



Reconciliation of probability measures

Jean Bérard, Nicolas Juillet

► To cite this version:

Jean Bérard, Nicolas Juillet. Reconciliation of probability measures. European Journal of Applied Mathematics, 2019, 30 (6), pp.1300-1310. 10.1017/S0956792518000682 . hal-02316702

HAL Id: hal-02316702

<https://hal.science/hal-02316702>

Submitted on 27 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RECONCILIATION OF PROBABILITY MEASURES

JEAN BÉRARD AND NICOLAS JUILLET

ABSTRACT. We discuss the reconciliation problem between probability measures: given $n \geq 2$ probability spaces $(\Omega, \mathcal{F}_1, \mathbb{P}_1), \dots, (\Omega, \mathcal{F}_n, \mathbb{P}_n)$ with a common sample space, does there exist an overall probability measure $\mathbb{P}$ on $\mathcal{F} = \sigma(\mathcal{F}_1, \dots, \mathcal{F}_n)$ such that, for all i , the restriction of $\mathbb{P}$ to $\mathcal{F}_i$ coincides with $\mathbb{P}_i$? General criteria for the existence of a reconciliation are stated, along with some counterexamples that highlight some delicate issues. Connections to earlier (recent and far less recent) work are discussed, and elementary self-contained proofs for the various results are given.

1. INTRODUCTION

Consider a finite number $n \geq 2$ of probability spaces all built upon the same sample space Ω , and denoted by $(\Omega, \mathcal{F}_1, \mathbb{P}_1), \dots, (\Omega, \mathcal{F}_n, \mathbb{P}_n)$. We ask whether it is possible to *reconcile* these n probability spaces, meaning that there exists a probability measure $\mathbb{P}$ on $\mathcal{F} = \sigma(\mathcal{F}_1, \dots, \mathcal{F}_n)$ such that, for all $1 \leq i \leq n$, the restriction of $\mathbb{P}$ to $\mathcal{F}_i$ coincides with $\mathbb{P}_i$. In such a case, we say that $\mathbb{P}$ provides a *reconciliation* of the probability measures $\mathbb{P}_1, \dots, \mathbb{P}_n$.

This is a natural problem from a modeling perspective, where several probabilistic models may be available, each describing a specific aspect of the situation under study. The question is then the existence of a probabilistic model which simultaneously incorporates the previous specific models into a global one. Scenario aggregation (see e.g. [4]) and coherent belief modeling (see e.g. [1]) are two examples where related (though not equivalent) problems appear.

Note that, in general, the σ -fields $\mathcal{F}_1, \dots, \mathcal{F}_n$ correspond to distinct but not completely unrelated families of events, which leads to additional constraints on a potential reconciliation $\mathbb{P}$ beyond the mere requirement that $\mathbb{P}|_{\mathcal{F}_i} = \mathbb{P}_i$ for all i . For instance, one may have two events $A_i \in \mathcal{F}_i$ and $A_j \in \mathcal{F}_j$ for which $A_i \cap A_j = \emptyset$ with $i \neq j$, so that any potential reconciliation $\mathbb{P}$ should satisfy $\mathbb{P}(A_i \cap A_j) = 0$. As a consequence, even in the simple case where all σ -fields are assumed to be finite, the existence of a reconciliation $\mathbb{P}$ is neither automatic nor a trivial question.

In this paper, we first state a characterization of the existence of a reconciliation in the case of finite σ -fields. For $n = 2$, a very simple criterion is obtained, while the corresponding criterion in the general case $n \geq 2$ looks more complicated. Through a counterexample, we show that as simple a criterion as in the case $n = 2$ cannot be expected to hold in general. We then

discuss the extension of these results to the case of more general (infinite) σ -fields. Through a counterexample, we show that additional conditions are required for such an extension to hold, then give one example of such conditions.

Note that the positive results stated in Theorem 1.1 and Theorem 1.2 can in fact be derived as corollaries of results obtained some decades ago by various authors, as we gradually became aware while completing the present study. The reconciliation problem as formulated above is neither the main motivation nor a notable example in these works, and we believe that the short self-contained proofs provided in the present paper are still of interest. Moreover, the counterexamples we provide highlight some interesting issues that are specific to the reconciliation problem.

The rest of the paper is organized as follows. Part 1.1 is devoted to the statement of the results (positive and negative). Connections with earlier work are discussed in Part 1.2. Finally, proofs of the various results are collected in Section 2.

1.1. Statement of results.

1.1.1. *The finite case.* Throughout this section, the σ -fields $\mathcal{F}_1, \dots, \mathcal{F}_n$ on Ω are assumed to comprise a finite number of events. Extensions to the infinite case are discussed in the next section.

The first result deals with the case of two probability spaces ($n = 2$), where we have the following characterization of when a reconciliation exists.

Theorem 1.1. *Assume that $\mathcal{F}_1$ and $\mathcal{F}_2$ comprise a finite number of events. Then it is possible to reconcile $(\Omega, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega, \mathcal{F}_2, \mathbb{P}_2)$ if and only if*

$$(1) \quad \mathbb{P}_1(E_1) \leq \mathbb{P}_2(E_2) \text{ for every } E_1 \subset E_2 \text{ with } E_1 \in \mathcal{F}_1 \text{ and } E_2 \in \mathcal{F}_2.$$

Note that condition (1) is clearly necessary since, given a reconciliation $\mathbb{P}$, one must have $\mathbb{P}_1(E_1) = \mathbb{P}(E_1) \leq \mathbb{P}(E_2) \leq \mathbb{P}_2(E_2)$ as soon as $E_1 \subset E_2$. The non-trivial part of the theorem lies in the fact that condition (1) is indeed sufficient to ensure the existence of a reconciliation. Also note that, taking complementary sets, condition (1) is equivalent to the symmetric condition $E_2 \subset E_1 \Rightarrow \mathbb{P}_2(E_2) \leq \mathbb{P}_1(E_1)$.

The next result deals with the general case $n \geq 2$. The existence of a reconciliation is characterized in terms of elementary integer-valued measurable functions. Consider integer numbers $m_1 \geq 1, \dots, m_n \geq 1$, and, for all $1 \leq i \leq n$, a family of m_i pairwise disjoint events $E_1^{(i)} \in \mathcal{F}_i, \dots, E_{m_i}^{(i)} \in \mathcal{F}_i$, and a family of m_i integer numbers $d_1^{(i)} \in \mathbb{Z}, \dots, d_{m_i}^{(i)} \in \mathbb{Z}$. Then set

$$(2) \quad f_i = \sum_{\ell=1}^{m_i} d_\ell^{(i)} \cdot \mathbf{1}_{E_\ell^{(i)}} \text{ and } f = \sum_{i=1}^n f_i.$$

Note that each f_i is $\mathcal{F}_i$ -measurable, and that one has

$$(3) \quad \int f_i d\mathbb{P}_i = \sum_{\ell=1}^{m_i} d_\ell^{(i)} \cdot \mathbb{P}_i(E_\ell^{(i)}).$$

Theorem 1.2. *Assume that $\mathcal{F}_1, \dots, \mathcal{F}_n$ comprise a finite number of events. Then it is possible to reconcile $(\Omega, \mathcal{F}_1, \mathbb{P}_1), \dots, (\Omega, \mathcal{F}_n, \mathbb{P}_n)$ if and only if, for all functions f and $f_1, \dots, f_n$ of the form given by (2), one has that*

$$(4) \quad \sum_{i=1}^n f_i = f \geq 0 \Rightarrow \sum_{i=1}^n \int f_i d\mathbb{P}_i \geq 0.$$

The necessity of condition (4) above is easy to see, for, given a reconciliation $\mathbb{P}$, and a non-negative function f , one must have

$$0 \leq \int f d\mathbb{P} = \sum_{i=1}^n \int f_i d\mathbb{P} = \sum_{i=1}^n \int f_i d\mathbb{P}_i.$$

As a consequence, the non-trivial part of Theorem 1.2 is the fact that condition (4) is indeed sufficient for the existence of a reconciliation $\mathbb{P}$.

Note that we insisted on giving a formulation in terms of integer-valued functions (instead of general real-valued functions) since we are interested in having a combinatorial interpretation of the criterion in terms of comparisons between probabilities of sets. Indeed, condition (4) in Theorem 1.1 has a very clear such combinatorial interpretation, and one may hope for an equally clear criterion in the general case $n \geq 2$. Since, in the case $n = 2$, the existence of a reconciliation can be checked by looking at a very specific subset of the conditions appearing in Theorem 1.2 (namely, those involving $m_1 = 1, m_2 = 1, d_1^{(1)} = -1, d_1^{(2)} = 1$), a natural question is whether, in the general case $n \geq 2$, it is still possible to characterize the existence of a reconciliation through a subset of conditions that involve only "small" integer values. Unfortunately, the following result shows that the answer is negative.

Theorem 1.3. *For all $K > 0$, there exists a triple of probability spaces $(\Omega, \mathcal{F}_1, \mathbb{P}_1), (\Omega, \mathcal{F}_2, \mathbb{P}_2), (\Omega, \mathcal{F}_3, \mathbb{P}_3)$, where $\mathcal{F}_i$ comprises a finite number of events for $i = 1, 2, 3$, such that:*

- *no reconciliation exists;*
- *condition (4) is satisfied whenever $|d_\ell^{(i)}| \leq K$ for all i, ℓ .*

Theorem 1.3 shows that, even when $n = 3$, there is no upper bound on how large the integer coefficients $d_\ell^{(i)}$ in (2) have to be in order to check the existence of a reconciliation. Note that, with $K = 1$ (and Theorem 1.1), the theorem also provides an example of a triple of probability spaces among which every pair admits a reconciliation, while no overall reconciliation exists for the triple.

1.1.2. *Extension to the infinite case.* We now consider the case where the σ -fields $\mathcal{F}_i$ on Ω may comprise an infinite number of events. We start with a negative result showing that one cannot extend the previous results to such a general case without additional assumptions.

Theorem 1.4. *There exists a pair $(\Omega, \mathcal{F}_1, \mathbb{P}_1)$, $(\Omega, \mathcal{F}_2, \mathbb{P}_2)$ such that condition (1) is satisfied, but for which no reconciliation exists.*

Remark 1.5. The proof of Theorem 1.1 shows that conditions (1) and (4) are in fact equivalent when $n = 2$. As a consequence, in the counterexample used to prove Theorem 1.4, condition (4) is also satisfied, showing that, in the infinite case, neither (1) nor (4) is sufficient to imply the existence of a reconciliation.

The next theorem states that, under a (reasonably mild and general) additional assumption on the σ -fields $\mathcal{F}_1, \dots, \mathcal{F}_n$, can indeed be extended.

Theorem 1.6. *Assume that the σ -fields $\mathcal{F}_1, \dots, \mathcal{F}_n$ (with $n = 2$ in the case of Theorem 1.1) are of the form $\mathcal{F}_i = \sigma(X_i)$, where X_i is a map from Ω to $\mathbb{R}^{d_i}$ (equipped with the Borel σ -field), with $d_i \geq 1$. Moreover, assume that the set $(X_1, \dots, X_n)(\Omega)$ is a closed set in $\mathbb{R}^{d_1 + \dots + d_n}$. Then the conclusions of Theorems 1.1 and 1.2 hold. Namely, in the case $n = 2$, there exists a reconciliation between $\mathbb{P}_1$ and $\mathbb{P}_2$ if and only if condition (1) holds, and, in the general case $n \geq 2$, there exists a reconciliation between $\mathbb{P}_1, \dots, \mathbb{P}_n$ if and only if condition (4) holds.*

The two following corollaries are immediate by-products of the theorem.

Corollary 1.7. *The conclusions of Theorems 1.1 and 1.2 hold when the σ -fields $\mathcal{F}_1, \dots, \mathcal{F}_n$ are generated by a countable number of atoms.*

Corollary 1.8. *Let V be a closed set of $\mathbb{R}^d \times \mathbb{R}^d$ and μ, ν two probability measures on $\mathbb{R}^d$. Then, using transport terminology, there exists a transport plan π from μ to ν concentrated on V , i.e. a measure $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with marginals μ and ν and $\pi(V) = 1$, if and only if for every Borel sets $A \subset \mathbb{R}^d$ and $B \subset \mathbb{R}^d$*

$$(5) \quad \mu(A) \leq \nu(B) \text{ as soon as } (A \times \mathbb{R}^d) \cap V \subset (\mathbb{R}^d \times B) \cap V.$$

Note that Corollary 1.8 is obtained from Theorem 1.6 by setting $\Omega = V$, with the random variables X_i being given by coordinate functions. Also, note that Theorem 1.4 is based on a construction where V is not closed and the conclusion of Corollary 1.8 fails. Take $\Omega = V = \{(x_1, x_2) \in \mathbb{R}^2; 0 < x_1 < x_2 < 1\}$, and let X_1 and X_2 denote the coordinate functions. Then no reconciliation exists if one asks that both X_1 and X_2 follow the uniform distribution on $]0, 1[$, see the proof of Theorem 1.4.

1.2. Connections with earlier work. Early references bearing directly on Theorem 1.1 are Fréchet [7] and Dall'Aglia [5] (where an unpublished result of Berge is also quoted). There, a general result is proved on the

existence of two-dimensional discrete distributions with prescribed marginals and upper bounds, from which Theorem 1.1 can be derived. Subsequent work by Kellerer (Satz 3.2 in [9]) in a general framework (finite dimensional distributions on abstract spaces) can be used to derive Theorem 1.2. Finally, key ingredients needed to prove Theorem 1.6 can be taken from Theorem 11 in Strassen [14]. We refer to the book [6] for additional references and a more detailed historical perspective on this body of work.

Here, a step-by-step self-contained approach to the proofs is given. Indeed, in our specific framework, Theorem 1.2 stems from a rather straightforward application of Farkas' Lemma. Theorem 1.1 is then derived from Theorem 1.2 through a simple reduction argument. Finally, Theorem 1.6 is what one more or less readily obtains by passing to the limit within a sequence of discrete approximations provided by Theorems 1.1 and 1.2.

To conclude this section, we would like to emphasize the connections of Corollary 1.8 (which is in fact a variant of Theorem 11 in [14], and also related with) with recent developments in optimal transport theory.

Although it appears a little less obvious, condition (5) is necessary in Corollary 1.8 for the same reason why (1) is in Theorem 1.1. Its natural interpretation is not probabilistic, but in terms of transport, as follows. We recall that $\pi(A \times B)$ represents the mass transported from A to B . The constraint on the capacity of the transport plan is given by V : no mass can travel from x to y if $(x, y) \notin V$. Gravel located in A with mass $\mu(A)$ can be displaced according to the capacity transport constraint encoded by V only to the set $B_0 = \text{proj}^2((A \times \mathbb{R}) \cap V)$. The storage size of this (universally measurable) set is $\nu(B_0)$. In order for the transport to be manageable, it has to be larger than $\mu(A)$, which is condition (5). Corollary 1.8 states that under this condition there exists a transport plan that satisfies the capacity constraint. As for the usual Monge–Kantorovich optimal transport problem, the goal in constrained versions of the problem is to obtain information on the minimizers of $\pi \mapsto \iint c d\pi$, where $c : \mathbb{R}^d \times \mathbb{R}^d$ is a lower semicontinuous cost function. In particular, is there a Monge solution, that is in the form $\pi = (\text{Id} \otimes T)_\# \mu$?

We are aware of two works in the transport literature that correspond to this problem. In [8] the special case of a constraint displacement vector has been encoded by $V = \{(x, y) \in \mathbb{R}^{2d} : y - x \in \vec{V}\}$ where $\vec{V} \subset \mathbb{R}^d$ is a closed convex set with some additional properties. The authors look at the shape of optimizers for a quadratic (constraint) cost

$$c(x, y) = \begin{cases} |y - x|^2 & \text{if } y - x \in \vec{V}, \\ +\infty & \text{if } y - x \notin \vec{V}. \end{cases}$$

They prove that, provided μ is absolutely continuous, any solution π^* of the Monge–Kantorovich transport problem is a Monge transport plan $\pi^* = (\text{Id} \otimes T)_\# \mu$ where $T(x) \in x + \vec{V}$, μ -almost surely, and hence, π^* is uniquely determined. This occurs under the assumption made that a transport plan

π with finite cost does exist, i.e., $\iint c d\pi < \infty$ for some admissible π . In this respect, condition (5) is namely required in order for the problem to have finite minimal total cost. Note that (5) reads in this case $\nu(A + \vec{V}) \geq \mu(A)$. The other work is [2] where $\vec{V}$ is the unit Euclidean ball $\mathcal{B}_1(0)$ and c is the so-called relativistic cost

$$c(x, y) = \begin{cases} 1 - \sqrt{1 - |y - x|^2} & \text{if } |y - x| \leq 1, \\ +\infty & \text{if } |y - x| > 1. \end{cases}$$

As c is bounded on its domain, (5) is a necessary and sufficient condition to have finite total cost. The authors also introduce $c_t := c(x/t, y/t)$ and the critical speed $T = T_{\mu, \nu} = \inf\{t \in \mathbb{R}_+ : \text{the total cost is finite for } c_t\}$. Therefore T is also the minimal t such that $\nu(A + \mathcal{B}_t(0)) \geq \mu(A)$ for every Borel set A . Let us finally mention [12, 11] where another Monge-Kantorovich problem under capacity constraint is investigated: a transport plan π is admissible if it possesses a density h on $X \times Y \subset \mathbb{R}^{d_1+d_2}$ that satisfies $0 < h \leq \bar{h}$ for a given $\bar{h}$. In this case the authors of [12] cite (p. 575) two functional criteria for the existence of an admissible transport plan. These are due to Kellerer [9] and Levin [13]. An equivalent set-based criterion due to Kellerer [10, Satz 4.2] directly transposes in the continuous settings the (already mentioned) ones obtained by Fréchet [7] and Dall'Aglio [5] in the discrete setting.

2. PROOFS

We start with the proof of Theorem 1.2, which consists merely in an application of Farkas' Lemma.

Proof of Theorem 1.2. As already noted after the statement of the theorem, condition (4) is clearly necessary for the existence of a reconciliation. We now assume that (4) holds, and prove the existence of a reconciliation.

Denote by H the (finite-dimensional) vector space formed by real-valued functions f on Ω of the form $f = f_1 + \dots + f_n$, where each f_i is an $\mathcal{F}_i$ -measurable real-valued function on Ω . Note that condition (4) is assumed to hold for such functions f when each f_i is integer-valued. Since we consider functions on a finite state space, condition (4) in fact holds for any f in H , as can be seen by approximating each f_i by functions with rational values, which in turn can be written as integer values divided by the g.c.d. of these values.

Now, for any $\omega \in \Omega$, denote by φ_ω the linear form on H defined by $\varphi_\omega(f) = f(\omega)$. On the other hand, denote by θ the linear form on H defined by $\theta(f) = \sum_{i=1}^n \int f d\mathbb{P}_i$. Condition (4) (extended to all functions in H) says that, if $f \in H$ is such that $\varphi_\omega(f) \geq 0$ for all $\omega \in \Omega$, then one has $\theta(f) \geq 0$. Farkas' Lemma then guarantees the existence of a family $(t_\omega)_{\omega \in \Omega}$ of *non-negative* real numbers such that, for all $f \in H$, one has the identity

$$(6) \quad \theta(f) = \sum_{\omega \in \Omega} t_\omega \cdot f(\omega).$$

Applying (6) with e.g. $f_1 \equiv 1$ and $f_2 = \dots = f_n \equiv 0$, we see that the non-negative numbers $(t_\omega)_{\omega \in \Omega}$ sum up to 1, so that we can define $\mathbb{P}(\{\omega\}) = t_\omega$. Given $A_i \in \mathcal{F}_i$, we can apply (6) with $f_i = \mathbf{1}_{A_i}$ and $f_j \equiv 0$ for $j \neq i$, and we deduce that $\mathbb{P}(A_i) = \mathbb{P}_i(A_i)$. $\square$

We now prove Theorem 1.1 from Theorem 1.2, through a suitable reduction argument.

Proof of Theorem 1.1. As already noted after the statement of the theorem, condition (1) is clearly necessary for the existence of a reconciliation. We now assume that (1) holds, and prove the existence of a reconciliation through Theorem 1.2.

Without loss of generality, we start with a function f of the form $f = f_1 + f_2$, with $f_1 = \sum_i \lambda_i \mathbf{1}_{A_i}$ and $f_2 = -\sum_j \mu_j \mathbf{1}_{B_j}$, where the λ_i and μ_j are in $\mathbb{Z}$, and where the $(A_i)_i$ (resp. $(B_j)_j$) is a finite family of events in $\mathcal{F}_1$ (resp. $\mathcal{F}_2$) that forms a finite partition of Ω . In the sequel, the λ_i and μ_j are called the coefficients of f .

Our goal is to establish condition (4) so as to apply Theorem 1.2. So we now have to prove that, if $f \geq 0$, then $\sum_i \lambda_i \mathbb{P}_1(A_i) - \sum_j \mu_j \mathbb{P}_2(B_j) \geq 0$.

Step 1: We may assume that f has coefficients in $\mathbb{N}$: Adding to f the function $0 = C - C = C(\sum_i \mathbf{1}_{A_i}) - C(\sum_j \mathbf{1}_{B_j})$ where $C = \max_{i,j}(|\lambda_i|, |\mu_j|)$, we see that

$$\sum_i (\lambda_i + C) \mathbf{1}_{A_i} - \sum_j (\mu_j + C) \mathbf{1}_{B_j}$$

has non-negative coefficients $\lambda'_i = \lambda_i + C$ and $\mu'_j = \mu_j + C$, while keeping $\sum \lambda'_i \mathbb{P}_1(A_i) + \sum \mu'_j \mathbb{P}_2(B_j) = \sum \lambda_i \mathbb{P}_1(A_i) + \sum \mu_j \mathbb{P}_2(B_j)$, since $\sum C \mathbb{P}_1(A_i) - \sum C \mathbb{P}_2(B_j) = 0$.

Step 2: We may assume that f has coefficients in $\{0, 1\}$: Consider f such that $f \geq 0$, and, given Step 1, assume that the coefficients in f satisfy $\lambda_i, \mu_j \in \mathbb{N}$. We set $\lambda'_i = \max(\lambda_i - 1, 0)$ and $\mu'_j = \max(\mu_j - 1, 0)$, and write $f = g + h$, with $g = \sum \lambda'_i \mathbf{1}_{A_i} - \sum \mu'_j \mathbf{1}_{B_j}$ and $h = \sum (\lambda_i - \lambda'_i) \mathbf{1}_{A_i} - \sum (\mu_j - \mu'_j) \mathbf{1}_{B_j}$. Since $\sum \lambda_i \mathbf{1}_{A_i} \leq \sum \mu_j \mathbf{1}_{B_j}$, we must have, for every (i, j) with $A_i \cap B_j \neq \emptyset$, the fact that $\lambda_i \leq \mu_j$, and therefore that $\lambda'_i \leq \mu'_j$ and $\lambda_i - \lambda'_i \leq \mu_j - \mu'_j$. As a consequence, we have $g \geq 0$ and $h \geq 0$, and g and h have non-negative integer coefficients. Moreover, unless they are already equal to 0 or 1, the coefficients of g and h are strictly smaller than the corresponding coefficients of f . Indeed, if $\lambda_i \geq 2$, one has $\lambda'_i < \lambda_i$ and $\lambda_i - \lambda'_i < \lambda_i$, and similarly for μ_j . As a consequence, after a finite number of iterations, we can write f as a sum of non-negative functions with integer coefficients in $\{0, 1\}$.

Conclusion: For a function $f \geq 0$ with coefficients in $\{0, 1\}$, the difference $\sum_i \lambda_i \mathbb{P}_1(A_i) - \sum_j \mu_j \mathbb{P}_2(B_j)$ is of the form $\mathbb{P}_1(E_1) - \mathbb{P}_2(E_2)$ with $E_2 \subset E_1$, so that condition (1) implies the fact that $\sum_i \lambda_i \mathbb{P}_1(A_i) - \sum_j \mu_j \mathbb{P}_2(B_j) \geq 0$. $\square$