(x)\}$ . One can see that the property (2LP) is satisfied in this case as well. Indeed, let  $X$  be the (countable) support of  $\mu$  and  $X = X_{<a} \cup X_{[a,b]} \cup X_{>b}$ , with  $X_{>a} := \{x \in X : x > a\}$  (the other two sets are defined analogously). Consider any numbering for those three sets, i.e.  $X_{>a} = (x_n^a)$ ,  $X_{>b} = (x_n^b)$  and  $X_{[a,b]} = (\bar{x}_n)$ . Hence by alternating elements of each sequence we get a numbering for  $X$ , given by  $(x_n) = (x_1^a, x_1^b, \bar{x}_1, \dots)$  which satisfies (2LP). Note that Hobson and Klimmek optimal coupling with  $\mu \wedge \nu = 0$  is a binomial tree. More on this support can be found in Example 5.15.

In the paper [4] Beiglböck and Juillet introduce the fundamental notion of left-monotone martingale transport plan (see Definition 1.4 therein) as follows: a martingale transport plan  $\pi$  on  $\mathbb{R} \times \mathbb{R}$  is called *left-monotone* if there exists a Borel set  $\Gamma \subset \mathbb{R} \times \mathbb{R}$  with  $\pi(\Gamma) = 1$  and such that whenever  $(x, y^-), (x, y^+), (x', y') \in \Gamma$  we cannot have

$$x < x' \quad \text{and} \quad y^- < y' < y^+. \quad (5.5)$$

In Theorem 5.1 in [4], it is proved that there exists a unique left-monotone transport plan in  $\mathcal{M}(\mu, \nu)$ , which is denoted by  $\pi_{lc}$  and called *left curtain*. The right-curtain  $\pi_{rc}$  is defined similarly just by replacing (5.5) with the following forbidden pattern: whenever  $(x, y^-), (x, y^+), (x', y') \in \Gamma$  we cannot have

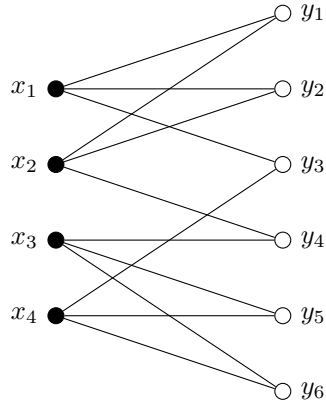
$$x > x' \quad \text{and} \quad y^- < y' < y^+. \quad (5.6)$$

**Proposition 5.8.** *Assume that there exists a strictly decreasing (resp. strictly increasing) numbering for  $S_X$ , i.e.  $S_X = (x_n)_{n \geq 1}$  with  $x_1 > x_2 > \dots$  (resp.  $x_1 < x_2 < \dots$ ). Then, the left (resp. right) curtain  $\pi_{lc}$  (resp.  $\pi_{rc}$ ) satisfies the property (2LP). In particular, it satisfies the WEP and so, if either  $S_X$  or  $S_Y$  is finite, it is extremal in  $\mathcal{M}(\mu, \nu)$ .*

*Proof.* Assume by contradiction that all numberings of  $S_X$  do not satisfy (2LP), hence the decreasing order  $x_1 > x_2 > \dots$  in particular does not fulfil it. Therefore, there exists  $k \geq 2$  such that  $Y(x_k) \cap (\cup_{1 \leq i \leq k-1} Y(x_i))$  contains three distinct points  $y', y'', y'''$  in  $Y$ . We can order them as  $y^u > y^m > y^l$ . There exist  $x_i$  with  $i = 1, \dots, k-1$  such that  $y^m \in Y(x_i)$ . Then, we have found that  $(x_k, y^j)$ , with  $j \in \{u, m, l\}$ , belongs to the support of the left curtain together with  $(x_i, y_m)$ , where we recall that  $x_i > x_k$ . This is exactly the forbidden mapping (5.5) in the left curtain definition (see also [4, Figure 1]). Hence, the left curtain support satisfies (2LP) and it satisfies the WEP (cf. Proposition 5.4). As a consequence, the left curtain is extremal in  $\mathcal{M}(\mu, \nu)$ . The proof for the right curtain is similar.  $\square$

Now we provide an example showing that the 2-link property is not necessary for the WEP.

**Example 5.9.** The picture below shows a subset  $S$  of  $X \times Y$ , with  $X = \{x_i\}_{i=1}^4$  and  $Y = \{y_j\}_{j=1}^6$ , which does not have the two-link property and nonetheless one can check by direct verification that the WEP is fulfilled.



We will come back to this picture in the next section (see Example 6.11).

### 5.2.1 Relation to graph theory

We conclude this section by showing that the 2-link property is very much related to the notion of  $k$ -degeneracy in graph theory as in, e.g., [26]. In particular, we see how this unexpected relation could provide alternative characterizations of (2LP) as well as a way to generate subsets of  $X \times Y$  satisfying (2LP). We will use a little terminology of graph theory, for whom we refer to Diestel's book [13].

Let  $G = (V, E)$  be a graph with vertex set  $V$  and edge set  $E$ . For our purposes, we allow  $G$  to have multiple edges. Moreover,  $V$  and  $E$  can be infinite countable sets. The degree  $d(v)$  of a vertex  $v \in V$  is the number of edges incident with  $v$ . The smallest degree among the vertices of  $G$  is called *minimum degree* of  $G$  and is denoted by  $\delta(G)$ . Moreover, a subgraph  $H$  of a graph  $G$  consists of a subset of the vertices of  $G$  and a subset of the edges of  $G$  which together form a graph. The *subgraph induced* by a set  $U$  of vertices of  $G$ , denoted by  $\langle U \rangle$ , has  $U$  as its vertex set and contains all edges of  $G$  incident with two vertices of  $U$ .

**Definition 5.10** ([26]). *A graph  $G$  is said to be  $k$ -degenerate, for  $k$  a nonnegative integer, if for each induced subgraph  $H$  of  $G$ , we have  $\delta(H) \leq k$ .*

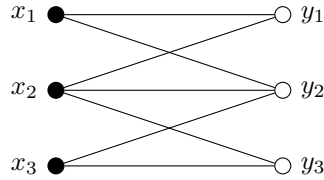
The following proposition states the equivalence between  $k$ -degeneracy of a graph, with  $k = 2$ , and a property very similar to (2LP). When  $G$  is finite and does not have multiple edges, this is just Proposition 1 in [26]. However, even when  $G$  can have countably many vertices and multiple edges, such an equivalence still holds. We provide the proof in the case  $k = 2$  for reader's convenience.

**Proposition 5.11** ([26]).  *$G = (V, E)$  is 2-degenerate if and only if the set of its vertices  $V$  admits an order  $V = (v_n)_{n \geq 1}$  such that  $d(v_1) \leq 2$  and, in the induced subgraph  $\langle \{v_1, \dots, v_{n-1}\} \rangle$  of  $G$ , we have  $d(v_n) \leq 2$ , for each  $n \geq 1$ .*

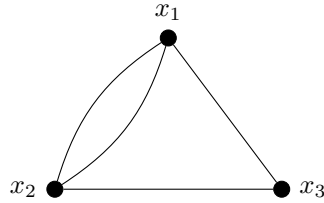
*Proof.* Assume that  $G$  is 2-degenerate. One can find such an ordering as follows: pick the vertex with the smallest degree, name it  $x_1$  and remove it from the graph. Repeat the

procedure with the remaining subgraph and iterate. Now, assume that the order is given and there exists an induced subgraph  $H$  of  $G$  with  $\delta(H) > 2$ . Choose  $n$  sufficiently large so that  $V(H)/\{v\} \subset \{v_1, \dots, v_{n-1}\}$ , where  $V(H)$  is the vertex set of  $H$  and  $v$  is one of its vertices. Now, since  $\delta(H) > 2$ , the degree of  $v$  in the induced subgraph  $\langle v_1, \dots, v_{n-1} \rangle$  is strictly bigger than 2, which contradicts the property of the ordering.  $\square$

We show how a subset  $S$  which has the 2-link property can be turned to a 2-degenerate graph and vice-versa. First, notice that  $S$  can be viewed as a (possibly infinite) bipartite undirected graph  $G = (V, E)$ , where  $V = X \times Y$  is the set of vertices and  $E = S$  is the set of edges so that  $e = xy$  is an edge of  $G$  if and only if  $(x, y) \in S$ , i.e.  $y \in Y(x)$ . For our purpose, let us define a simpler graph with vertex set  $X$ . Let  $G^X = (V^X, E^X)$  be a graph with  $V^X = X$  and  $E^X$  is such that  $x_1x_2 \in E^X$  if and only if  $x_1y$  and  $x_2y$  belong to  $E$  for some  $y \in Y$ , with the constraint that the same  $y$  cannot be used more than twice. Notice that the graph  $G^X$  can have multiple edges. Moreover, different  $G^X$  can be constructed starting from the same graph  $G$ . Let us illustrate this construction in the following example: let  $G$  be as in the picture below:



Notice that  $G$  satisfies (2LP). One possible graph  $G^X$  produced out of  $G$  as described above is given by



the others can be obtained from the latter by re-labeling the vertices, in other terms such graphs are isomorphic.<sup>2</sup> Let us denote  $\tilde{G}^X$  the associated equivalence class. The following equivalence is an immediate consequence of Proposition 5.11 and the way  $G^X$  has been defined.

**Proposition 5.12.** *Let  $S$  be a subset of  $X \times Y$  and let  $G$  be the corresponding graph. Then,  $S$  (or equivalently  $G$ ) satisfies (2LP) if and only if  $\tilde{G}^X$  is 2-degenerate.*

<sup>2</sup>Let  $G = (V, E)$  and  $G' = (V', E')$  be two graphs. We call  $G$  and  $G'$  isomorphic if there exists a bijection  $\eta : V \rightarrow V'$  such that  $xy \in E$  if and only if  $\eta(x)\eta(y) \in E'$  for all  $x, y \in V$ . Such a map is called isomorphism (cf. [13, Sec. 1.1]).

The article [26] contains many examples of  $k$ -degenerate (finite) graphs, that could be used to generate supports of extremal measures in  $\mathcal{M}(\mu, \nu)$ . A relevant class of examples is the class of connected 3-regular graphs (all vertices have exactly three neighbours). Such graphs are not 2-degenerate themselves, but deleting any of their vertices leaves a 2-degenerate graph. A way to generate graphs fulfilling (2LP) is the following: given a 2-degenerate graph  $G^X = (V^X, E^X)$  with vertex set  $V^X = X$ , we define a graph  $G = (V, E)$  with vertex set  $V = X \times Y$  and edge set  $E$  obtained by splitting any edge  $e = xx' \in E^X$  into two edges  $xy$  and  $x'y$  for some  $y \in Y$ . The new graph  $G$  satisfies (2LP) by construction.

A deeper understanding of (2LP) and 2-degeneracy within graph theory goes beyond the scope of this paper. We postpone the study of this interplay for future research.

### 5.3 Erasable sets

In this section we provide another sufficient condition for the WEP. While the 2-link property imposes a kind of compatibility condition among paths when adding more and more points, the condition given here is based on erasing paths in a certain way. This is motivated by the analogous property in the non martingale case (cf [28], Theorem 2.3).

**Lemma 5.13.** *Let  $S \subset X \times Y$  and let  $U = \{(x, y) \in S : |X(y)| = 1 \text{ or } |Y(x)| \leq 2\}$ . Then the WEP holds for  $S$  if and only if it holds for  $S \setminus U$ .*

*Proof.* It is clear that if the WEP holds for  $S$ , then it holds for  $S \setminus U$ . Now, assume that the WEP holds for  $S \setminus U$ . Let  $(x, y) \in U$ . Consider first the case when  $|X(y)| = 1$  and  $x \in X(y')$  for some  $y'$  such that  $(x, y') \in S \setminus U$ . In this case the value  $\psi(y)$  can be taken as  $\psi(y) := f(x, y) - \varphi(x) - h(x)(y - x)$  where  $\varphi(x)$  and  $h(x)$  are given by the WEP for  $S \setminus U$ . On the other hand, if  $|Y(x)| = 2$  with  $Y(x) = \{y_1, y_2\}$ , then  $\varphi(x)$  and  $h(x)$  are uniquely determined by the values of  $\psi$  on  $Y(x)$  via the following equations:

$$h(x) = \frac{\psi(y_1) - \psi(y_2)}{y_1 - y_2}, \quad \varphi(x) = \frac{x - y_2}{y_1 - y_2} \psi(y_1) + \frac{y_1 - x}{y_1 - y_2} \psi(y_2). \quad (5.7)$$

If  $|Y(x)| = 1$ , i.e.  $Y(x) = \{y\}$ , then the choice  $\varphi(x) = \psi(y) + f(x, y)$  and  $h(x) = 0$  allows to satisfy the WEP along the path  $(x, y)$ .  $\square$

Let us define for a subset  $S \subset X \times Y$  the following *erasure* transformations:

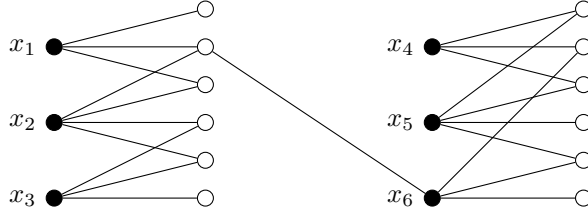
$$\begin{aligned} \mathcal{E}_{1,x}(S) &= S \setminus \{(x, y) \in S : |X(y)| = 1\}, \\ \mathcal{E}_{k,y}(S) &= S \setminus \{(x, y) \in S : |Y(x)| = k\}, \quad k = 1, 2, \end{aligned}$$

and finally  $\mathcal{E} = \mathcal{E}_{2,y} \circ \mathcal{E}_{1,y} \circ \mathcal{E}_{1,x}$ .

**Definition 5.14** (Erased sets and fully erasable sets). *A set  $S$  is called erased if  $\mathcal{E}(S) = S$ . Moreover, it is called fully erasable if  $\mathcal{E}^n(S) \downarrow \emptyset$  as  $n \rightarrow \infty$ , i.e. for all  $(x, y) \in S$  there exists  $n \in \mathbb{N}$  such that  $(x, y) \notin \mathcal{E}^n(S)$  (where by convention we set  $\mathcal{E}^0 = id$ ).*

**Example 5.15** (Hobson and Klimmek [19] “Trinomial tree” (cont’d)). Hobson and Klimmek trinomial tree (as defined in Example 5.7) is fully erasable: indeed  $|Y(x)| = 1$  if  $x \leq a$  or  $x \geq b$ . For  $a < x < b$ , it follows from the definition of the transition probabilities that  $|X(x)| = 1$ . After applying the maps  $\mathcal{E}_{1,x}$  and  $\mathcal{E}_{1,y}$  we are therefore left with binomial meshes, and eventually  $\mathcal{E}(S) = \emptyset$ .

**Example 5.16.** This is a non-trivial example of a fully erasable set  $S$ :



We have the following:

**Proposition 5.17.** *Assume  $S$  is fully erasable. Then the WEP holds for  $S$ .*

*Proof.* Notice that  $S$  is fully erasable if and only if  $S = \cup_{n \geq 0} \mathcal{E}^n(S)^c$  (with the convention  $\mathcal{E}^0 = id$ ). Hence to prove that the WEP holds for  $S$  we can proceed by induction over  $n$  as follows. First  $\mathcal{E}^0(S)^c = \emptyset$ , so the WEP holds for  $n = 0$ . Assume that the WEP holds for  $\mathcal{E}^n(S)^c$  and let us prove that it holds for  $\mathcal{E}^{n+1}(S)^c$  as well. By definition of erasure transformations, any  $(x, y) \in \mathcal{E}^{n+1}(S)^c \setminus \mathcal{E}^n(S)^c = \mathcal{E}^n(S) \setminus \mathcal{E}^{n+1}(S)$  satisfies  $|X(y)| = 1$  or  $|Y(x)| \in \{1, 2\}$ . Hence to extend the WEP from  $\mathcal{E}^n(S)^c$  to the path  $(x, y)$ , we can proceed as in the proof of Lemma 5.13.  $\square$

**Remark 5.18.** Being fully erasable is not a necessary condition for the WEP: one can directly check that the set in Example 5.9 is not fully erasable. On the other hand, it can be proved that it is necessary in some special cases (see Proposition 6.24).

Note that removing a path  $(x, y)$  such that  $|X(y)| = 1$  may prevent the remaining set from being the support of a martingale measure. Indeed, consider the set  $\{(x, y_i) : i = 1, 2\}$  with  $0 < y_1 < x < y_2$ . We shall need in the sequel the following weakenings of the notion of erased set, which do preserve the martingale property:

**Definition 5.19** (1-erased and 2-erased sets). *A set  $S$  is called 1-erased (resp. 2-erased) if  $\mathcal{E}_{1,y}(S) = S$  (resp.  $\mathcal{E}_{2,y} \circ \mathcal{E}_{1,y}(S) = S$ ), or, equivalently, if  $|Y_S(x)| \geq 2$  (resp.  $|Y_S(x)| \geq 3$ ) for all  $x \in S_X$ .*

**Remark 5.20.** According to the definition above, 1-erased or 2-erased sets may have right-hand free paths, i.e. paths  $(x, y)$  such that  $|X(y)| = 1$ , unlike erased sets  $S$  for which any point in  $S_X$  is connected through  $S$  to at least three points in  $Y$ , and any points in  $S_Y$  is connected to at least two points in  $X$  through  $S$ . This implies that erased sets have at least two points in their projection onto  $X$ , and three points on their projections onto  $Y$ .



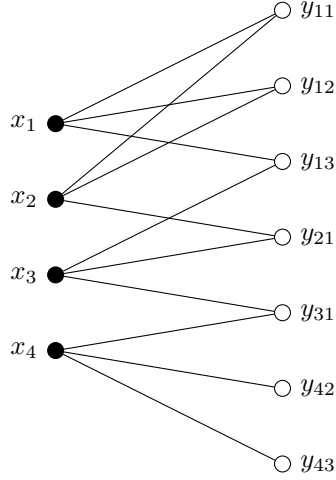
We conclude this section by investigating the relation between the notions of fully erasability and the 2-link property. It turns out that in the finite case they are equivalent, while when  $X$  is infinite one can easily build an example of a support satisfying the latter and not the former.

**Proposition 5.21.** *Assume  $S_X$  is finite. Then  $S$  is fully erasable if and only if it has the 2-link property.*

*Proof.* Let  $|X| = n$  for some nonnegative integer  $n \geq 1$ . Assume that  $S$  satisfies the 2-link property. Hence,  $S$  can be constructed as the union of finitely many sets  $(S_k)_{k=1}^n$  as follows:  $(S_1)_X$  is a singleton and at each further step  $S_k$  is obtained by adding to  $(S_{k-1})_X$  a new point, say  $x_k$ , such that the property (2LP) is fulfilled. Now, starting from the bottom of such a construction, notice that any pair  $(x_n, y)$  can be erased by applying the transformation  $\mathcal{E}$  since the pairs  $(x_n, y)$  with  $|X(y)| = 1$  will be cancelled first and then any other pairs  $(x_n, y)$  will follow since, after the first cancellation, they would satisfy  $|Y(x_n)| \leq 2$ . Iterating  $\mathcal{E}$  will have the same effect on every other pairs  $(x_k, y)$ ,  $y \in Y(x_k)$ ,  $1 \leq k \leq n - 1$ , of the support  $S$ , which will be reduced to the empty set. Hence  $S$  is fully erasable.

Now, assume that  $S$  is fully erasable. Since  $S_X$  is finite,  $S$  is fully erasable if and only if  $\mathcal{E}^n(S) = \emptyset$  for some  $n \geq 1$ . The empty set trivially satisfies the 2-link property. Now, we can proceed by induction. We assume that  $\mathcal{E}^k(S)$  satisfies the 2-link property and we want to prove that  $\mathcal{E}^{k-1}(S)$  does as well. By definition of the erasure transformation  $\mathcal{E} = \mathcal{E}_{2,y} \circ \mathcal{E}_{1,y} \circ \mathcal{E}_{1,x}$ , we have that  $\mathcal{E}^k(S)$  has been obtained by erasing from  $\mathcal{E}^{k-1}(S)$  some pair  $(x, y) \in S$  in the following order: first those satisfying  $|X(y)| = 1$ , second those with  $|Y(x)| = 1$  and finally those having  $|Y(x)| = 2$ . The key observation is that adding them up to  $\mathcal{E}^k(S)$  to go back to  $\mathcal{E}^{k-1}(S)$  transfer the 2-link property to the bigger set  $\mathcal{E}^{k-1}(S)$ . Hence, by the induction principle, we can conclude that  $S = \mathcal{E}^0(S)$  satisfies the 2-link property.  $\square$

**Example 5.22.** Here we show how to construct a set  $S \subset X \times Y$ , with  $X$  and  $Y$  countable subsets of  $\mathbb{R}_+$ , which satisfies the 2-link property and which is not fully erasable. We want the support  $S$  to satisfy  $|Y(x)| \geq 3$  and  $|X(y)| \geq 2$  for all  $(x, y) \in S$ . We start from some  $x_1 \in X$  with  $Y(x_1) = \{y_{11}, y_{12}, y_{13}\}$ . Then we add a second point  $x_2 \in X \setminus \{x_1\}$  with two links with  $x_1$  and one free  $y$ -point attached to it, i.e.  $Y(x_2) = \{y_{11}, y_{12}, y_{21}\}$  for some  $y_{21} \in Y \setminus Y(x_1)$ . We continue the construction in such a way that  $X(y_{21})$  has at least two points in  $X$ . Hence, we add  $x_3 \in X \setminus \{x_1, x_2\}$  with, for instance,  $Y(x_3) = \{y_{13}, y_{21}, y_{31}\}$ . So far, the 2-link property is fulfilled by construction. Now, consider the left-hand free point of  $x_3$ , i.e.  $y_{31}$ , and add a fourth point  $x_4$  such that  $Y(x_4) = \{y_{41}, y_{42}, y_{31}\}$  and so on. By iterating we will eventually get a set  $S$  with the required properties. The next picture illustrates the first four steps of the construction:



## 6 WEP and 2-nets

The goal of this section is to provide a characterization of sets  $S \subset X \times Y$  satisfying the WEP. The main result is stated in Theorem 6.17, which is based on the new notion of 2-net (introduced in Definition 6.6). Intuitively, a 2-net is a subset of  $X \times Y$  where the WEP is defined without ambiguity, i.e. modulo an affine function so that a 2-net has intrinsically the corresponding two degrees of freedom, whence its name.

Now, we introduce the preliminary intuitive notion of connectedness in the following definition. Given a binary relation  $R$  on  $S$ , we recall that the transitive closure of  $R$  is defined as the smallest transitive relation over  $S$  containing  $R$ .

**Definition 6.1.** *Let  $S \subset X \times Y$ . We say that two paths  $(x, y)$  and  $(x', y')$  in  $S$  are neighbours if either  $x = x'$  or  $y = y'$ . The transitive closure of this relation is an equivalence relation. The corresponding equivalence classes of  $S$  will be called connected subclasses (of  $S$ ). A set  $S$  with a single subclass will be called connected.*

This notion of connectedness induces the following decomposition property, which will allow us to work with 1-erased connected sets without loss of generality.

**Proposition 6.2.** *Let  $S \subset X \times Y$  be a 1-erased set and let  $S = \cup_{n \geq 1} S_n$  be its decomposition into mutually disjoint connected subclasses. Then each  $S_n$ , for  $n \geq 1$ , is 1-erased. Moreover, the WEP holds for  $S$  if and only if it holds for each set  $S_n$ ,  $n \geq 1$ .*

*Proof.* Assume that  $S$  is 1-erased and that  $S = \cup_n S_n$  is its decomposition into mutually disjoint connected subclasses. Assume by contradiction that  $S_n$  is not 1-erased for some  $n \geq 1$ . Hence there exists  $(x, y) \in S_n$  such that  $\{y\} = Y_{S_n}(x)$ . Since  $S$  is 1-erased we have  $|Y_S(x)| \geq 2$ , hence there exists a point  $y_0$  in  $Y_S(x) \setminus Y_{S_n}(x)$ . Moreover  $(x, y_0)$  and  $(x, y)$  are neighbours, which implies that  $S_n$  cannot be a connected subclass of  $S$ .

If the WEP holds for  $S$  then it clearly holds for each subclass  $S_n$  as well. Assume now that the WEP holds for every subclass  $S_n$ ,  $n \geq 1$ . Hence every function  $f : X \times Y \rightarrow \mathbb{R}$  satisfies the WEP over each subclass  $S_n$ , i.e.

$$f(x, y) = \varphi_n(x) + h_n(x)(y - x) - \psi_n(y), \quad (x, y) \in S_n, \quad n \geq 1,$$

for some functions  $\varphi_n, h_n : (S_n)_X \rightarrow \mathbb{R}$  and  $\psi_n : (S_n)_Y \rightarrow \mathbb{R}$ . Since the subclasses  $S_n$  are disjoint, we can safely paste such functions together and get the WEP for  $f$  over the whole set  $S$ .  $\square$

## 6.1 $S$ -affine functions and 2-nets: definitions and properties

From now on we will work under the following standing assumption:

**Assumption 6.3.**  $S \subset X \times Y$  is a 1-erased set, i.e.  $\mathcal{E}_{1,y}(S) = S$  or equivalently  $|Y(x)| \geq 2$  for all  $x \in S_X$ .

**Definition 6.4** ( $S$ -affine functions). A function  $\psi : S_Y \rightarrow \mathbb{R}$  is called  $S$ -affine if it is affine on each set  $Y_S(x)$ , i.e., for all  $x \in S_X$  there exist  $\varphi(x), h(x)$  such that

$$\psi(y) = \varphi(x) + h(x)(y - x), \quad \forall y \in Y_S(x). \quad (6.1)$$

$\text{Aff}(S)$  denotes the set of all  $S$ -affine functions.

Basically, an  $S$ -affine function is a function which coincide on every portion  $Y(x)$ ,  $x \in S_X$ , with a truly affine function whose slope and intercept might depend on  $x$ . Obviously, affine functions are  $S$ -affine for any  $S$ . Moreover  $\text{Aff}(S)$  is a vector space.

**Remark 6.5.** Note that the functions  $\varphi, h$  in (6.1) are uniquely defined from  $\psi$  because  $|Y_S(x)| \geq 2$ , which is due to the fact that  $S$  is assumed 1-erased. Indeed, take  $y_1, y_2 \in Y_S(x)$  with  $y_1 \neq y_2$ . Being  $\psi$   $S$ -affine, we have in particular that

$$\psi(y_i) = \varphi(x) + h(x)(y_i - x), \quad i = 1, 2.$$

This is a linear system of two equations with two unknowns  $h(x), \varphi(x)$ , that can be solved explicitly giving

$$h(x) = \frac{\psi(y_1) - \psi(y_2)}{y_1 - y_2}, \quad \varphi(x) = \frac{x - y_2}{y_1 - y_2} \psi(y_1) + \frac{y_1 - x}{y_1 - y_2} \psi(y_2). \quad (6.2)$$

The other important ingredient of this section is the new notion of 2-net, which we introduce in the following definition.

**Definition 6.6** (2-net). A 1-erased set  $A \subset X \times Y$  is called 2-net if every  $A$ -affine function is affine.

The following property follows from the definition of 2-net:

**Proposition 6.7.** *Every 2-net  $A$  is connected.*

*Proof.* Assume by contradiction that  $A$  is not connected, i.e. there exist at least two disjoint connected subclasses, say  $A_1, A_2, \dots$ . Consider some  $A_i$ -affine function  $\psi_i$ , for  $i \geq 1$ . Since  $A_i$  is a 2-net for all  $i$ , we have  $\psi_i(y) = \alpha_i + \beta_i y$ , for all  $y \in (A_i)_Y$ , for some constant  $\alpha_i, \beta_i \in \mathbb{R}$ . Define  $\psi(y) := \sum_{i \geq 1} \psi_i(y) \mathbf{1}_{\{y \in (A_i)_Y\}}$ . This is an  $A$ -affine function, which is not affine. Therefore,  $A$  must be connected.  $\square$

The next lemmas give, respectively, an equivalent characterization of 2-nets and a sort of stability property, according to which adding points to a given 2-net preserves the 2-net structure.

**Lemma 6.8.** *A set  $A$  is a 2-net if and only if for all  $\psi \in \text{Aff}(A)$  as in (6.1) we have*

$$h(x) = h(x'), \quad \forall (x, x') \in (A_X)^2.$$

*Proof.* The direct implication is obvious. Let us prove the other direction. Let  $\psi \in \text{Aff}(A)$  and let  $\beta$  denote the common value of  $h(x)$  for  $x \in A_X$ . Then  $\psi(y) = \varphi(x) + \beta(y - x)$  for some function  $\varphi(x)$ , or yet  $\varphi(x) - \beta x = \psi(y) - \beta y$ . For  $x$  and  $x'$  with  $Y_A(x) \cap Y_A(x') \neq \emptyset$  this yields  $\varphi(x) - \beta x = \varphi(x') - \beta x'$ . Since every two points in  $A$  are connected,  $\varphi(x) - \beta x = \alpha$  on  $A$  for some constant  $\alpha$  and the proof is completed.  $\square$

**Lemma 6.9.** *Let  $A, B$  be two 2-nets such that  $|A_Y \cap B_Y| \geq 2$ . Then  $A \cup B$  is a 2-net.*

*Proof.* Let  $\psi$  be a  $A \cup B$ -affine function. Let  $\alpha + \beta y$  be the affine function matching  $\psi$  on  $A_Y$  and let  $\gamma + \delta y$  be the affine function matching  $\psi$  on  $B_Y$ . Since  $|A_Y \cap B_Y| \geq 2$  we have  $\alpha = \gamma$  and  $\beta = \delta$ , hence  $\psi$  is affine on  $A_Y \cup B_Y = (A \cup B)_Y$ .  $\square$

The following two examples clarify the relation between 2-nets and the 2-link property.

**Example 6.10.** Any subset  $A$  satisfying (2LP) with equality is a 2-net. Indeed, assume that there exists a numbering  $A_X = (x_n)_{n \geq 1}$  such that

$$|Y(x_n) \cap \bigcup_{i=1}^{n-1} Y(x_i)| = 2, \quad n \geq 1. \quad (6.3)$$

To show that  $A$  is a 2-net we proceed by induction. First,  $\{x_1\} \times Y(x_1)$  is trivially a 2-net. Assume now that  $A_{n-1} := \{(x_i, y) : y \in Y(x_i), i = 1, \dots, n\}$  is a 2-net. Since (6.3) holds for all  $n$ , we can apply Lemma 6.9 yielding that  $A_{n-1} \cup (\{x_n\} \times Y(x_n))$  is a 2-net. Therefore,  $A$  is a 2-net.

**Example 6.11.** The support described in Example 5.9 is also a 2-net. Indeed, both sets  $\{x_1, x_2\}$  and  $\{x_3, x_4\}$  are 2-nets and they are connected to each other with exactly two links. Hence, their union is a 2-net by Lemma 6.9.

## 6.2 $S$ -maximal 2-nets

In this section we introduce the notion of  $S$ -maximal 2-net and give some properties that will reveal useful later in this section.

**Definition 6.12** (Maximal 2-net). *A 2-net  $A \subset S$  is  $S$ -maximal if for any 2-net  $A' \subset S$  such that  $A \subset A'$ , we have  $A = A'$ .*

**Proposition 6.13.**  *$S$ -maximal 2-nets exist.*

*Proof.* The existence of maximal 2-nets is guaranteed by an application of Zorn's Lemma (e.g. 1.7 in [2]). Let  $\mathcal{A}$  denote the class of all 2-nets in  $S$  and let  $\mathcal{A}'$  any subclass of  $\mathcal{A}$ , totally ordered with respect to set inclusion, i.e. for any two elements  $A'_1, A'_2 \in \mathcal{A}'$  we have either  $A'_1 \subset A'_2$  or  $A'_2 \subset A'_1$ . We need to prove that  $\mathcal{A}'$  has an upper bound in  $\mathcal{A}$ . Consider  $A_0 := \cup_{A' \in \mathcal{A}'} A'$ , which by definition contains any 2-net in  $\mathcal{A}'$ . To conclude, it remains to show that  $A_0$  is a 2-net. In order to do so, take an  $S$ -affine function  $f$ . By definition,  $f$  coincide with an affine function on every 2-net  $A'$  with possibly different intercepts and slopes  $\varphi_{A'}, h_{A'}$ . Consider two 2-nets in  $\mathcal{A}'$ , say  $A'_1, A'_2$ . Since they are totally ordered, we have  $A'_1 \subset A'_2$  or the opposite. Both situations imply  $\varphi_{A'_1} = \varphi_{A'_2}$  and  $h_{A'_1} = h_{A'_2}$ . Therefore, being  $A'_i, i = 1, 2$ , arbitrary, we have that slopes and intercepts of  $f$  are the same on every 2-net  $A' \in \mathcal{A}'$ . Since this is true for all  $S$ -affine functions  $f$ , we conclude that  $A_0$  is a 2-net and Zorn's Lemma applies.  $\square$

**Proposition 6.14.** *Let  $A, B \subset S$  be two  $S$ -maximal distinct 2-nets. The following properties holds:*

(i) *for all  $z \in S_X \setminus A_X$ , we have*

$$|Y(z) \cap Y(A)| \leq 1;$$

(ii)  *$A_X \cap B_X = \emptyset$  and  $|A_Y \cap B_Y| \leq 1$ .*

*Proof.* Property (i) is a direct consequence of Lemma 6.9 and Definition 6.12.

Regarding the properties in (ii): assume that there exists  $z \in A_X \cap B_X$ . Hence  $A \cup B$  is connected since both are and they have a point in common. Take an  $S$ -affine function  $\psi$ . Since both  $A$  and  $B$  are 2-nets, such a function is affine on  $A$  and  $B$  separately, with slopes and intercepts, respectively,  $\varphi_A, h_A$  and  $\varphi_B, h_B$ . Moreover,  $\varphi_A = \varphi_B$  and  $h_A = h_B$  since  $A_X$  and  $B_X$  have the point  $\{z\}$  in common: indeed since every 2-net is 1-erased, we have  $|Y(z) \cap A_Y| \geq 2$  and  $|Y(z) \cap B_Y| \geq 2$ . Therefore,  $\psi$  is affine on  $A \cup B$  and being  $\psi$  arbitrary we have that  $A \cup B$  is a 2-net, so contradicting the assumption that  $A$  and  $B$  are  $S$ -maximal. Hence,  $A_X \cap B_X = \emptyset$ . Now, assume that  $|A_Y \cap B_Y| \geq 2$ . Proceeding as in the proof of Lemma 6.9, we can prove that  $A \cup B$  is a 2-net strictly bigger than both  $A$  and  $B$  since they are disjoint, so contradicting the fact that they are  $S$ -maximal.  $\square$

### 6.3 Saturated 2-nets and the WEP

In this section we study the relation between the WEP and the 2-nets of  $S$ . By definition, the WEP for a given function  $f$  is defined only up to an  $S$ -affine function. Recall that we say that  $\text{WEP}(f)$  holds on a set  $A$  if we have

$$f(x, y) = \varphi_A(x) + h_A(x)(y - x) - \psi_A(y), \quad (x, y) \in A, \quad (6.4)$$

for some functions  $\varphi_A, h_A, \psi_A$ . Since on 2-nets all  $S$ -affine functions are affine, we have immediately the following proposition, whose proof is straightforward and therefore it is omitted.

**Proposition 6.15.** *Let  $f : X \times Y \rightarrow \mathbb{R}$  be a given function. Assume that  $\text{WEP}(f)$  holds on a 2-net  $A$ . Then the corresponding decomposition (6.4) is defined up to an affine function on  $A$ .*

Let us introduce now the following:

**Definition 6.16.** *We say that a subset  $T \subset S$  is saturated in  $S$  if*

$$\forall (x, y) \in S, x \in T_X \text{ and } y \in T_Y \implies (x, y) \in T$$

In other words, no path can be added to a saturated set.

We recall that in our setting a mesh  $M$  is any subset of  $S$  with  $|M_X| = 1$ . We also use the notation  $M(x) = M$  if  $M_X = \{x\}$ , i.e.  $M(x) = \{x\} \times Y(x)$ . We can finally state the main result of this section.

**Theorem 6.17.** *Let  $S$  be any subset of  $X \times Y$ . If the WEP holds for  $S$ , then any 2-net  $A$  in  $S$  is saturated. Conversely, if every 2-net  $A$  in  $S$  is saturated there exists a sequence of subsets  $S_n \uparrow S$  as  $n \rightarrow \infty$  such that the WEP holds for any  $S_n$ ,  $n \geq 1$ .*

*Proof.* We prove first that the WEP for  $S$  implies the saturatedness of every 2-net  $A$  in  $S$ . Let  $A$  be a 2-net and assume it is not saturated, i.e. there is  $(x_0, y_0) \in A_X \times A_Y \subset S$  and  $(x_0, y_0) \notin A$ . Using the WEP property on  $S$ , let  $(\varphi, \psi, h)$  be the WEP decomposition for the function  $f(x, y) := \mathbf{1}_{(x_0, y_0)}(x, y)$ , i.e.

$$f(x, y) = \varphi(x) + h(x)(y - x) - \psi(y), \quad (x, y) \in S.$$

This implies in particular that

$$\psi(y_0) = \varphi(x_0) + h(x_0)(y_0 - x_0) - 1.$$

Moreover, notice that  $f \equiv 0$  on  $A$ . Hence on  $A$  we have  $\psi(y) = \varphi(x) + h(x)(y - x)$ , so that being  $A$  a 2-net yields that  $\varphi(x) = \alpha + \beta x$  and  $h(x) = \beta$  for some real constant  $\alpha, \beta$ . Therefore we have  $f(x_0, y_0) = \alpha + \beta x_0 + \beta(y_0 - x_0) - \psi(y_0) = 0$ , giving a contradiction.

To complete the proof, we need to show that saturatedness of all 2-nets in  $S$  implies that there exists a sequence of sets  $S_n \uparrow S$ , each of them fulfilling the WEP. By Proposition

6.2, we can assume without loss of generality that  $S$  is connected (cf. Definition 6.1). Let  $f$  be an arbitrary function  $f : X \times Y \rightarrow \mathbb{R}$ . We prove that  $f$  satisfies the WEP using the induction principle as follows. We construct a sequence of subsets  $S_n \subset S$  such that  $S_n \uparrow S$  as  $n \rightarrow \infty$  and the WEP holds on every  $S_n$ . We start from the mesh  $S_1 = M(x)$  for some  $x \in S_X$ . The WEP trivially holds for  $f$  over  $M(x)$ . Now, assume that  $f$  has the WEP on  $S_n$ , i.e.

$$f(x, y) = \varphi_n(x) + h_n(x)(y - x) - \psi_n(y), \quad (x, y) \in S_n,$$

for some functions  $\varphi_n, h_n : (S_n)_X \rightarrow \mathbb{R}$  and  $\psi_n : (S_n)_Y \rightarrow \mathbb{R}$ . Moreover, we assume that  $S_n$  can be decomposed in finitely many  $S_n$ -maximal 2-nets, i.e.  $S_n = \cup_{i=1}^{k_n} A_i$  where  $A_i$  are distinct  $S_n$ -maximal 2-nets. Notice that since the  $A_i$ 's are 2-nets, any  $S_n$ -affine functions  $\xi(y)$  equals  $\alpha_i + \beta_i y$  on each  $A_i$ , and the constant coefficients  $\alpha_i, \beta_i$  satisfy the  $\alpha_i + \beta_i y = \alpha_j + \beta_j y$  on the (possibly empty) intersections  $(A_i)_Y \cap (A_j)_Y$  for all  $i, j$ .

The induction step consists in describing simultaneously how to construct a new set  $S_{n+1} \supset S_n$  (with the same properties as  $S_n$ ) and how to extend the WEP for  $f$  from  $S_n$  to the bigger set  $S_{n+1}$ .

To do so, consider a point  $x' \in S \setminus S_n$ . We want to extend the WEP for  $f$  from  $S_n$  to  $S_{n+1} := S_n \cup M(x')$ . We can discard immediately the case  $Y(x') \cap (S_n)_Y = \emptyset$ . Indeed, in this case the WEP can be trivially extended from  $S_n$  to  $S_n \cup M(x')$ . Therefore, we need to distinguish only between the following two situations.

- (a) Let  $Y(x') \cap (S_n)_Y = \{y_1\}$  for some  $y_1 \in Y$ . Hence we set arbitrarily

$$(\varphi_{n+1}(x'), h_{n+1}(x')) = (\psi_n(y_1), 0)$$

and this choice uniquely determines the values of  $\psi$  on  $Y(x') \setminus (S_n)_Y$ . The mesh  $M(x')$  constitutes a  $S_{n+1}$ -maximal 2-net, and we are done.

- (b) Otherwise, there exist at least two distinct points  $y_1, y_2$  in  $Y(x') \cap (S_n)_Y$ . In this case, define  $(\varphi_{n+1}(x'), h_{n+1}(x'))$  via the formulas (6.2) on the corresponding binomial mesh  $M_{bin}(x') := \{(x', y_i) : i = 1, 2\}$ , so that  $f$  satisfies the WEP on the set  $D := D_{n+1} := S_n \cup M_{bin}(x')$

The key point lies in considering the decomposition of  $D$  in  $D$ -maximal 2-nets before extending the WEP further. Let  $A_1^D, \dots, A_d^D$  be such a decomposition. We can assume without loss of generality that  $M_{bin}(x) \subset A_1^D$  (by relabelling if necessary).

Consider now a new path  $(x, y_3) \in M(x') \setminus M_{bin}(x')$  with right-hand value  $y_3 \in (S_n)_Y$  (if  $y_3$  does not belong to  $(S_n)_Y$ , the WEP for  $f$  can be trivially extended). Since  $A_1^D$  is a 2-net, it is saturated by assumption and so  $y_3$  can not belong to  $(A_1^D)_Y$ . Hence by relabelling again the  $D$ -maximal 2-nets in decomposition of  $D$  (if necessary) we can assume that  $y_3 \in (A_2^D)_Y$ .

In order to extend the WEP for  $f$  to the path  $(x', y_3)$ , we modify the function  $\psi_n$  by adding a  $D$ -affine function  $\chi$  which is null on  $(A_1^D)_Y$  (because  $(\varphi_n(x'), h_n(x'))$  depend on the values of  $\psi_n$  on  $(A_1^D)_Y$ ) and satisfies  $(\psi_n + \chi)(y_3) = \varphi_n(x') + h_n(x')(y_3 - x')$ . Such a function does not exist if and only if every  $D$ -affine function which is null on

$(A_1^D)_Y$  is also null at the point  $y_3$ . If this holds true, then  $\{(x', y_1), (x', y_3)\} \cup A_1^D$  is a 2-net, and therefore it is saturated, which contradicts the existence of the path  $(x', y_2)$ . Adding a  $D$ -affine function to  $\psi$  will not alter the WEP for  $f$  on  $D$ , and therefore  $f$  fulfils the WEP on the extended set  $D \cup \{(x', y_3)\}$  as well.

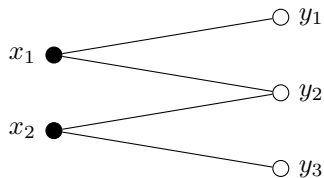
We move on like that, by updating the decomposition in maximal 2-nets before extending the WEP to a new path in  $M(x')$  with right-hand value in  $(S_n)_Y$ . Proceeding this way and after at most  $2k_n$  iterations (by the intersection Lemma 5.1), we get the WEP for  $f$  on the whole set  $S_{n+1}$ . By construction, the decomposition in  $S$ -maximal 2-nets of  $S_{n+1}$  is finite. Therefore, the proof is complete.  $\square$

**Remark 6.18.** Notice that if  $S \subset X \times Y$  is not necessarily 1-erased, we can always apply the main theorem above to the set  $\mathcal{E}_{1,x}(S)$ , which is 1-erased.

**Remark 6.19.** The main theorem above can be reformulated as follows when  $S_X$  is finite: the WEP holds on  $S$  if and only if  $S$  does not have any *non-saturated 2-nets*. Hence, it seems that in our formulation the non-saturated 2-nets play the same role that the cycles have in characterizing extremality of measures with given marginals (without the martingale property as in [12]), i.e. both are forbidden patterns in the supports of their respective extremal measures.

We conclude this section with some simple examples illustrating the content of the main Theorem 6.17.

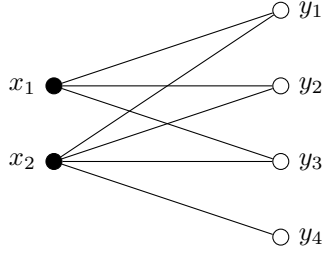
**Example 6.20.** Consider a very simple situation with  $X = \{x_1, x_2\}$ ,  $Y = \{y_1, y_2, y_3\}$  and where the paths in  $S$  are given in the following graph:



This clearly satisfies (2LP), so that WEP holds. On the other hand it is not a 2-net. However, the 2-nets in  $S$  are easily found. They are all the paths  $\{(x_i, y_j)\}$  with  $i = 1, 2$  and  $j = i, i + 1$ , as well as  $A_1 = \{(x_i, y_j) : j = i, i + 1\}$  for  $i = 1, 2$ . All such 2-nets are saturated. Hence Theorem 6.17 applies giving the WEP and as a consequence that any measure  $Q \in \mathcal{M}(\mu, \nu)$  with that set as a support is extremal.

**Example 6.21.** According to Theorem 6.17 above, we have that the WEP holds for a 1-erased set  $S$ , with  $|S_X| < \infty$ , if and only if it does not contain non-saturated 2-nets. An example of non-saturated 2-net is given in the following picture:





**Example 6.22.** The set in Example 5.9 is a maximal 2-net that fulfills the condition in Theorem 6.17. Thus WEP holds. We recall that none of the sufficient conditions previously discussed in Section 5 work in that example.

#### 6.4 Some complementary results on 2-nets and WEP

In this section we gather some consequences and complements of Theorem 6.17. Firstly, we have the following useful byproduct of the first part of Theorem 6.17, which extends the intersection Lemma 5.1 to the situation where one of the image sets  $Y(x_i)$  in its statement is replaced with a 2-net:

**Corollary 6.23** (Extended intersection Lemma). *Assume that the WEP holds for  $S$  and let  $A$  be a 2-net in  $S$ . Then for any  $z \in S_X \setminus A_X$ ,  $|Y(z) \cap A_Y| \leq 2$ .*

*Proof.* Assume the contrary, then the set  $A \cup \{(z, y_i) : i = 1, 2\}$  where the  $y_i$ 's belong to the intersection  $Y(z) \cap A_Y$ , is a 2-net, and therefore is saturated. Hence there can not be a third point in the intersection and we have a contradiction.  $\square$

Using the extended intersection lemma we are now able to prove that when  $|S_Y|$  is small enough, full erasability and the WEP are equivalent:

**Proposition 6.24.** *Assume  $|S_Y| \leq 5$ . Then the WEP holds for  $S$  if and only if  $S$  is fully erasable.*

*Proof.* We already know (cf. Proposition 5.17) that, for a given set  $S$ , full erasability implies the WEP. Hence, it suffices to prove the opposite implication. We can assume, without loss of generality, that  $S$  is 2-erased, so that  $|Y(x)| \geq 3$ , for all  $x \in S_X$ . Let  $x_1 \in S_X$ . We distinguish three different cases.

(i) Case  $|S_Y| = 3$ : Then  $|Y(x_1)| = 3$  and by the intersection Lemma 5.1 there can not be another point in  $S_X$ , so that the paths of  $M(x_1)$  are isolated and  $S$  is fully erasable.

(ii) Case  $|S_Y| = 4$ : If  $|Y(x_1)| = 4$  then we can conclude as above. If  $|Y(x_1)| = 3$ , let  $x_2 \in S_X$ . By the intersection lemma,  $|Y(x_2)| = 3$ ,  $Y(x_2)$  intersects  $Y(x_1)$  in exactly two points. Therefore  $M(x_1) \cup M(x_2)$  is a 2-net and by the extended intersection Lemma 6.23 there can not be another point in  $S_X$ . Now since the mesh  $M(x_2)$  has an isolated branch, i.e. a path  $(x_2, y)$  with  $X(y_2) = \{x_2\}$ , it is erasable, and we can then erase  $M(x_1)$  whose paths are isolated.

(iii) Case  $|S_Y| = 5$ : If  $|Y(x_1)| = 5$ , we can conclude as in case (i) and if  $|Y(x_1)| = 4$  the proof is the same as in the previous case (ii). If  $|Y(x_1)| = 3$ , let  $x_2 \in S_X$ . By the intersection lemma, either  $|Y(x_2)| = 4$  and  $Y(x_2)$  intersects  $Y(x_1)$  in two points, or  $|Y(x_2)| = 3$ . In the former case,  $M(x_1) \cup M(x_2)$  is a 2-net and there can not be another point in  $S_X$ . Hence we conclude as above. In the latter case, we distinguish two sub-cases. If  $Y(x_2)$  intersects  $Y(x_1)$  in two points, then  $M(x_1) \cup M(x_2)$  is a 2-net. Therefore a third point  $x_3 \in S_X$  has the property that  $Y(x_3)$  intersects  $(Y(x_1) \cup Y(x_2))_Y$  in exactly two points, and we have eventually a 2-net, say  $A$ , with  $A_Y = Y$ , so there can not be another point in  $S_X$ . It is readily checked that  $S$  is fully erasable, starting by  $Y(x_3)$ , then  $Y(x_2)$  and  $Y(x_1)$ . If  $Y(x_2)$  intersects  $Y(x_1)$  in one point, then a third point  $x_3 \in S_X$  intersects either  $Y(x_1)$  in two points and  $Y(x_2)$  in one point or the contrary. Therefore we have eventually a 2-net with full  $Y$  projection, so there can not be another point in  $S_X$ . It is readily checked that  $S$  is fully erasable, starting from the meshes  $M(x_1)$  or  $M(x_2)$  whose  $Y$ -projection have only one intersection point with  $Y(x_3)$ , then continuing with  $M(x_3)$  and  $M(x_1)$ .  $\square$

The following proposition is a slight complement to Theorem 6.17. It describes a situation where we can conclude that the WEP holds for an increasing limit of sets  $S_n$ :

**Proposition 6.25.** *Let  $(S_n)_{n \geq 1}$  be an increasing sequence such that the WEP holds for each  $S_n$ , and let  $S = \cup_{n \geq 1} S_n$ . If for each  $n \geq 1$ , any  $S_n$ -affine function is the restriction to  $S_n$  of an  $S_{n+1}$ -affine function, then the WEP holds for  $S$ .*

*Proof.* Let  $f$  be a real-valued function on  $S$  and for any  $n \geq 1$  let  $f_n$  be its restriction to  $S_n$ . Then there is some triplet  $(\varphi_1, h_1, \psi_1)$  such that  $f_1(x, y) = \varphi_1(x) + h_1(x)(y - x) - \psi_1(y)$  on  $S_1$ . Assume by induction that there exists a sequence of triples  $(\varphi_p, h_p, \psi_p)_{1 \leq p \leq n}$ , for  $n \geq 1$ , such that  $\varphi_{p+1}|_{(S_p)_X} = \varphi_p$  for all  $1 \leq p \leq n - 1$  if  $n \geq 2$ , and the same for the other two functions. Then since the WEP holds for  $S_{n+1}$ , we have for some triple  $(\varphi_{n+1}^0, h_{n+1}^0, \psi_{n+1}^0)$

$$f_{n+1}(x, y) = \varphi_{n+1}^0(x) + h_{n+1}^0(x)(y - x) - \psi_{n+1}^0(y) \quad \text{on } S_{n+1}.$$

In particular

$$(\varphi_{n+1}^0 - \varphi_n)(x) + (h_{n+1}^0 - h_n)(x)(y - x) - (\psi_{n+1}^0 - \psi_n)(y) = 0$$

on  $S_n$  and  $q_n := \psi_{n+1}^0|_{(S_n)_Y} - \psi_n$  is an  $S_n$ -affine function. Let  $t_{n+1}$  be an  $S_{n+1}$ -affine function whose restriction to  $S_n$  is  $q_n$ . We have on  $S_{n+1}$

$$t_{n+1}(y) = r_{n+1}(x) + s_{n+1}(x)(y - x)$$

for suitable functions  $r_{n+1}, s_{n+1}$ . Defining

$$\varphi_{n+1} := \varphi_{n+1}^0 - r_{n+1}, \quad h_{n+1} := h_{n+1}^0 - s_{n+1}, \quad \psi_{n+1} := \psi_{n+1}^0 - t_{n+1},$$

yields  $\varphi_{n+1}|_{(S_n)_X} = \varphi_{n+1}^0|_{(S_n)_X} - (\varphi_{n+1}^0|_{(S_n)_X} - \varphi_n)(x) = \varphi_n$ , and in the same way we get  $h_{n+1}|_{(S_n)_X} = h_n$  and  $\psi_{n+1}|_{(S_n)_Y} = \psi_n$ .

It follows that the functions  $\varphi := \lim_{n \rightarrow \infty} \varphi_n$ ,  $h := \lim_{n \rightarrow \infty} h_n$  and  $\psi := \lim_{n \rightarrow \infty} \psi_n$  are well defined on the whole set  $S$ . Indeed, for any  $(x, y) \in S$  there exists  $k \geq 1$  such that  $(x, y) \in S_k$ . Hence  $\lim_{n \rightarrow \infty} \varphi_n(x) = \varphi_k(x)$  and similarly for the other two limits. Finally, we have  $f(x, y) = \varphi(x) + h(x)(y - x) - \psi(y)$  on  $S$ .  $\square$

**Remark 6.26.** The previous proposition provides another proof that the WEP holds for a set satisfying the 2-link property (cf. Proposition 5.4): indeed consider any sequence  $(x_i)_{i \geq 1}$  in  $\mathbb{R}_{++}$  such that (2LP) holds true, so that  $Y(x_{n+1})$  has at most two intersection points with  $(S_n)_Y$  where  $S_n$  is the union of the meshes stemming from the predecessors  $(x_1, x_2, \dots, x_n)$ . Then any  $S_n$ -affine function  $\psi_n$  can be extended to an  $S_{n+1}$ -affine function through some affine function on  $Y(x_{n+1})$  matching  $\psi_n$  on such intersection points. More precisely, let  $y \mapsto \gamma + \delta y$  be an affine function defined on  $Y(x_{n+1})$ . Let us define  $\psi_{n+1}(y) := \psi_n(y)$ , for  $y \in \cup_{i=1}^n Y(x_i)$ , and  $\psi_{n+1}(y) := \gamma + \delta y$ , for  $y \in Y(x_{n+1})$ . To conclude, we just observe that, since  $Y(x_{n+1})$  has up to two points in common with  $(S_n)_Y$ , it is clearly possible to choose  $\delta, \gamma$  so that  $\psi_{n+1}|_{(S_n)_Y} = \psi_n$ .

Finally, we have the following:

**Corollary 6.27.** *Let  $(S_n)_{n \geq 1}$  be an increasing sequence of 2-nets such that the WEP holds for each  $S_n$ , and let  $S = \cup_{n \geq 1} S_n$ . Then the WEP holds for  $S$ .*

*Proof.* The result follows readily from Proposition 6.25 above and from the fact that any affine function on  $(S_n)_Y, n \geq 1$ , is the restriction to  $(S_n)_Y$  of the affine function defined on  $(S_{n+1})_Y$  with the same slope and intercept.  $\square$

## 7 Cycles and extremality

In this section we examine the relation between extremality of a measure  $Q$  in  $(\mu, \nu)$  and existence of cycles in its support. This is motivated by a following well-known geometrical characterization of extremal probabilities (without the martingale property) with given marginals that we have already mentioned in the introduction: let  $\mu, \nu$  be given marginals with countable supports, then a measure  $Q$  with marginals  $\mu, \nu$  is extremal if and only if its support does not contain any cycle. For clarity and for later use, we recall the relevant notion of (classical) cycle in the following

**Definition 7.1.** *Let  $S \subset X \times Y$ . A (classical) cycle  $\mathcal{C}$  in  $S$  is any finite sequence of paths  $\mathcal{C} = (x_i, y_i)_{i=1}^{2n} \subset S$  with  $n \geq 1$ , such that:*

1. *either  $y_{2i} = y_{2i-1}, x_{2i+1} = x_{2i}, x_{2i} \neq x_{2i-1}$ , and  $y_{2i+1} \neq y_{2i}$ , or the same condition with  $x$  and  $y$  interchanged;*
2.  *$x_1 = x_{2n}$  (in which case  $y_1 = y_2$ ) or  $y_1 = y_{2n}$  (in which case  $x_1 = x_2$ ) and  $x_i \neq x_j$  and  $y_i \neq y_j$  for  $1 \leq i \leq j - 3 < 2n - 3$ .*

**Notation 7.2.** Since a cycle is a sequence of paths, there is a natural order along the cycle. For a given path  $(x_i, y_i)$  in the cycle, either  $x_{i+1} \neq x_i$  and we will say that  $(x_i, y_i)$  is an *outgoing* path from  $x_i$ , or  $x_{i+1} = x_i$  and we will say that it is an *incoming* path. We use the convention  $x_{i+1} = x_1$  if  $i = 2n$ . By relabelling if necessary, we can assume without loss of generality that  $(x_1, y_1)$  is an outgoing path from  $x_1$ . Then we can enumerate the cycle starting from  $x_1$  as follows:  $(x_1, y_1), (x_2, y_2), \dots, (x_{2n}, y_{2n})$ , where  $y_1 = y_2$  and  $x_{2n} = x_1$ .

We will ease the notation by denoting a cycle directly by the sequence of its points, where the last one coincide with the first one by convention:  $(x_1, y_2, x_3, y_4, \dots, x_{2n})$ . Note also that a cycle will be identified with its support and orientation: for instance the cycles  $(x_1, y_1, x_2, y_3, x_1)$  and  $(x_2, y_3, x_1, y_1, x_2)$  are the same cycle, and  $(x_1, y_3, x_2, y_1, x_1)$  has the same support, but opposite orientation.

Proofs of the equivalence between extremality of  $Q$  and absence of cycles in the support of  $Q$  can be found in [25, 28]. The main idea is that if such a cycle exists the measure  $Q$  can be perturbed along the cycle while preserving the marginals as follows: let  $\alpha > 0$  be a given parameter, set

$$Q_1(x_i, y_i) = Q(x_i, y_i) + (-1)^i \alpha, \quad Q_2(x_i, y_i) = Q(x_i, y_i) - (-1)^i \alpha, \quad 1 \leq i \leq 2n, \quad (7.1)$$

and  $Q_1(x, y) = Q_2(x, y) = Q(x, y)$  otherwise. Hence, since  $\alpha$  can be chosen sufficiently small so that  $Q_k$ , for  $k = 1, 2$ , are probability measures, we have  $Q_k \in \mathcal{P}(\mu, \nu)$ ,  $k = 1, 2$ , and  $Q = (Q_1 + Q_2)/2$ . Whence  $Q$  is not extremal in  $\mathcal{P}(\mu, \nu)$ .

In this section we investigate to which extent this idea can be exploited in our martingale context. We will introduce first a very natural notion of cycles in our context, *cycles of 2-meshes* and we will end this section by a generalization of this notion in terms of classical cycles.

## 7.1 Cycles of 2-meshes

Let us start by revisiting the proof of Lemma 5.2 (Intersection Lemma under extremality), where we have constructed a perturbation of the initial probability  $Q \in \mathcal{M}(\mu, \nu)$  in the subset  $A = \{(x_i, y_j)\}_{i \in \{1, 2\}, j \in \{1, 2, 3\}}$ . It turns out that this perturbation can be seen in a different perspective. In fact, it can be viewed as a perturbation along a *cycle of 2-meshes*. A given 2-mesh  $M := \{(x_1, y_i) : i = 1, 2\}$  can be clearly seen as an element of the product space  $X \times Y^2$ . With a slight abuse of notation we will sometimes write  $M = (x_1; y_1, y_2)$ . We define then a cycle of 2-meshes in a natural way:

**Definition 7.3.** *A cycle of 2-meshes is a cycle in  $X \times Y^2$ .*

Hence, the decomposition of the set  $A$  in a cycle of 2-meshes is

$$(x_1; y_1, y_2), (x_2; y_1, y_2), (x_2; y_2, y_3), (x_1; y_2, y_3). \quad (7.2)$$

or, using the notation 7.2 applied to product space  $X \times Y^2$ ,  $(x_1, (y_1, y_2), x_2, (y_2, y_3), x_1)$ . Now the key observation is the following: associate to each 2-mesh  $(x; y, y')$  a perturbation of total mass  $\alpha$  dispatched as  $p$  on the path  $(x, y)$  and  $q$  on the path  $(y, y')$ , so that  $\alpha = p + q$ , i.e.

$$Q(x, y) + p, \quad Q(x, y') + q, \quad p + q = \alpha,$$

for  $p, q \in [0, 1]$ . In order for such a perturbation to preserve the martingale property, we impose

$$py + qy' = 0,$$

giving

$$q = \frac{\alpha y}{y - y'}, \quad p = \frac{-\alpha y'}{y - y'},$$

so that given  $\alpha$  there is a unique possible choice for  $p, q$ , which do not depend on the origin point of the 2-mesh. Along the cycle (7.2) of 2-meshes, in order to preserve the mass  $\nu(y)$  at each point  $y \in Y$ , we choose the following sequence of perturbations:  $\alpha, -\alpha, \alpha, -\alpha$ . This also grants that the total mass of the perturbation at each point in  $X$  is zero. Choosing  $\alpha$  small enough, such a procedure leads to a new probability measure, say  $Q^\alpha \in \mathcal{M}(\mu, \nu)$ . Finally, applying a perturbation with opposite signs, i.e.  $-\alpha, \alpha, -\alpha, \alpha$ , we get another probability, say  $\tilde{Q}^\alpha \in \mathcal{M}(\mu, \nu)$ , such that  $Q = (Q^\alpha + \tilde{Q}^\alpha)/2$ . This contradicts the extremality of  $Q$  in  $\mathcal{M}(\mu, \nu)$ .

The construction is exactly the same for a cycle of 2-meshes of any finite length. Therefore, we can sum up what we have just obtained in the following, where we identify a point  $(x, (y_1, y_2))$  in  $X \times Y^2$  with the subset  $\{(x, y_1), (x, y_2)\}$  of  $X \times Y$ .

**Proposition 7.4** (Perturbation along a cycle of 2-meshes). *Let  $Q \in \mathcal{M}(\mu, \nu)$  be extremal. Then the support of  $Q$  does not contains any cycle of 2-meshes.*

A similar implication holds with the WEP replacing extremality:

**Proposition 7.5.** *Let  $S$  be a subset of  $X \times Y$  and assume that the WEP holds for  $S$ . Then  $S$  does not contain any cycle of 2-meshes.*

*Proof.* Assume that the WEP holds for  $S$  and that  $S$  contains a cycle of 2-meshes  $M^1, \dots, M^n$  for some  $n \geq 2$ . Pick any function  $f : S \rightarrow \mathbb{R}$  and let  $(\varphi, h, \psi)$  be a decomposition of  $f$  as in (4.1). Within any 2-mesh  $M^i = (x_i; y_{i,1}, y_{i,2})$ , one has

$$\frac{f(x_i, y_{i,2}) - f(x_i, y_{i,1})}{y_{i,2} - y_{i,1}} = h(x_i) + \frac{\psi(y_{i,2}) - \psi(y_{i,1})}{y_{i,2} - y_{i,1}}, \quad i = 1, \dots, n.$$

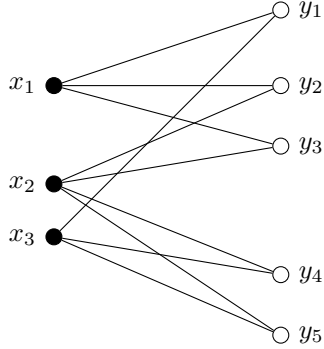
Summing up along the cycle of such 2-meshes, notice that the  $h$  term (resp. the  $\psi$  term) will cancel whenever consecutive 2-meshes have their  $x$ -points (resp. one or more of their  $y$ -points) in common. We get therefore the equality

$$0 = \sum_i (-1)^i \frac{f(x_i, y_{i,2}) - f(x_i, y_{i,1})}{y_{i,2} - y_{i,1}}.$$

Since the function  $f : S \rightarrow \mathbb{R}$  is arbitrary, we get a contradiction.  $\square$

Therefore, absence of cycles of 2-meshes is necessary for both extremality and WEP. At this point it is very natural to ask if the converse statement is also true. Unfortunately, albeit being a natural notion to consider in a martingale setting, it turns out that it is not sufficient for neither extremality nor WEP as the following example shows.

**Example 7.6.** Let  $X = \{x_1, x_2, x_3\}$  and  $Y = \{y_i\}_{i=1,\dots,5}$  be in decreasing order, i.e.  $x_1 > x_2 > x_3 > 0$  and  $y_1 > \dots > y_5 > 0$ . Moreover, we assume  $y_2 = x_1$  and  $y_4 < x_3 < x_2 < y_3$ . Consider the following pattern:



This cannot be the support of an extremal martingale measure with given marginals, since the WEP does not hold. Moreover, it can also be checked by direct inspection that it does not contain any cycles of 2-meshes.

Thus, this seems to imply that the notion of cycle of 2-meshes is too strong. Finding a more general pattern, admitting a perturbation which preserves the marginals and the martingale property, is the topic of the next section.

## 7.2 Perturbation along a pool of cycles

First, we observe that if a probability measure is not extremal in  $\mathcal{M}(\mu, \nu)$ , it is also not extremal in  $\mathcal{P}(\mu, \nu)$ , so there will be a classical cycle in its support (cf. [25, 28]). We know that a one-parameter perturbation is naturally attached to a cycle as in (7.1), in such a way that the marginals are preserved. There is no hope for this single perturbation to preserve *also* the martingale property: Indeed the martingale condition at any point  $x$  in the  $X$ -section of the cycle reads  $\pm\alpha(y_1 - y_2) = 0$  where  $y_1$  and  $y_2$  are the two distinct points in  $Y$  such that the paths  $(x, y_1)$  and  $(x, y_2)$  belong to the cycle. The idea we are going to exploit is that combining many cycles in a suitable way should add sufficiently many degrees of freedom to the perturbation in order to fulfil the martingale property.

### 7.2.1 Revisiting cycles of 2-meshes

As a warm-up, we revisit the notion of cycle of 2-meshes in terms of classical cycles. Consider a cycle of 2-meshes  $M_1, \dots, M_{2n}$ , for some  $n \geq 1$ , as in Definition 7.3. Each consecutive pair of 2-meshes  $M_i, M_{i+1}$  with the same  $Y$ -section can be viewed as a classical cycle  $\mathcal{C}_i$  (with length 4), so that a cycle of 2-meshes with length  $2n$  can be seen as a set of  $n$  classical cycles, each with length 4. If we attach a sufficiently small perturbation parameter  $\alpha_i > 0$  to each of these cycles as in (7.1) the marginals will be preserved.

We use the following notation for the points of any cycle  $\mathcal{C}_i$ : its left-hand points are given by  $(\mathcal{C}_i)_X = \{x_i, x_{i+1}\}$ , while the right-hand points are  $(\mathcal{C}_i)_Y = \{y_{i,1}, y_{i,2}\}$ . Notice that the cycle property implies in particular that  $y_{i,2} = y_{i+1,1}$  for all  $i$ .

Let us check if the martingale property is also preserved under the perturbation given by the parameters  $\alpha_i$  as above. We start from a left-hand point  $x_1$  and a cycle  $\mathcal{C}_1$ . The other left-hand point of  $\mathcal{C}_1$  will be  $x_2$ , which is also left-hand point  $x_2$  of the cycle  $\mathcal{C}_2$ . The martingale property at  $x_2$  is fulfilled if and only if

$$-\alpha_1(y_{1,1} - y_{2,1}) + \alpha_2(y_{2,1} - y_{2,2}) = 0,$$

and similarly for the other points. Eventually we obtain the martingale condition at the point  $x_1$  as

$$-\alpha_n(y_{n,1} - y_{n,1}) + \alpha_1(y_{1,1} - y_{2,1}) = 0.$$

The key observation is that this last equation, *because of the cycle property*, is obtained as a sum of the  $n - 1$  previous ones. Indeed:

$$\begin{aligned} 0 &= \sum_{i=1}^n \alpha_i((y_{i,1} - y_{i,2}) - (y_{i,1} - y_{i,2})) \\ &= -\alpha_1(y_{1,1} - y_{2,1}) + \alpha_2(y_{2,1} - y_{2,2}) + \cdots + (-\alpha_n(y_{n,1} - y_{n,1}) + \alpha_1(y_{1,1} - y_{2,1})). \end{aligned}$$

We get therefore a system of  $n - 1$  equations with  $n$  unknowns, which is readily solved by induction in this case, taking for instance  $\alpha_1$  as free parameter. Therefore, we have obtained a perturbation preserving both the marginals and the martingale property.

## 7.2.2 Generalization to arbitrary cycles

We can now generalize the previous pattern to cycles of any length, in the following way: consider  $n$  classical cycles  $\mathcal{C}_i$  with  $i = 1, \dots, n$ , each of arbitrary length, with the property that the union of the  $X$ -sections of the cycles contains exactly  $n$  distinct points  $x_1, \dots, x_n$ , i.e.  $\bigcup_{i=1}^n (\mathcal{C}_i)_X = \{x_1, \dots, x_n\}$ . Let  $\gamma_{i,j} = y_{i,j} - y_{i,j+1}$  be the difference between the right-hand point of the outgoing path from  $x_j$  and the right-hand point of the incoming path to  $x_j$  along the cycle  $\mathcal{C}_i$ .

We start with a statement relating the WEP and a certain pattern of cycles for some subset  $S \subset X \times Y$ :

**Proposition 7.7.** *Assume that a set  $S$  in  $X \times Y$  contains, for some  $n \geq 2$ , a set of  $n$  classical cycles  $\mathcal{C}_i$  such that:*

1.  $|\bigcup_{i=1}^n (\mathcal{C}_i)_X| = n$ ;
2. *the cycles are free, i.e. each cycle  $\mathcal{C}_i$  contains a path which does not belong to any other cycle  $\mathcal{C}_k$ ,  $k \neq i$ .*

*Then the WEP does not hold for  $S$ .*

*Proof.* Assume that  $\text{WEP}(f)$  holds for any function  $f$ . Let  $h_j$  be the coefficient of  $(y - x)$  in the WEP decomposition attached to the point  $x_j$ . We have along each cycle  $\mathcal{C}_i$ , with the notations above:

$$\sum_j h_j \gamma_{i,j} = \hat{f}_i$$

where  $\hat{f}_i$  is the sum of the values of  $f$  along the paths of the cycles, counted with a positive sign if path goes from  $X$  to  $Y$  along the cycle, and with a negative sign otherwise. By the cycle property we have  $\sum_j \gamma_{i,j} = 0$  so the identity vector belongs to the kernel of the matrix  $\Gamma = (\gamma_{i,j})$ . By the rank theorem, the image of the matrix  $\Gamma$  is of dimension at most  $n - 1$ . It remains to observe that the assumptions 2 in the statement (i.e. cycles are free) implies that the vectors  $(\hat{f}_i)_{1 \leq i \leq n}$  when  $f$  varies in the set of all real-valued functions defined on  $X \times Y$ , span the whole  $\mathbb{R}^n$ . Therefore, for functions  $f$  such that the vector  $\hat{f}_i$  does not belong to the image of  $\Gamma$ , the above relation does not hold, whence a contradiction.  $\square$

Let us now go back to the construction of a martingale perturbation. Attach to each cycle  $\mathcal{C}_i$  a perturbation as in (7.1) with parameter  $\alpha_i$ . So a (classical) perturbation associated to the vector  $\alpha_1, \dots, \alpha_n$  can be built along the  $n$  cycles  $\mathcal{C}_1, \dots, \mathcal{C}_n$  by choosing sufficiently small parameters  $\alpha_i$ .

Let us investigate now, exactly as above, the martingale conditions at the left-hand points  $x_j$ . The contribution of the cycle  $\mathcal{C}_i$  to the martingale condition at point  $x_j$  will be  $\alpha_i \gamma_{i,j}$  where the classical cycle condition entails  $\sum_j \gamma_{i,j} = 0$ , for each  $i$ , and the martingale condition at the point  $x_j$  reads  $\sum_i \alpha_i \gamma_{i,j} = 0$ . Exactly as in Section 7.2.1 we have therefore

$$0 = \sum_i \alpha_i \sum_j \gamma_{i,j} = \sum_j \sum_i \alpha_i \gamma_{i,j} \tag{7.3}$$

so that the martingale condition at any point  $x_j$  is entailed by the martingale conditions at all the other left-hand points.

We are left with  $n - 1$  equations for  $n$  unknown, and by the rank theorem the solution is a vector space of dimension at least 1, so that by taking a sufficiently small element in this space we get a perturbation preserving the martingale property. Now it remains to prove that this perturbation is not zero.<sup>3</sup> So we need an additional hypothesis, which is given by assumption 2 in the previous proposition (i.e. freeness of cycles): it guarantees indeed that any non-zero solution vector of 7.3 is associated to a non-zero perturbation, since for each index  $i$  there is a path which is perturbed by  $\alpha_i$  only, and not by a linear combination of the components of  $\alpha$ . We have just proved the following

**Proposition 7.8.** *Let  $Q \in \mathcal{M}(\mu, \nu)$ . Assume that the support of  $Q$  satisfies the assumptions of Proposition 7.7 with free cycles  $\mathcal{C}_1, \dots, \mathcal{C}_n$ . Then  $Q$  is not extremal.*

<sup>3</sup>A simple example where this would happen is given by a set of two cycles with twice the same cycle: our approach would lead to a single equation in two unknowns  $\alpha_1, \alpha_2$ , with a one dimensional solution space given by  $\alpha_1 + \alpha_2 = 0$ . The resulting perturbation in this case is the sum of the perturbations  $\alpha_1$  and  $\alpha_2$  along the cycle, hence the zero perturbation.



**Example 7.9.** *The pattern in Example 7.6 satisfies the hypotheses of Propositions 7.7 and 7.8: the three (classical) cycles can be taken as*

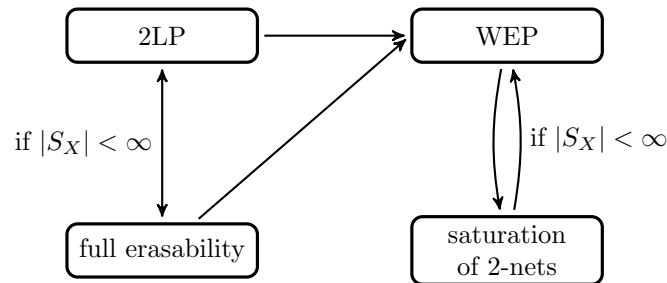
$$(x_1, y_2, x_2, y_3, x_1), \quad (x_2, y_4, x_3, y_5, x_2), \quad (x_1, y_1, x_3, y_4, x_2, y_2, x_1).$$

*It can also be checked that in all the finite examples of extremal points stated in this paper there are at most  $n - 1$  cycles with  $n$  left-hand points.*

We leave the converse statement, i.e. if  $Q$  is not extremal then there is necessarily such a configuration of cycles in its support, as a conjecture.

## 8 Conclusion

In this paper, motivated by the recent literature in model-free finance, we have investigated the properties of the supports of extremal martingale measures with given marginals. Using the Douglas-Lindenstrauss-Naimark Theorem, we have provided an equivalence between extremality of some martingale measure  $Q$  with given marginals and the denseness in  $L^1(Q)$  of a suitable linear subspace, which has a natural financial interpretation as the set of all semi-static strategies. Furthermore, we have studied the combinatorial properties of the supports of such extremal measures in the countable case. More precisely, we have focused on a pointwise version of the weak PRP, called WEP, which implies the extremality when one of the two marginals has finite support. Then we have introduced three combinatorial properties called “full erasability”, 2LP and “saturation of 2-nets”, and we have proved the following implications (among others):



Moreover, we have also started to study the role of cycles in relation to extremality and identified some forbidden patterns, generalizing the notion of (classical) cycles, in the supports of extremal measures. Many examples have been provided in order to illustrate all those notions and how they differ from each other. Many problems remain open, such as finding a more general notion of cycles whose absence from the support would guarantee extremality (compare to the non-martingale case), the relation with graph theory and, more importantly, to what extent those implications can be extended to the non-countable case, e.g. when the marginals have absolutely continuous densities. They are all left for future research.

## References

- [1] B. Acciaio, M. Larsson, and W. Schachermayer. “The space of outcomes of semi-static trading strategies need not be closed.” arXiv preprint arXiv:1606.00631 (2016), to appear in *Finance and Stochastics*.
- [2] C. D. Aliprantis, K. C. Border. *Infinite Dimensional Analysis*. Springer (1994).
- [3] M. Beiglböck, P. Henry-Labordère, F. Penkner. “Model-independent bounds for option prices: a mass transport approach”. *Finance and Stochastics* 17.3 (2013), 477-501.
- [4] M. Beiglböck, N. Juillet. “On a problem of optimal transport under marginal martingale constraints”. *Annals of Probability* 44(1), (2016), 42–106.
- [5] V. Beneš, J. Štěpán. “The support of extremal probability measures with given marginals”. *Mathematical Statistics and Probability Theory*. Springer Netherlands (1987), 33–41.
- [6] S. Bianchini, L. Caravenna. “On the extremality, uniqueness and optimality of transference plans.” *Bull. Inst. Math. Acad. Sin. (N.S.)*, 4(4), (2009), 353–455.
- [7] G. Birkhoff. “Three observations on linear algebra.” *Univ. Nac. Tucumn. Revista A* 5 (1946), 147-151.
- [8] D. T. Breeden, R. H. Litzenberger. “Prices of state-contingent claims implicit in option prices.” *Journal of Business*, (1978), 621–651.
- [9] L. Campi. “A note on extremality and completeness in financial markets with infinitely many risky assets.” *Rendiconti del Seminario Matematico della Università di Padova* 112 (2004): 181–198.
- [10] L. Campi. “Arbitrage and completeness in financial markets with given  $N$ -dimensional distributions.” *Decisions in Economics and Finance* 27.1 (2004): 57–80.
- [11] C. Dellacherie. “Une représentation intégrale des surmartingales à temps discret”. *Publ. Inst. Statist. Univ. Paris* 17.2 (1968), 1-17.
- [12] J. L. Denny. “The support of discrete extremal measures with given marginals”. *The Michigan Mathematical Journal* 27.1 (1980), 59–64.
- [13] R. Diestel. *Graph theory*. Grad. Texts in Math, Springer (2005).
- [14] R. G. Douglas. “On extremal measures and subspace density”. *The Michigan Mathematical Journal* 11.3 (1964), 243–246.
- [15] A. Galichon, P. Henry-Labordere, N. Touzi. “A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options.” *The Annals of Applied Probability* 24.1 (2014): 312–336.

- [16] P. Henry-Labordere, N. Touzi. “An Explicit Martingale Version of Brenier’s Theorem”. To appear in *Finance and Stochastics*.
- [17] K. Hestir, S.C. Williams. “Supports of doubly stochastic measures”. *Bernoulli* (1995), 217–243.
- [18] D. Hobson. “Robust hedging of the lookback option.” *Finance and Stochastics* 2.4 (1998): 329–347.
- [19] D. Hobson, M. Klimmek. “Robust price bounds for the forward starting straddle”. *Finance and Stochastics*, 19.1 (2015), 189–214.
- [20] D. Hobson, A. Neuberger. “Robust bounds for forward start options.” *Mathematical Finance* 22.1 (2012): 31–56.
- [21] H. G. Kellerer. “Verteilungsfunktionen mit gegebenen Marginalverteilungen”. *Probability Theory and Related Fields*, 3.3 (1964), 247–270.
- [22] A. Kłopotowski, M. G. Nadkarni, K. P. S. Bhaskara Rao. “When is  $f(x_1, x_2, \dots, x_n) = u_1(x_1) + u_2(x_2) + \dots + u_n(x_n)$ ?” *Proceedings of The Indian Academy of Sciences-Mathematical Sciences*, Vol. 113, No. 1. Indian Academy of Sciences (2003).
- [23] A. Kłopotowski, M. G. Nadkarni, K. P. S. Bhaskara Rao. “Geometry of good sets in  $n$ -fold Cartesian product.” *Proceedings of the Indian Academy of Sciences-Mathematical Sciences*, Vol. 114, No. 2. Indian Academy of Sciences (2004).
- [24] J. Jacod, A.N. Shiryaev. “Local martingales and the fundamental asset pricing theorems in the discrete-time case.” *Finance and stochastics*, 2.3 (1998), 259–273.
- [25] G. Letac. “Representation des mesures de probabilité sur le produit de deux espaces dénombrables, de marges données.” *Illinois Journal of Mathematics*, 10.3 (1966), 497–507.
- [26] D. R. Lick, A. T. White. “ $k$ -Degenerate graphs.” *Canadian J. of Mathematics*, 22 (1970), 1082–1096.
- [27] J. Lindenstrauss. “A remark on extreme doubly stochastic measures.” *American Mathematical Monthly* (1965), 379–382.
- [28] H. G. Mukerjee. “Supports of extremal measures with given marginals.” *Illinois Journal of Mathematics*, 29.2 (1985), 248–260.
- [29] M. A. Naimark. “On extremal spectral functions of a symmetric operator.” *Dokl. Akad. Nauk SSSR*. Vol. 54. No. 7 (1946).
- [30] D. Revuz, M. Yor. *Continuous martingales and Brownian motion*. Vol. 293. Springer Science & Business Media, 2013.

- [31] V. Strassen. “The existence of probability measures with given marginals”. *The Annals of Mathematical Statistics* (1965), 423–439.
- [32] C. Villani. *Optimal Transport: Old and New*. Vol. 338. Springer Science & Business Media, 2008.