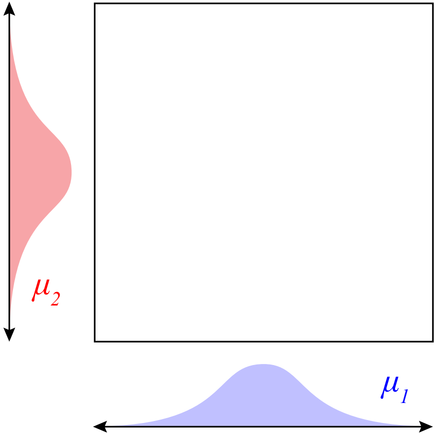
A Strong Duality Principle for Total Variation and Equivalence Couplings

Adam Quinn Jaffe

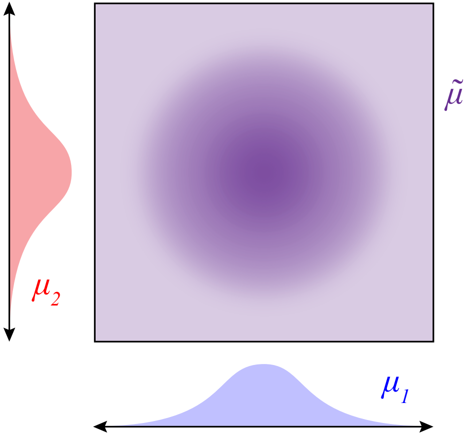
McGill DDC Seminar
February 20, 2024

I. Some vignettes

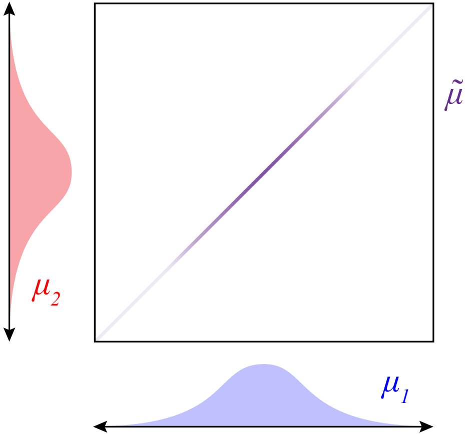
Given probability measures μ_1, μ_2 on $(X, \mathcal{F})$, a *coupling of* μ_1, μ_2 is a probability measure $\tilde{\mu}$ on $(X \times X, \mathcal{F} \otimes \mathcal{F})$ with $\tilde{\mu} \circ \pi_i^{-1} = \mu_i$ for $i = 1, 2$.



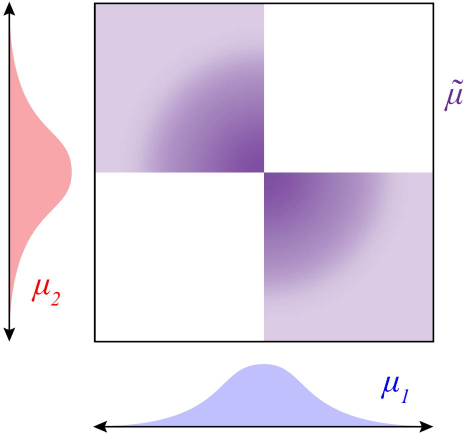
Given probability measures μ_1, μ_2 on $(X, \mathcal{F})$, a *coupling of* μ_1, μ_2 is a probability measure $\tilde{\mu}$ on $(X \times X, \mathcal{F} \otimes \mathcal{F})$ with $\tilde{\mu} \circ \pi_i^{-1} = \mu_i$ for $i = 1, 2$.



Given probability measures μ_1, μ_2 on $(X, \mathcal{F})$, a *coupling of* μ_1, μ_2 is a probability measure $\tilde{\mu}$ on $(X \times X, \mathcal{F} \otimes \mathcal{F})$ with $\tilde{\mu} \circ \pi_i^{-1} = \mu_i$ for $i = 1, 2$.



Given probability measures μ_1, μ_2 on $(X, \mathcal{F})$, a *coupling of* μ_1, μ_2 is a probability measure $\tilde{\mu}$ on $(X \times X, \mathcal{F} \otimes \mathcal{F})$ with $\tilde{\mu} \circ \pi_i^{-1} = \mu_i$ for $i = 1, 2$.



For fixed probability measures μ_1, μ_2, we write $\Pi(\mu_1, \mu_2)$ for the space of all couplings of μ_1 and μ_2, which is a convex subset of $\mathcal{P}(X \times X)$.

Theorem (folklore)

If $(X, \mathcal{F})$ *is a standard Borel space and* $\Delta = \{(x, x) : x \in X\}$ *denotes the diagonal, then for all probability measures* P, P' *on* $(X, \mathcal{F})$*, the following are equivalent:*

(i) $P(A) = P'(A)$ *for all* $A \in \mathcal{F}$.

(ii) *There exists a coupling* $\tilde{P} \in \Pi(P, P')$ *satisfying* $\tilde{P}(\Delta) = 1$.

Exact history is hard to track down (Lindvall 2002).

Theorem (Thorisson 1996)

If G is a locally compact Polish group acting measuably on a Polish space X, with $E_G \subseteq X \times X$ its orbit equivalence relation and $\mathcal{I}_G \subseteq \mathcal{F}$ its invariant σ-algebra, then for all Borel probability measures P, P' on X, the following are equivalent:

(i) $P(A) = P'(A)$ *for all* $A \in \mathcal{I}_G$.

(ii) *There exists a coupling* $\tilde{P} \in \Pi(P, P')$ *satisfying* $\tilde{P}(E_G) = 1$.

Extends prior work for shift invariance (Aldous-Thorisson 1993).

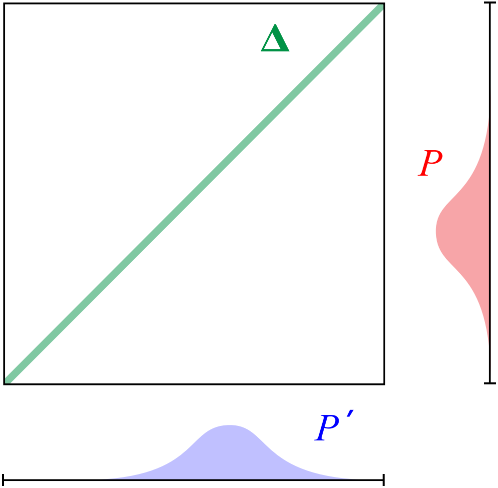
Theorem (Griffeath 1974, Pitman 1976)

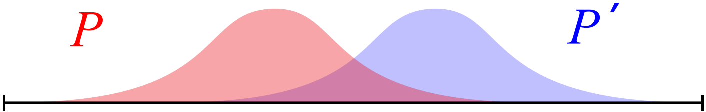
If $E_1 := \bigcup_{n\in\mathbb{N}}\{(x,x') : (x_n, x_{n_1}, \ldots) = (x_n, x_{n+1}, \ldots)\}$ *denotes the equivalence relation of eventual equality and* $\mathcal{T} := \bigcap_{n\in\mathbb{N}} \sigma(x_n, x_{n+1}, \ldots)$ *denotes the tail* σ*-algebra on* $\mathbb{R}^{\mathbb{N}}$*, then for all Borel probability measures* P, P' *on* $\mathbb{R}^{\mathbb{N}}$*, the following are equivalent:*

(i) $P(A) = P'(A)$ *for all* $A \in \mathcal{T}$*.*

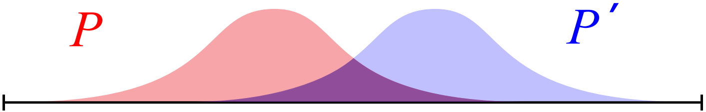
(ii) *There exists a coupling* $\tilde{P} \in \Pi(P, P')$ *satisfying* $\tilde{P}(E_1) = 1$*.*

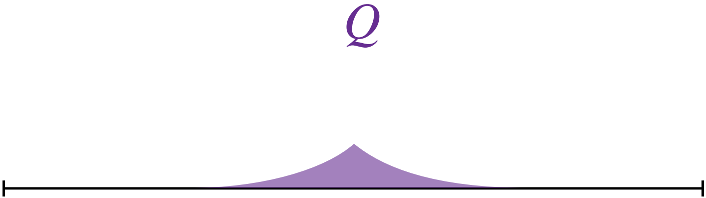
$\tilde{P}$?



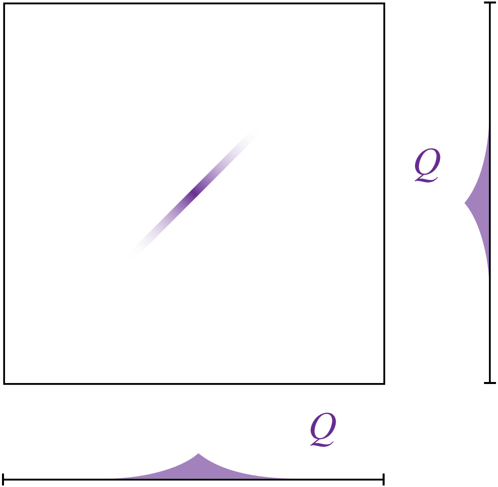

P
P'

$$Q = P \wedge P'$$

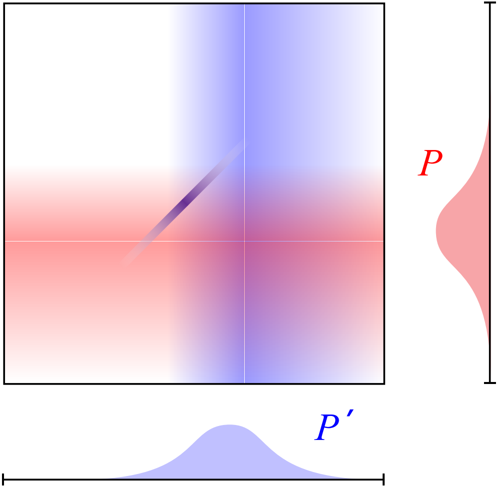



Q

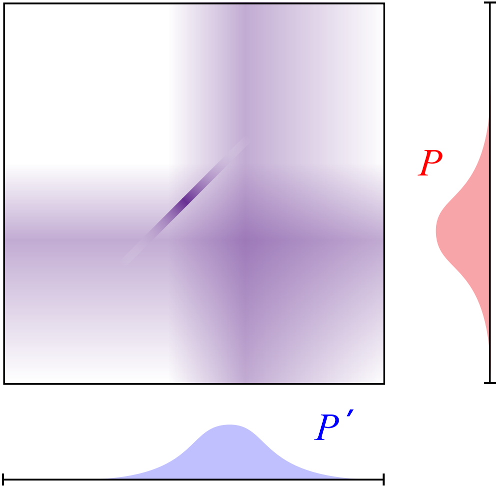
$$\tilde{Q} = Q \circ (i,i)^{-1}$$

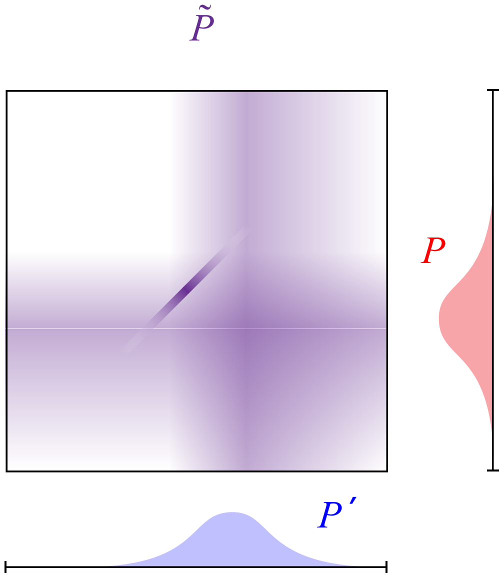


$$\tilde{Q} + \gamma\, M \otimes M'$$



$$\tilde{Q} + \gamma\, M \otimes M'$$




$\tilde{P}$
P
P'

Theorem (folklore)

If $(X, \mathcal{F})$ *is a standard Borel space and* $\Delta = \{(x,x) : x \in X\}$ *denotes the diagonal, then for all probability measures* P, P' *on* $(X, \mathcal{F})$*, the following are equivalent:*

(i) $P(A) = P'(A)$ *for all* $A \in \mathcal{F}$.

(ii) *There exists a coupling* $\tilde{P} \in \Pi(P, P')$ *satisfying* $\tilde{P}(\Delta) = 1$.

Theorem (Thorisson 1996, Georgii 1997)

If G is a locally compact Polish group acting measuably on a Polish space X, with $E_G \subseteq X \times X$ its orbit equivalence relation and $\mathcal{I}_G \subseteq \mathcal{F}$ its invariant σ-algebra, then for all Borel probability measures P, P' on X, the following are equivalent:

(i) $P(A) = P'(A)$ *for all* $A \in \mathcal{I}_G$.

(ii) *There exists a coupling* $\tilde{P} \in \Pi(P, P')$ *satisfying* $\tilde{\mu}(E_G) = 1$.

Theorem (Griffeath 1974, Pitman 1976)

If $E_1 := \bigcup_{n\in\mathbb{N}}\{(x,x') : (x_n, x_{n_1}, \ldots) = (x_n, x_{n+1}, \ldots)\}$ *denotes the equivalence relation of eventual equality and* $\mathcal{T} := \bigcap_{n\in\mathbb{N}} \sigma(x_n, x_{n+1}, \ldots)$ *denotes the tail* σ*-algebra on* $\mathbb{R}^{\mathbb{N}}$*, then for all Borel probability measures* P, P' *on* $\mathbb{R}^{\mathbb{N}}$*, the following are equivalent:*

(i) $P(A) = P'(A)$ *for all* $A \in \mathcal{T}$.

(ii) *There exists a coupling* $\tilde{P} \in \Pi(P, P')$ *satisfying* $\tilde{P}(E_1) = 1$.

Theorem (?)

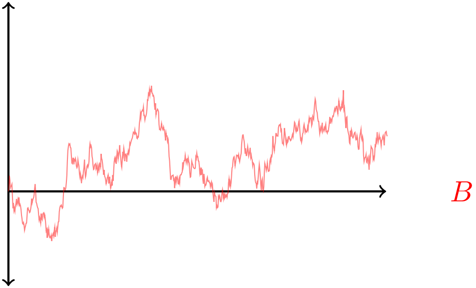
If $(X, \mathcal{F})$ *is a standard Borel space and* $(E, \mathcal{G})$ *is a certain pair where* $E \subseteq X \times X$ *is a sufficiently nice equivalence relation on* X *and* $\mathcal{G}$ *is a sufficiently nice sub-*σ*-algebra of* $\mathcal{F}$ *then for all probability measures* P, P' *on* $(X, \mathcal{F})$*, the following are equivalent:*

(i) $P(A) = P'(A)$ *for all* $A \in \mathcal{G}$.

(ii) *There exists a coupling* $\tilde{P} \in \Pi(P, P')$ *satisfying* $\tilde{P}(E) = 1$.

II. Stochastic processes

Brownian motion is a Borel probability measure W on $C_0([0,\infty);\mathbb{R})$ which in some sense represents the canonical distribution of a random continuous function.



(Universal scaling limit of centered random walks on $\mathbb{R}$.)

Brownian motion with drift $\theta \in \mathbb{R}$ is the Borel probability measure W^θ on $C_0([0,\infty);\mathbb{R})$ given by the pushforward of $\mathbb{W}$ by the map $T^\theta(\{B(t)\}_{t\geq 0}) := \{B(t) + \theta t\}_{t\geq 0}$.

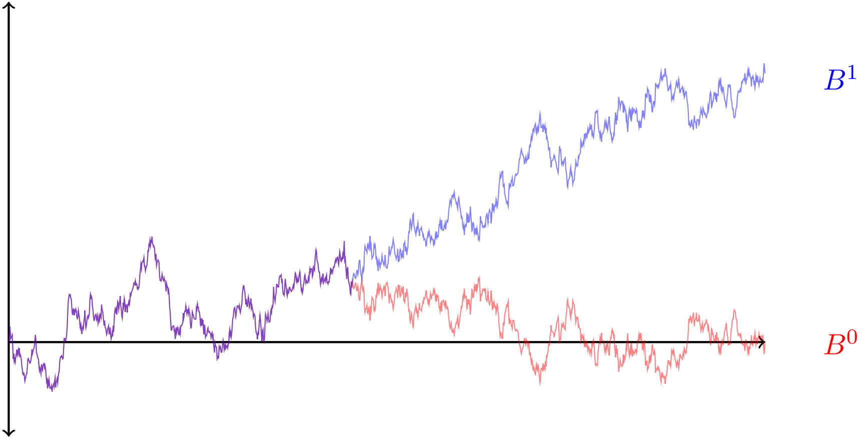
Each T^θ gives rise to a coupling $\tilde{W} \in \Pi(W, W^\theta)$ defined as the pushforward of W by (id, T^θ). But this coupling is not very interesting.

Theorem (Ernst-Kendall-Roberts-Rosenthal 2019)

For any $\theta \in \mathbb{R}$, *there exists* $\tilde{W} \in \Pi(W, W^\theta)$ *satisfying* $\tilde{W}(E_{0+}) = 1$, *where*

$$E_{0+} := \bigcup_{t>0} \{(B, B') : \{B_s\}_{0 \le s \le t} = \{B'_s\}_{0 \le s \le t}\}$$

is called the germ equivalence relation.


B^1
B^0

Definition

Say that a pair of Borel probability measures (P, P') on $C_0([0,\infty);\mathbb{R})$ has the *germ coupling property (GCP)* if there exists $\tilde{P} \in \Pi(P, P')$ with $\tilde{P}(E_{0+}) = 1$.

Definition

Say that a pair of Borel probability measures (P, P') on $C_0([0,\infty);\mathbb{R})$ has the *germ coupling property (GCP)* if there exists $\tilde{P} \in \Pi(P, P')$ with $\tilde{P}(E_{0+}) = 1$.

This is a form of "local equivalence" of two stochastic processes.

Know that (W, W^θ) has the GCP for all $\theta \in \mathbb{R}$.

Which other pairs have the GCP?

I. **Some vignettes**
II. **Stochastic processes**
III. **Problem statement**
IV. **Optimal transport**
V. **Results**

III. Problem statement

Notation:

- $(X, \mathcal{F})$ standard Borel space,
- $\mathcal{P}(X, \mathcal{F})$ space of probability measures on $(X, \mathcal{F})$,
- $\Pi(P, P')$ space of couplings of $P, P' \in \mathcal{P}(X, \mathcal{F})$,
- E equivalence relation on X, and
- $\mathcal{G}$ sub-σ-algebra of $\mathcal{F}$.

Definition

Say $(E, \mathcal{G})$ is *strongly dual* if E is Borel and if we have

$$\max_{A \in \mathcal{G}} \big| P(A) - P'(A) \big| = \min_{\tilde{P} \in \Pi(P,P')} (1 - \tilde{P}(E))$$

for all $P, P' \in \mathcal{P}(X, \mathcal{F})$. (Here, "max" and "min" assert that the supremum and infimum are achieved.)

Which pairs $(E, \mathcal{G})$ are strongly dual?

We have some useful reductions.

Definition

Say $(E, \mathcal{G})$ is *weakly dual* if E is Borel and if we have

$$\max_{A \in \mathcal{G}} \left| P(A) - P'(A) \right| \leq \inf_{\tilde{P} \in \Pi(P, P')} (1 - \tilde{P}(E))$$

for all $P, P' \in \mathcal{P}(X, \mathcal{F})$.

Lemma

If $(E, \mathcal{G})$ is strongly dual then $(E, \mathcal{G})$ is weakly dual.

Definition

Say $(E, \mathcal{G})$ is *quasi-strongly dual* if, for all $P, P' \in \mathcal{P}(X, \mathcal{F})$, the following are equivalent:

(i) $P(A) = P'(A)$ for all $A \in \mathcal{G}$.

(ii) There exists a coupling $\tilde{P} \in \Pi(P, P')$ satisfying $\tilde{P}(E) = 1$.

Lemma

$(E, \mathcal{G})$ is strongly dual iff $(E, \mathcal{G})$ is quasi-strongly dual and E is Borel.

Definition
For an equivalence relation E on X, we write

$$E^* := \{A \in \mathcal{F} : \forall (x, x') \in E (x \in A \Leftrightarrow x' \in A)\}$$

for the *E-invariant σ-algebra.*

Lemma
If $(E, \mathcal{G})$ is strongly dual for some $\mathcal{G}$, then (E, E^*) is strongly dual.

Say that E is *strongly dualizable* if (E, E^*) is strongly dual.

Which Borel equivalence relations E are strongly dualizable?

Definition

For an equivalence relation E on X, we write

$$E^* := \{A \in \mathcal{F} : \forall (x, x') \in E (x \in A \Leftrightarrow x' \in A)\}$$

for the E-*invariant* σ-*algebra.*

Lemma

If $(E, \mathcal{G})$ *is strongly dual for some* $\mathcal{G}$*, then* (E, E^*) *is strongly dual.*

Say that E is *strongly dualizable* if (E, E^*) is strongly dual.

Which Borel equivalence relations E are strongly dualizable?

IV. Optimal transport

Suppose that the supply of some good in $\mathbb{R}^k$ is distributed as P, and that the demand is distributed as P', and that we need to match supply to demand while minimizing the total transport cost.

This can be formulated as the *Monge problem:*

$$\inf_{\substack{T:\mathbb{R}^k\to\mathbb{R}^k \\ P\circ T^{-1}=P'}} \int_{\mathbb{R}^k} \|x - T(x)\|^2 \mathrm{d}P(x).$$

Unfortunately, sometimes solutions do not exist.

Instead one can consider the *Kantorovich problem*

$$\min_{\tilde{P}\in\Pi(P,P')} \int_{\mathbb{R}^k} \|x - x'\|^2 \mathrm{d}\tilde{P}(x, x')$$

which is analytically much nicer:

Properties of Kantorovich problem:

Always has a minimizer

Convex optimization problem

Equivalent to the *Kantorovich dual problem*

$$\max_{\substack{\phi:\mathbb{R}^k\to\mathbb{R}\\ \phi \text{ convex}}} \left(\int_{\mathbb{R}^k} \phi \mathrm{d}P + \int_{\mathbb{R}^k} \phi^* \mathrm{d}P' \right)$$

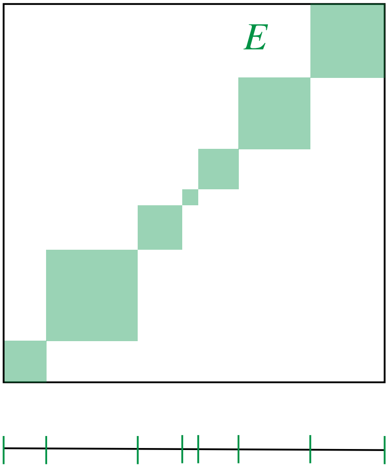
meaning they have the same optimal value.

Many nice properties of the Kantorovich problem hold for a suitable cost function $c : X \times X \to \mathbb{R}$

$$\min_{\tilde{P} \in \Pi(P,P')} \int_X c(x,x') \mathrm{d}\tilde{P}(x,x')$$

on a Polish space X (Rachev-Rüschendorf 1998, Villani 2009).

Notice that our problem is exactly a Kantorovich problem for $c = 1 - \mathbb{1}_E$, meaning: free to move within an equivalence class, and constant cost to move between equivalence classes.


E

If E is closed in $X \times X$, then we can apply existing results on Kantorovich duality to deduce strong duality.

However, existing results require c to be lower semi-continuous which is equivalent to E being closed in $X \times X$, so we cannot generalize past this.

This is not enough, since most of the interesting examples in probability, we need to allow E to be F_σ in $X \times X$.

Need new tools!

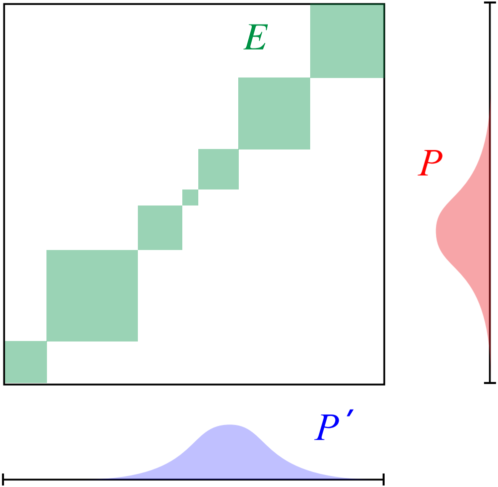
V. Results

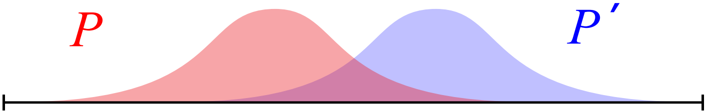
Theorem (AQJ)

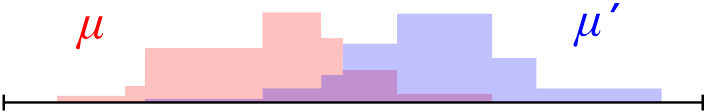
Every smooth equivalence relation is strongly dualizable.

Idea of proof: Do the folklore coupling in (Ω, E^*) then "smooth things over" with conditional expectations.

$\tilde{P}$?

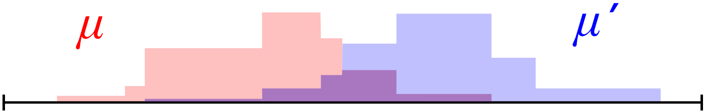



P
P′



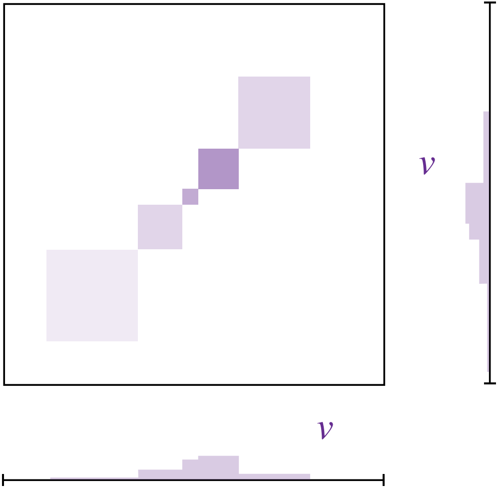
μ
μ'

$$\nu = \mu \wedge \mu'$$

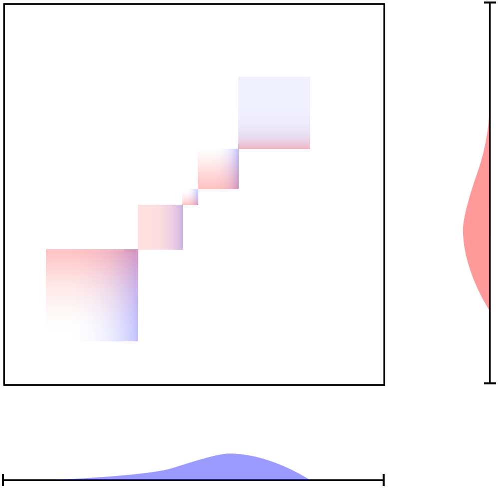


v

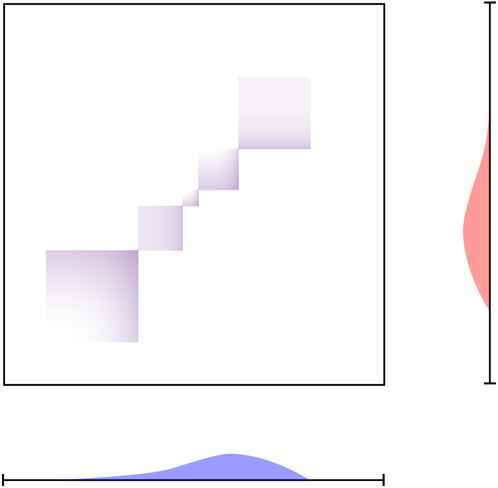
$$v \circ (i, i)^{-1}$$



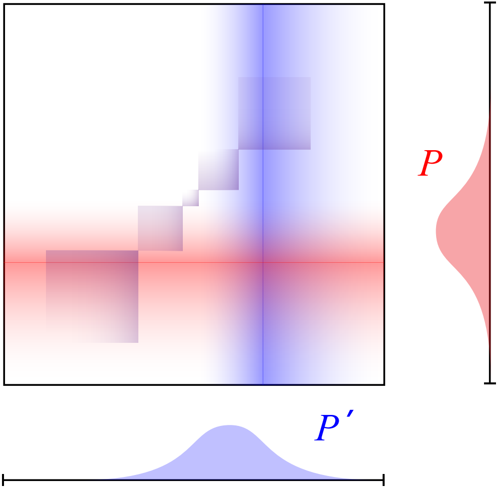
$$\tilde{Q}(\,\cdot \times \cdot \mid E^*) \approx \mu(\,\cdot \mid E^*)\mu'(\,\cdot \mid E^*)$$



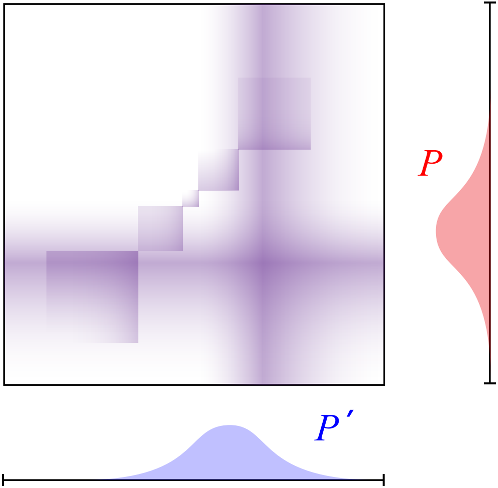
$\tilde{Q}$

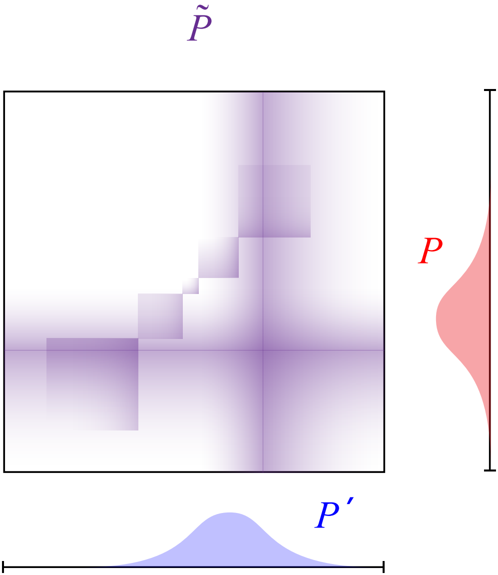


$$\tilde{Q} + \gamma\, M \otimes M'$$



$$\tilde{Q} + \gamma\, M \otimes M'$$




$\tilde{P}$
P
P'

A key step along the way is the following:

Lemma (AQJ)

The following are equivalent:

(i) E *is smooth.*

(ii) E^* *is countably generated.*

(iii) $E \in E^* \otimes E^*$.

Since E_1 is strongly dualizable, we know that smoothness is not necessary.

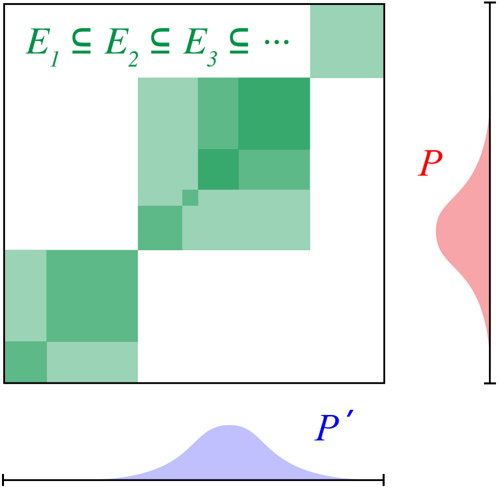
Instead, we have the following closure result:

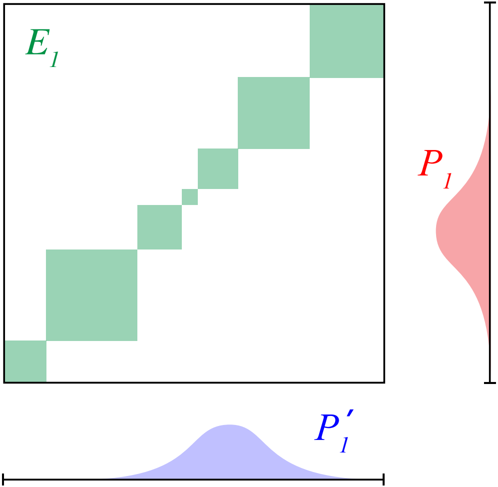
Theorem (AQJ)

A countable increasing union of strongly dualizable equivalence relations is strongly dualizable.

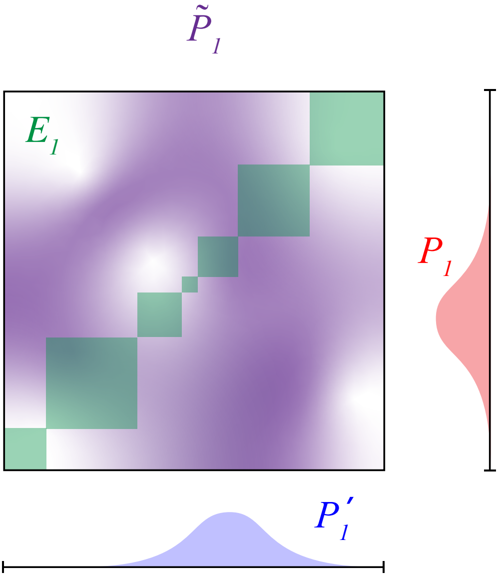
Idea of proof: Apply strong duality, and iterate.

$\tilde{P}$?

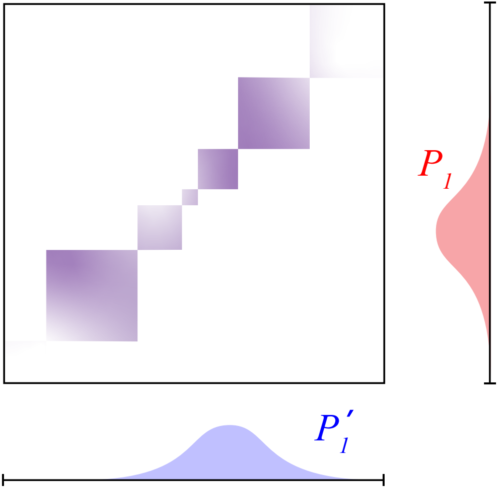


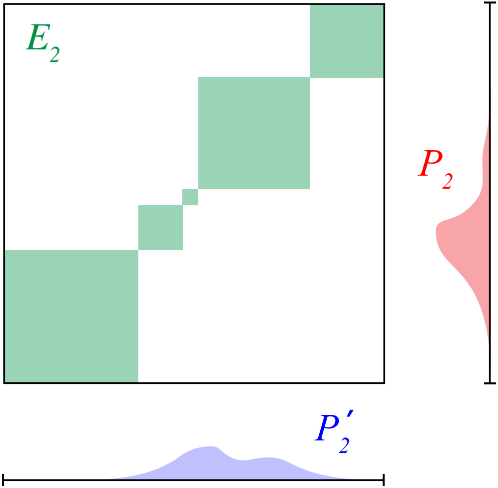


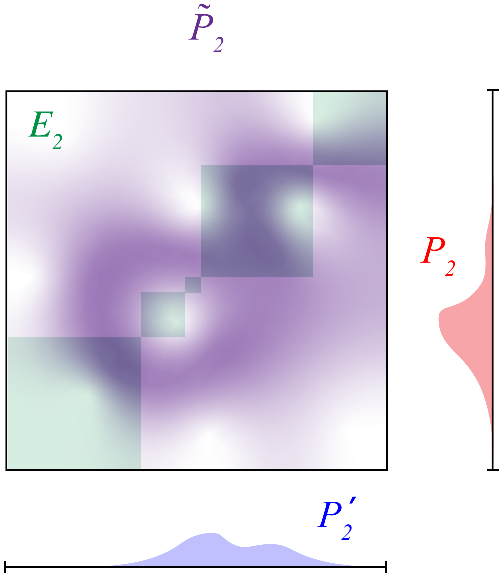
E_1
P_1
P_1'


$\tilde{P}_1$
E_1
P_1
P'_1

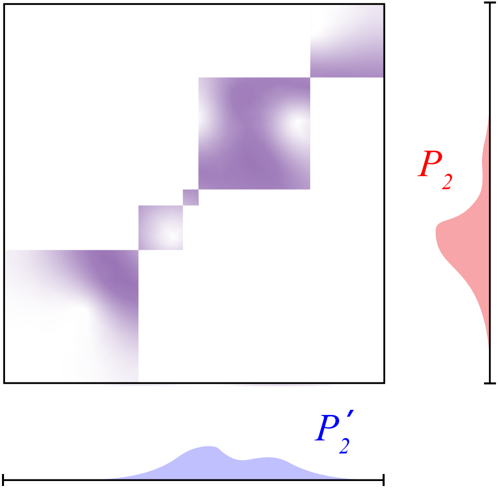
$$\tilde{P}_1(\,\cdot \cap E_1)$$

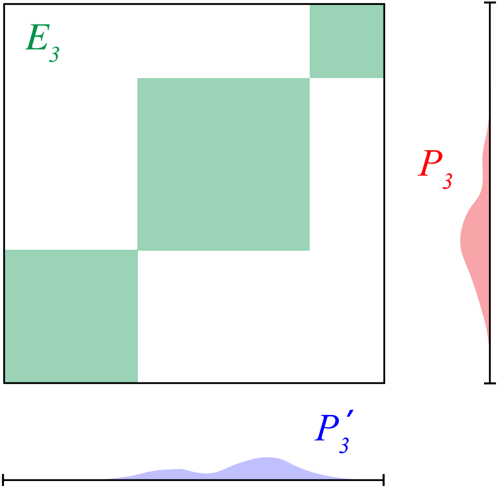



E_2
P_2
P_2'

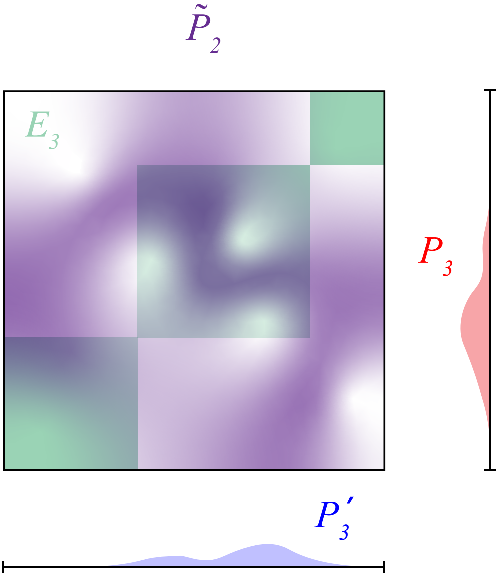

$\tilde{P}_2$
E_2
P_2
P'_2

$\tilde{P}_2(\,\cdot \cap E_2)$

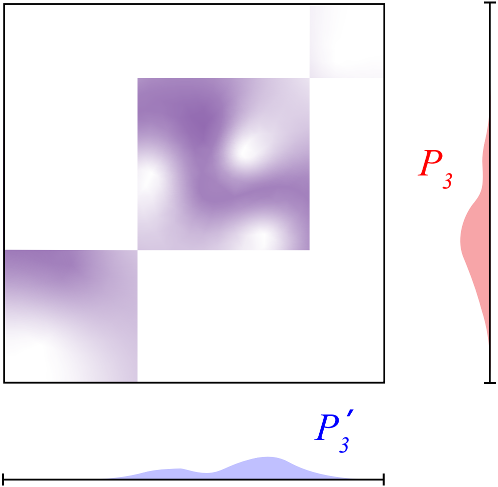




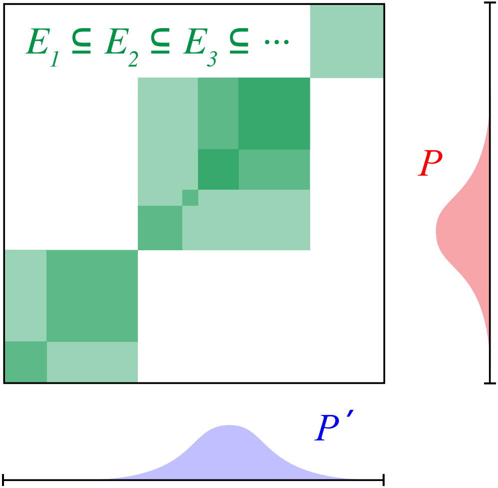
E_3
P_3
P_3'


$\tilde{P}_2$
E_3
P_3
P_3'

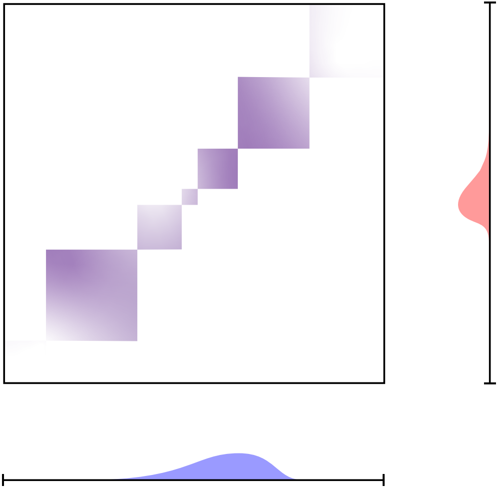
$\tilde{P}_3(\,\cdot \cap E_3)$



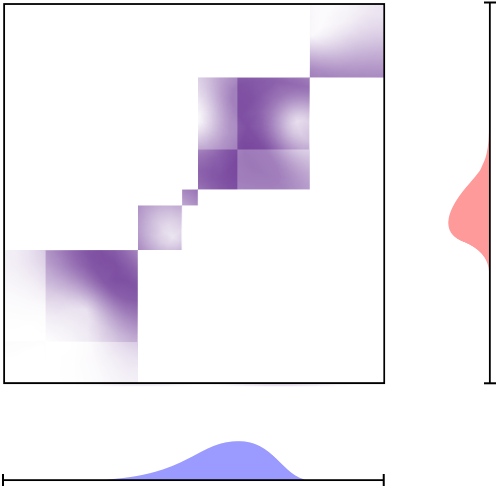
$\tilde{P}$?



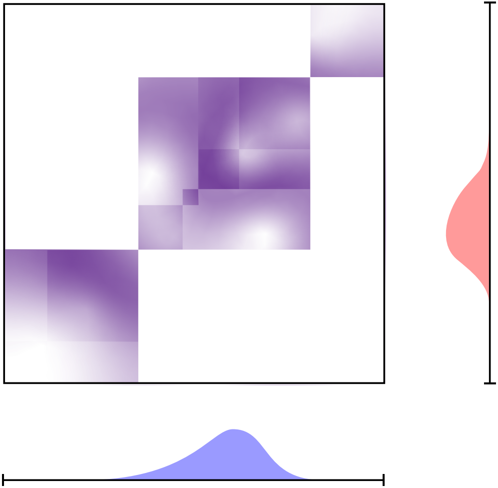
$\tilde{P}_1(\,\cdot \cap E_1)$



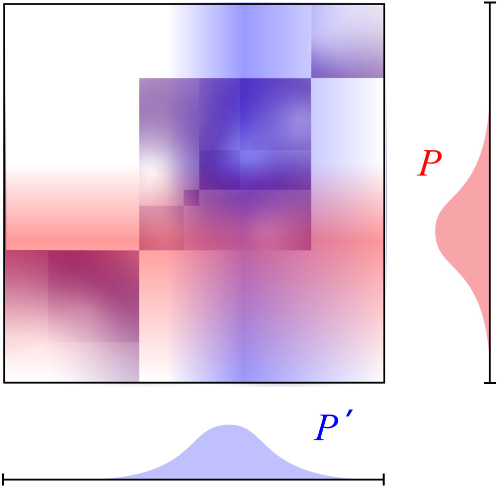
$$\sum_{n=1}^{2} \tilde{P}_n(\cdot \cap E_n)$$



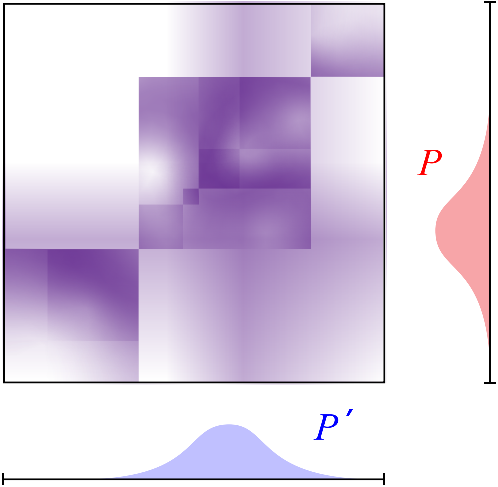
$$\sum_{n=1}^{3} \tilde{P}_n(\,\cdot \cap E_n)$$

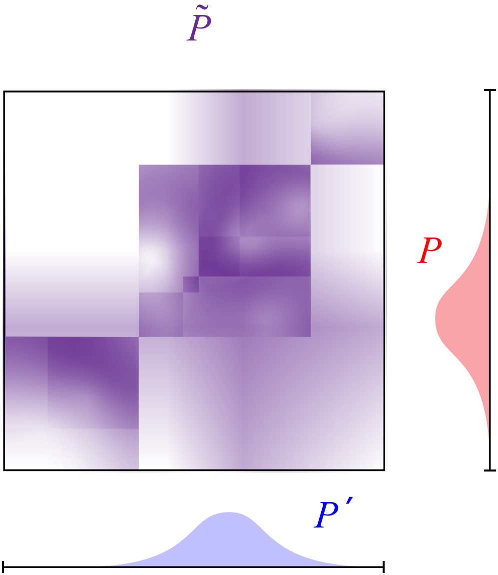


$$\sum_{n=1}^{\infty} \tilde{P}_n(\,\cdot \cap E_n) + \gamma\, M \otimes M'$$



$$\sum_{n=1}^{\infty} \tilde{P}_n(\,\cdot \cap E_n) + \gamma\, M \otimes M'$$




$\tilde{P}$
P
P'

Corollary (AQJ)

Hypersmooth equivalence relations are strongly dualizable.

Corollary (AQJ)

Borel orbit equivalence relations (of locally compact Polish groups acting measurably on standard Borel spaces) are strongly dualizable.

These results establish strong dualizability for most Borel equivalence relations occurring in probability.

Back to our motivation from stochastic processes:

Corollary (AQJ)

A pair of probability measures (P, P') *on* $C_0([0,\infty);\mathbb{R})$ *has the germ coupling property iff* $P(A) = P'(A)$ *for all* $A \in \mathcal{F}_{0+}$*, where*

$$\mathcal{F}_{0+} := \bigcap_{t>0} \sigma(x_s : 0 \leq s < t)$$

is the germ σ*-algebra.*

In stochastic calculus we already have many tools to study the germ σ-algebra. (Blumenthal zero-one law, Girsanov theorem, etc.)

Now we have some intereresting probabilistic consequences:s

Theorem (AQJ)

If $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to [0, \infty)$ are Lipschitz continuous and $x_0 \in \mathbb{R}$ is arbitrary, then the strong solution X of the SDE

$$\begin{cases} dX_t = \mu(X_t)dt + \sigma(X_t)dB_t \text{ for } 0 \leq t \leq 1 \\ X_0 = x_0, \end{cases} \tag{1}$$

forms the GCP with a Brownian motion if and only if $\sigma \equiv 1$ on a neighborhood of x_0.

Open questions:

- Is every Borel equivalence relation strongly dualizable?
- What is the complexity of the GCP relation (or TCP relation) on $\mathcal{P}(C_0([0,\infty);\mathbb{R}))$, or special subsets thereof?
- Is there a useful dual formulation for coupling problems for other relations? (AQJ-Raban 2024+)
- Are there concrete characterizations of the GCP for other classes of Markov processes? (Chu-AQJ 2024+, Hummel-AQJ 2024+)

Thank you!

References

D. Aldous and H. Thorisson. Shift-coupling. *Stochastic Process. Appl.*, 44(1):1-14, 1993.

Y. Chu and A. Q. Jaffe. Nearly-maximal germ couplings of Lévy processes. *in progress*.

P. A. Ernst, W. S. Kendall, G. O. Roberts, and J. S. Rosenthal. MEXIT: Maximal un-coupling times for stochastic processes. *Stochastic Process. Appl.*, 129(2):355-380, 2019.

H.-O. Georgii. Orbit coupling. *Ann. Inst, H. Poincaré Probab. Statist.*, 33(2):253-268, 1997.

S. Goldstein. Maximal coupling. *Z. Wahrsch. Verw. Gebiete.*, 46(2):193-204, 1978.

D. Griffeath. A maximal coupling for Markov chains. *Z. Wahrsch. Verw. Gebiete,* 31:95-106, 1974.

S. Hummel and A. Q. Jaffe. Constructing maximal germ couplings of Brownian motion with drift. *arxiv pre-print*.

A. Q. Jaffe. A strong duality principle for equivalence couplings and total variation. *Electron. J. Probab.*, 28, 1-33.

A. Q. Jaffe. and D. Raban. Kantorovich duality for some topologically degenerate zero-one costs. *in progress*.

T. Lindvall. Lectures on the coupling method. Dover Publications, Inc., Mineola, NY, 2002.

J. W. Pitman. On couplings of Markov chains. *Z. Wahrsch. Verw. Gebiete,* 35(4):315-322, 1976.

L. Pratelli and P. Rigo. Some duality results for equivalence couplings and total variation. *arix pre-print*.

S. Rachev and L. Rüschendorf. *Mass Transportation Problems*. Volume 1 of *Probability and its Applications*. Springer-Verlag, New York, 1998.

H. Thorisson. Transforming random elements and shifting random fields. *Ann. Probab.*, 24(4):2057-2064, 1996.

C. Villani. *Optimal Transport: Old and New*. Volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.