A Strong Duality Principle for Total Variation and Equivalence Couplings

Adam Quinn Jaffe

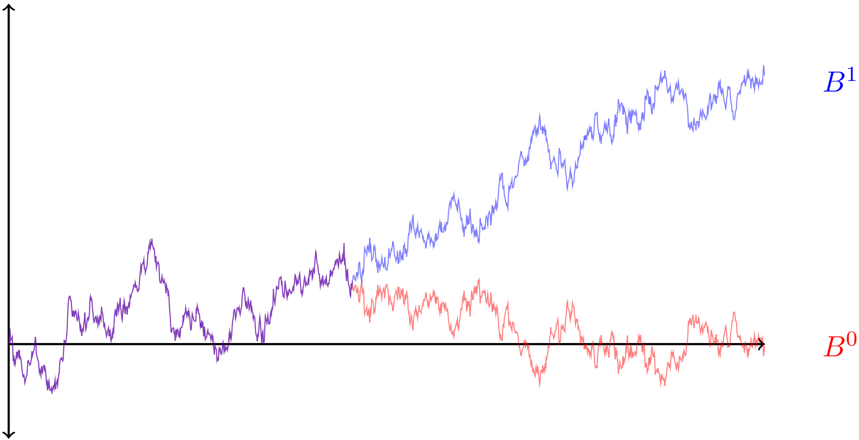
I. Stochastic processes

Theorem (Ernst-Kendall-Roberts-Rosenthal, 2019)

For any $\theta_1, \theta_2 \in \mathbb{R}$, one can construct a probability space supporting

- *a Brownian motion $B^{\theta_1} = \{B_t^{\theta_1}\}_{t \geq 0}$ with drift θ_1,*
- *a Brownian motion $B^{\theta_2} = \{B_t^{\theta_2}\}_{t \geq 0}$ with drift θ_2, and*
- *a random time T with $T > 0$ almost surely,*

such that $B_t^{\theta_1} = B_t^{\theta_2}$ for all $0 \leq t \leq T$ almost surely.


B^1
B^0

In words:

- BM with drift "starts out as" a BM without drift.
- BMs with drift are all "locally equivalent" at time zero.
- The drift of a BM cannot be detected, if it is only observed up to an adversarially-chosen time.

Explicit construction based on Itô excursion theory.

Definition

Say that a pair of Borel probability measures (P, P') on $D([0,\infty);\mathbb{R})$ has the *germ coupling property (GCP)* if one can construct a probability space supporting

- a stochastic process $X = \{X_t\}_{t\geq 0}$ with law P,
- a stochastic process $X' = \{X'_t\}_{t\geq 0}$ with law P', and
- a random time T with $T > 0$ almost surely,

such that $X_t = X'_t$ for all $0 \leq t \leq T$ almost surely.

Know that $(W^{\theta_1}, W^{\theta_2})$ has the GCP for all $\theta_1, \theta_2 \in \mathbb{R}$, where W^θ denotes the law of BM with drift $\theta \in \mathbb{R}$.

Say P has the *Brownian GCP* if (P, W^0) has the GCP.

Which other pairs have the GCP?

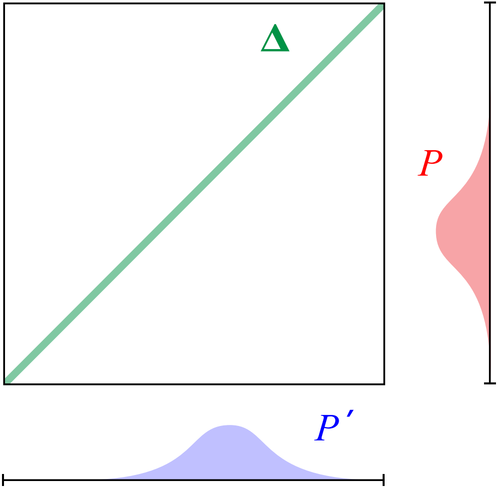
II. Some vignettes

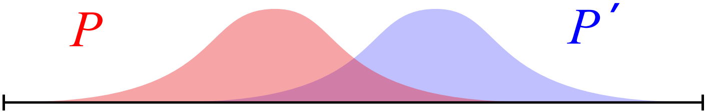
- Ω a Polish space,
- $\Delta := \{(x, x) \in \Omega \times \Omega : x \in \Omega\}$ the diagonal in $\Omega \times \Omega$,
- P, P' two Borel probability measures on Ω, and
- $\Pi(P, P')$ the space of all couplings of P and P'.

Then (folklore):

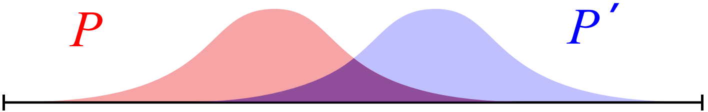
$$\sup_{A \in \mathcal{B}(\Omega)} \left| P(A) - P'(A) \right| = \min_{\tilde{P} \in \Pi(P,P')} (1 - \tilde{P}(\Delta)).$$

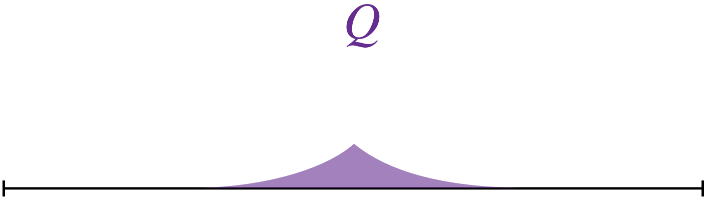
$\tilde{P}$?



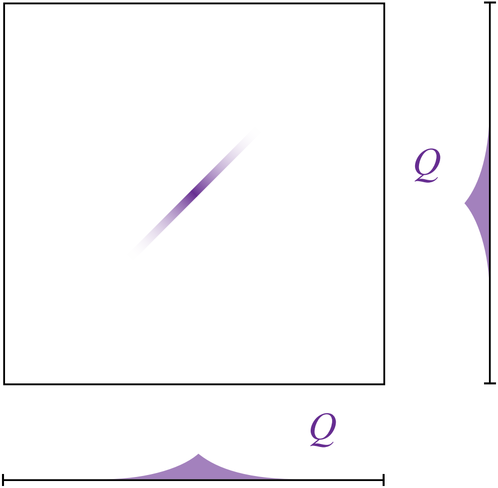

P
P′

$$Q = P \wedge P'$$

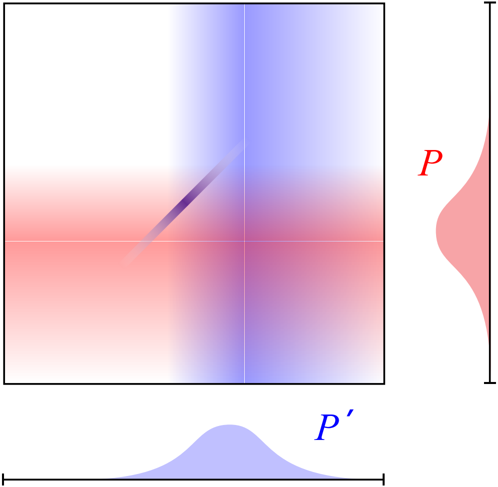



Q

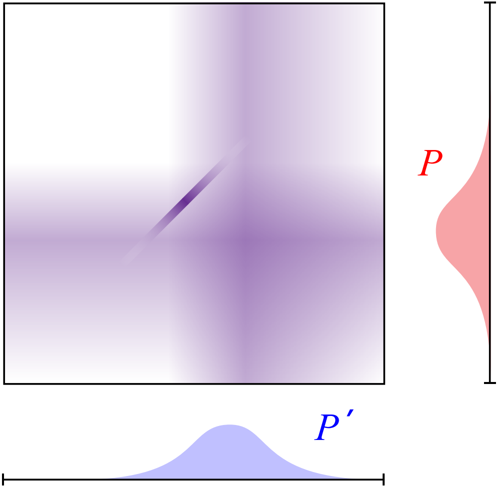
$$\tilde{Q} = Q \circ (i,i)^{-1}$$

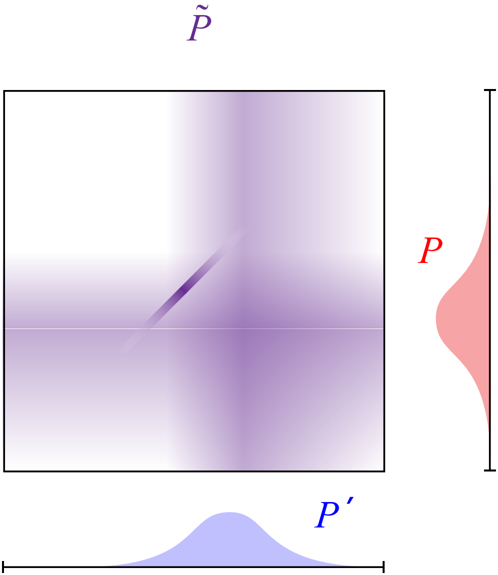


$$\tilde{Q} + \gamma\, M \otimes M'$$



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





$\tilde{P}$
P
P'

- Ω a Polish space,
- $\Delta := \{(x, x) \in \Omega \times \Omega : x \in \Omega\}$ the diagonal in $\Omega \times \Omega$,
- P, P' two Borel probability measures on Ω, and
- $\Pi(P, P')$ the space of all couplings of P and P'.

Then (folklore):

$$\sup_{A \in \mathcal{B}(\Omega)} \left| P(A) - P'(A) \right| = \min_{\tilde{P} \in \Pi(P, P')} (1 - \tilde{P}(\Delta)).$$

- $\Omega := S^{\mathbb{N}}$ the space of sequences for a finite set S,
- $E_0 := \bigcup_{n\in\mathbb{N}}\{(x, x') \in \Omega \times \Omega : (x_n, x_{n+1}, \ldots) = (x'_n, x'_{n+1}, \ldots)\}$ the equivalence relation of eventual equality,
- $\mathcal{T} := \bigcap_{n\in\mathbb{N}} \sigma(x_n, x_{n+1}, \ldots)$ the tail σ-algebra,
- P, P' two Borel probability measures on Ω, and
- $\Pi(P, P')$ the space of all couplings of P and P'.

Then (Griffeath 1974, Pitman 1976, Goldstein 1978):

$$\sup_{A\in\mathcal{T}} \left|P(A) - P'(A)\right| = 0 \text{ if and only if } \min_{\tilde{P}\in\Pi(P,P')} (1 - \tilde{P}(E_0)) = 0.$$

- $\Omega := S^{\mathbb{N}}$ the space of sequences for a finite set S,
- $\theta : \Omega \to \Omega$ the left-shift operation,
- $E_{\mathbb{Z}} := \bigcup_{n\in\mathbb{Z}}\{(x,x') \in \Omega \times \Omega : \theta^n(x) = x'\}$ the equivalence relation of shift-equivalence,
- $\mathcal{I}_{\mathbb{Z}} = \{A \in \mathcal{B}(\Omega) : \theta^{-1}(A) = A\}$ the shift-invariant σ-algebra,
- P, P' two Borel probability measures on Ω, and
- $\Pi(P, P')$ the space of all couplings of P and P'.

Then (Aldous-Thorisson 1993):

$$\sup_{A\in\mathcal{I}_{\mathbb{Z}}} \left|P(A) - P'(A)\right| = 0 \text{ if and only if } \min_{\tilde{P}\in\Pi(P,P')} (1 - \tilde{P}(E_{\mathbb{Z}})) = 0.$$

Also have generalizations to sufficiently regular group and semigroup actions (Thorisson 1996, Georgii 1997).

$$\sup_{A \in \mathcal{B}(\Omega)} \left| P(A) - P'(A) \right| = \min_{\tilde{P} \in \Pi(P, P')} (1 - \tilde{P}(\Delta)).$$

$$\sup_{A\in\mathcal{T}} \left|P(A) - P'(A)\right| = \min_{\tilde{P}\in\Pi(P,P')} (1 - \tilde{P}(E_0)).$$

$$\sup_{A \in \mathcal{I}_{\mathbb{Z}}} \left| P(A) - P'(A) \right| = \min_{\tilde{P} \in \Pi(P,P')} (1 - \tilde{P}(E_{\mathbb{Z}})).$$

$$\sup_{A \in \mathcal{G}} \left| P(A) - P'(A) \right| \stackrel{?}{=} \min_{\tilde{P} \in \Pi(P,P')} (1 - \tilde{P}(E)).$$

Many probability settings lead to the E*-coupling problem*

$$\inf_{\tilde{P}\in\Pi(P,P')} (1 - \tilde{P}(E)),$$

In general this problem is hard to solve and there are not many general-purpose tools available.

On the other hand, the $\mathcal{G}$*-total variation problem*

$$\sup_{A\in\mathcal{G}} |P(A) - P'(A)|$$

is typically easy to analyze for probabilists.

These optimization problems are closely related! In fact, we'll see that they are often dual, in the sense of mathematical optimization.

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

III. Problem statement

Notation:

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

Definition

Say E is *measurable* if $E \in \mathcal{F} \otimes \mathcal{F}$. Say $(E, \mathcal{G})$ is *strongly dual* if E is measurable and if we have

$$\sup_{A \in \mathcal{G}} \left| P(A) - P'(A) \right| = \min_{\tilde{P} \in \Pi(P,P')} (1 - \tilde{P}(E)).$$

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

Roughly speaking, $(\Delta, \mathcal{B}(\Omega)), (E_0, \mathcal{T})$, and $(E_{\mathbb{Z}}, \mathcal{I}_{\mathbb{Z}})$ are strongly dual.

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

If E is given, then there is a natural choice of $\mathcal{G}$:

Lemma (AQJ)

If $(E, \mathcal{G})$ is strongly dual for some $\mathcal{G}$, then (E, E^*) is strongly dual, where E^* is the E-invariant σ-algebra

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

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

Connection to optimal transport?

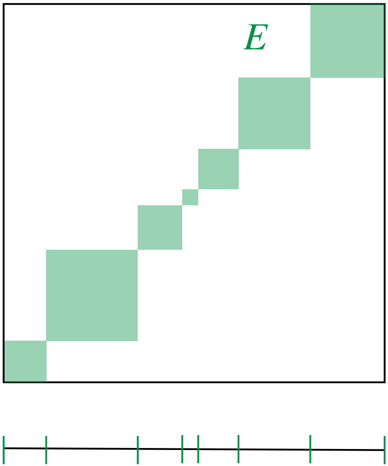
Note that the E-coupling problem

$$\inf_{\tilde{P} \in \Pi(P, P')} (1 - \tilde{P}(E)).$$

is exactly a Monge-Kantorovich problem with cost function

$$c(x, x') = 1 - \mathbb{1}\{(x, x') \in E\}.$$

In words: cost 0 to move within an equivalence class, and cost 1 to move between equivalence classes.


E

Classical Monge-Kantorovich theory (Rachev-Rüschendorf 1998, Villani 2009) requires topological regularity: Ω is a Polish space, $\mathcal{F}$ is its Borel σ-algebra, and c is lower semi-continuous.

In our setting, this requires E to be closed in $\Omega \times \Omega$; in this case, Kantorovich duality and some standard tricks can show that E is strongly dualizable.

However, most interesting equivalence relations, from the point of view of probability, are F_σ (countable union of closed) in $\Omega \times \Omega$.

IV. Results

Some useful reductions:

It is easy to show that we always have *weak dualizability*, that

$$\sup_{A \in E^*} \left| P(A) - P'(A) \right| \leq \inf_{\tilde{P} \in \Pi(P,P')} (1 - \tilde{P}(E)),$$

for all $P, P' \in \mathcal{P}(\Omega, \mathcal{F})$. The difficult part is showing the reverse inequality and that the inf is attained.

We say that E is *quasi-strongly dualizable* if for all $P, P' \in \mathcal{P}(\Omega, \mathcal{F})$ the following are equivalent:

- $P(A) = P'(A)$ for all $A \in E^*$
- There exists $\tilde{P} \in \Pi(P, P')$ and $N \in \mathcal{F} \otimes \mathcal{F}$ with $\tilde{P}(N) = 0$ and $(\Omega \times \Omega) \setminus E \subseteq N$.

Then E is strongly dualizable if and only if it is measurable and quasi-strongly dualizable.

Some basic descriptive set theory:

A measurable space $(S, \mathcal{S})$ is called a *standard Borel space* if there exists a Polish topology τ on Ω such that $\mathcal{S} = \mathcal{B}(\tau)$.

An equivalence relation E on a standard Borel space $(\Omega, \mathcal{F})$ is called *smooth* if there exists a standard Borel space $(S, \mathcal{S})$ and a measurable function $\phi : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ such that $(x, x') \in E$ is equivalent to $\phi(x) = \phi(x')$.

Roughly speaking, E is smooth if and only if the quotient Ω/E can be given a natural standard Borel structure.

Lemma (AQJ)

The following are equivalent:

(i) E *is smooth.*

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

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

The equivalence between (i) and (ii) is classical, but the equivalence with (iii) appears to be novel.

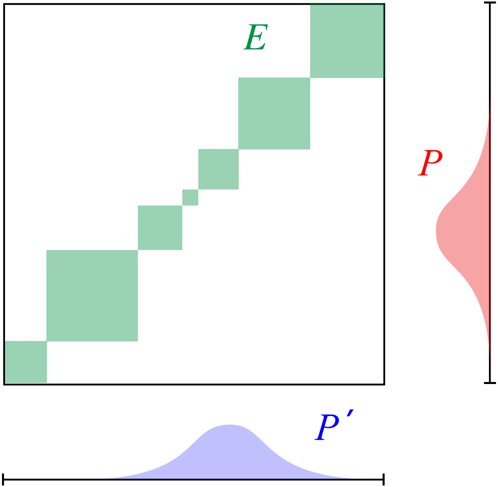
Theorem (AQJ)

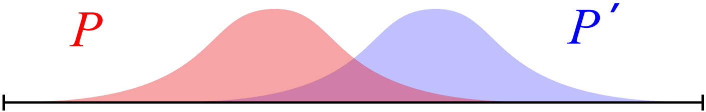
Every smooth equivalence relation is strongly dualizable.

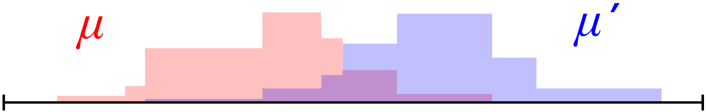
Consequence: All equivalence relations with G_δ (countable intersection of open) equivalence classes are strongly dualizable.

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

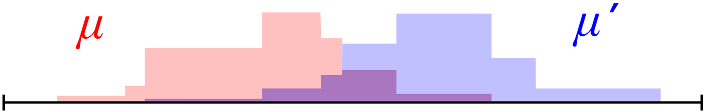
$\tilde{P}$?




P
P′

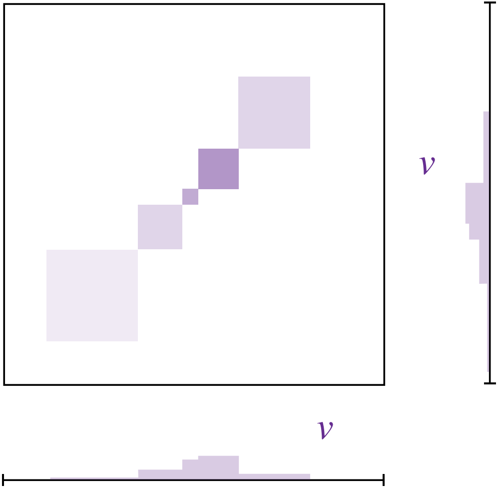

μ
μ'

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

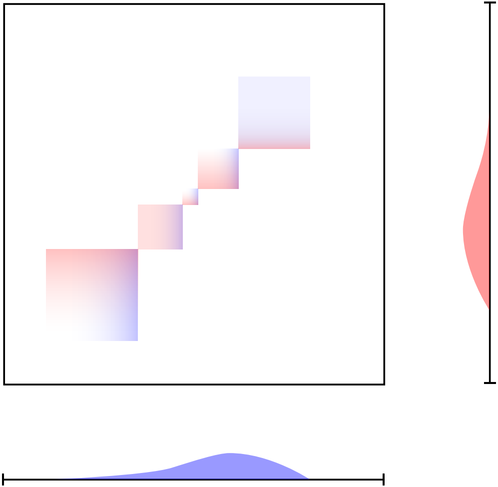


v

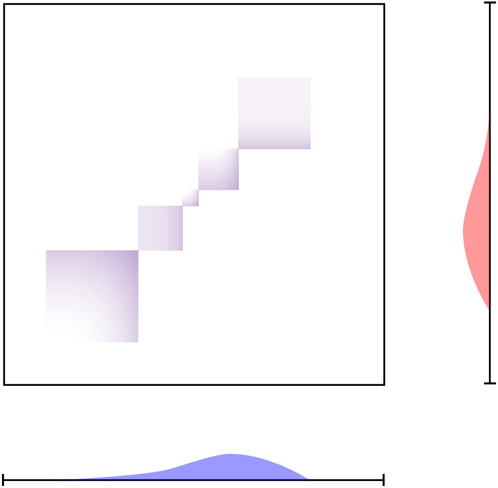
$$v \circ (i,i)^{-1}$$



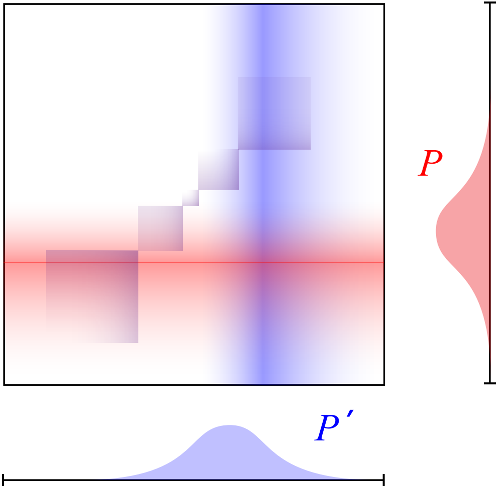
$$\tilde{Q}(\cdot \times \cdot \mid E^*) \approx \mu(\cdot \mid E^*)\mu'(\cdot \mid E^*)$$



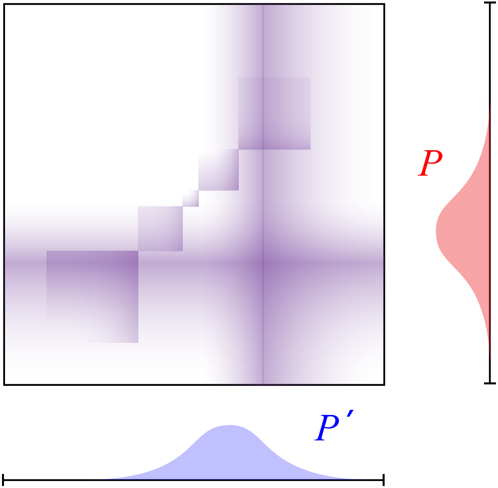
$\tilde{Q}$

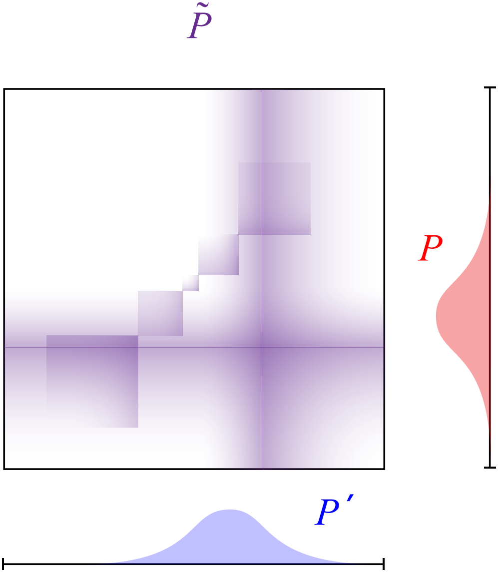


$$\tilde{Q} + \gamma\, M \otimes M'$$



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




$\tilde{P}$
P
P'

However, many equivalence relations of interest are not smooth.

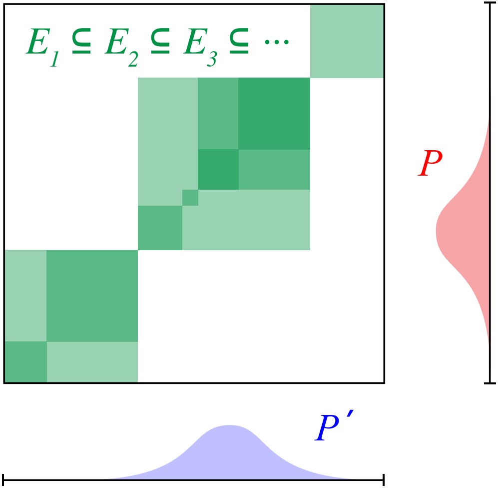
Instead, we have the following closure result:

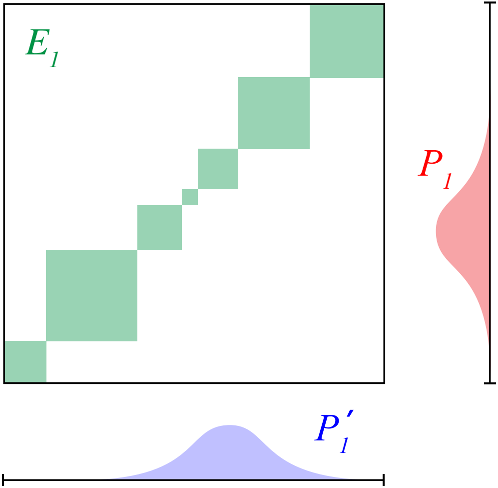
Theorem (AQJ)

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

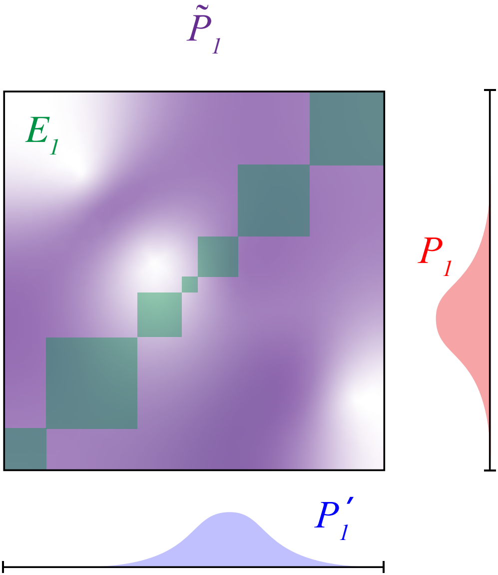
Idea of proof: Apply strong duality, and iterate.

$\tilde{P}$?



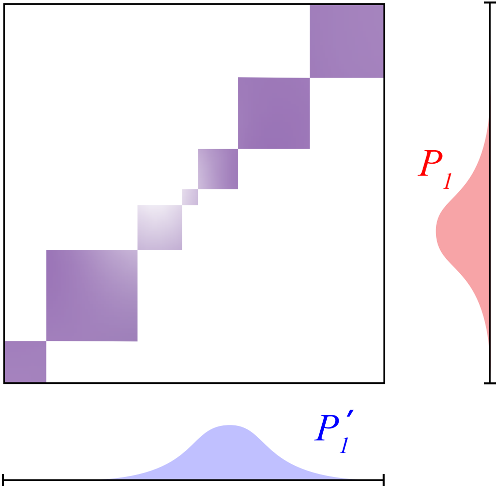


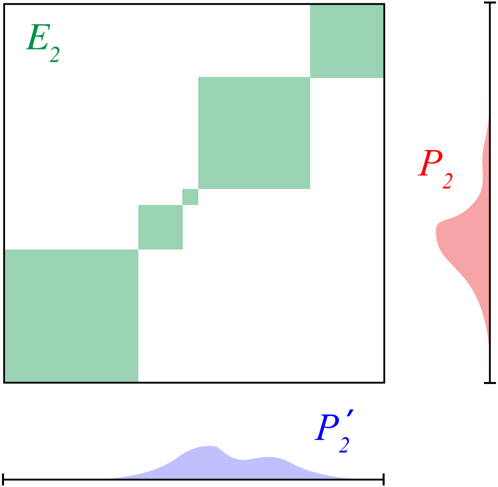
E_1
P_1
P_1'

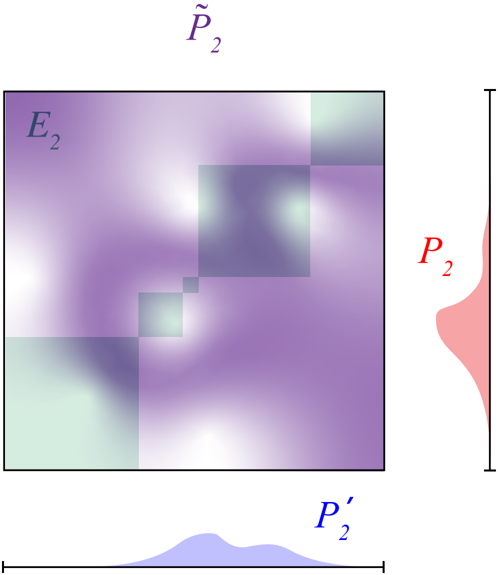


$\tilde{P}_1$
E_1
P_1
P'_1

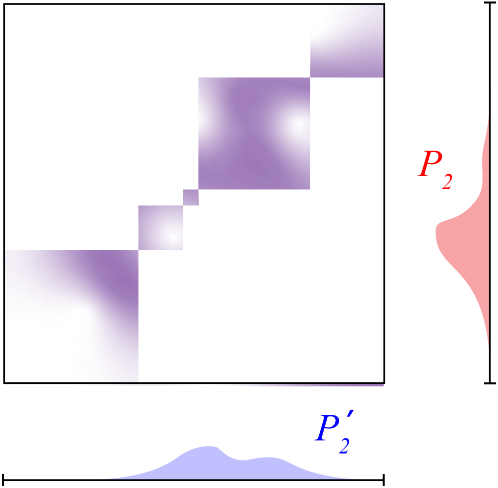
$$\tilde{P}_1(\,\cdot \cap E_1)$$

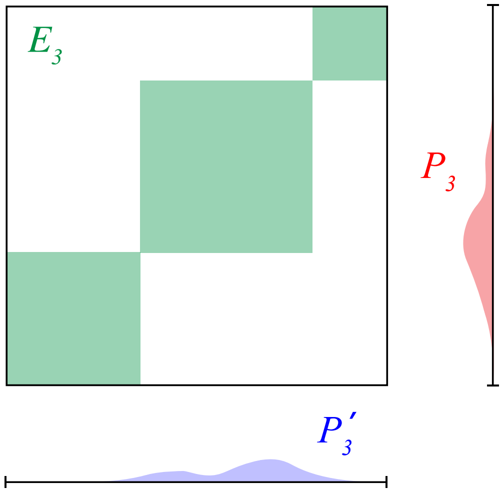


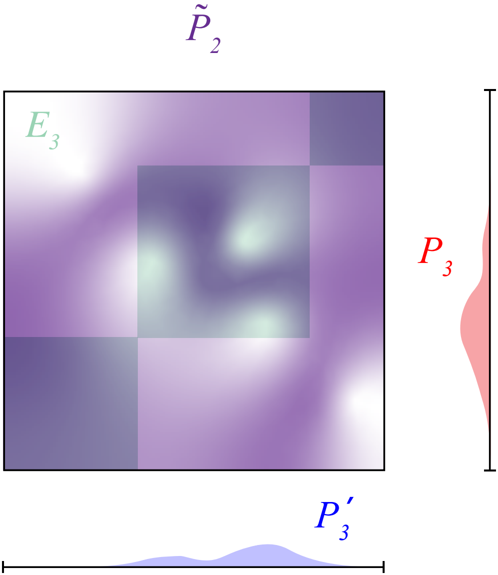

E_2
P_2
P_2'


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

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

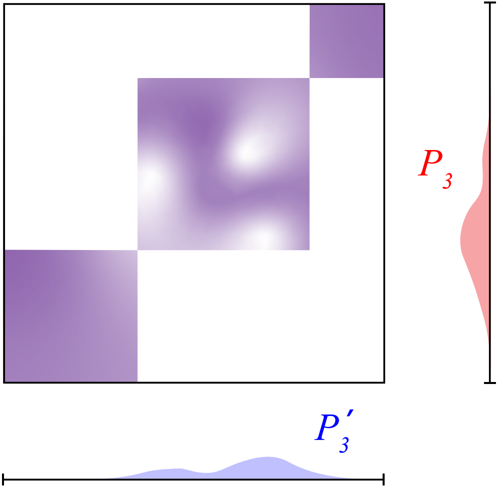



E_3
P_3
P_3'

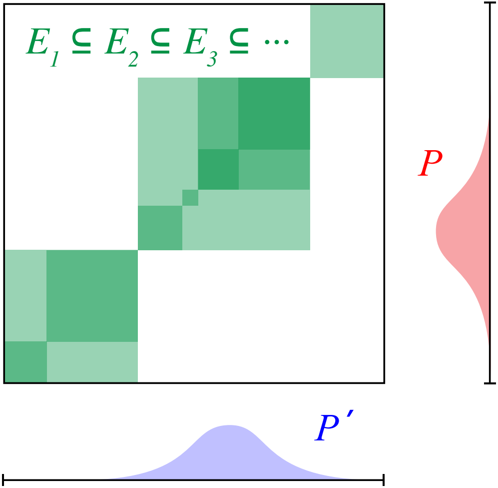


$\tilde{P}_2$
E_3
P_3
P_3'

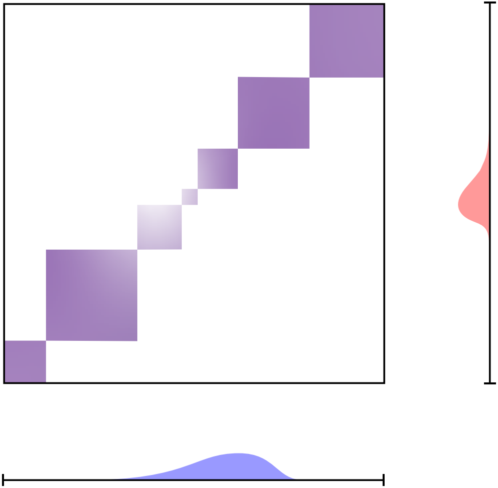
$\tilde{P}_3(\,\cdot \cap E_3)$



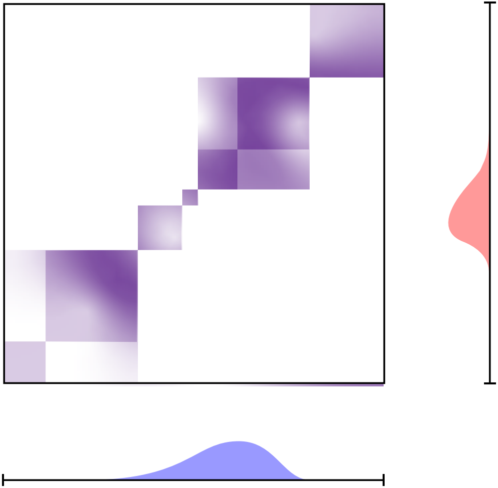
$\tilde{P}$?



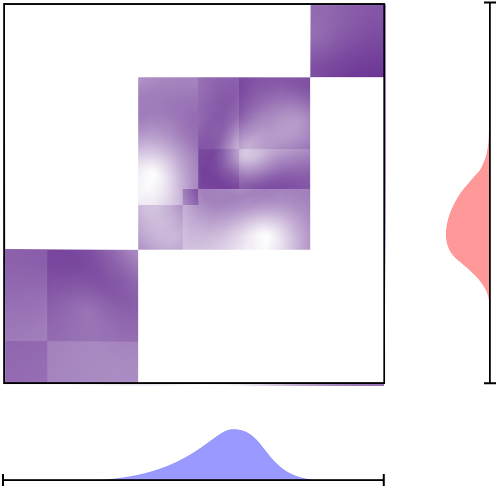
$\tilde{P}_1(\,\cdot \cap E_1)$



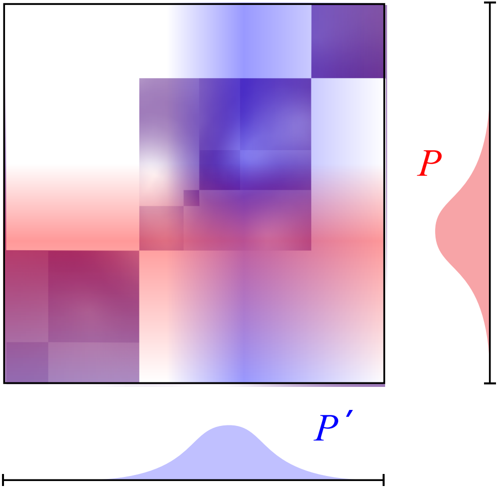
$$\sum_{n=1}^{2} \tilde{P}_n(\,\cdot \cap E_n)$$



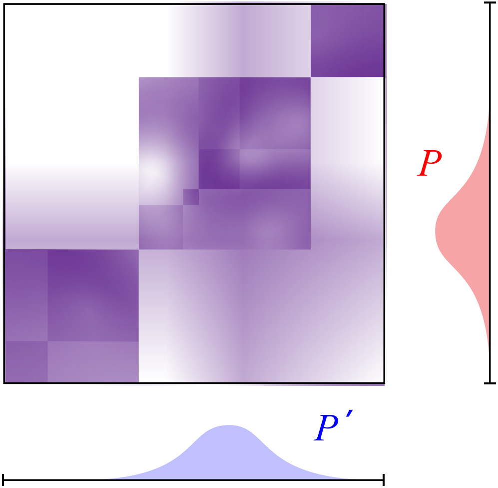
$$\sum_{n=1}^{3} \tilde{P}_n(\,\cdot \cap E_n)$$

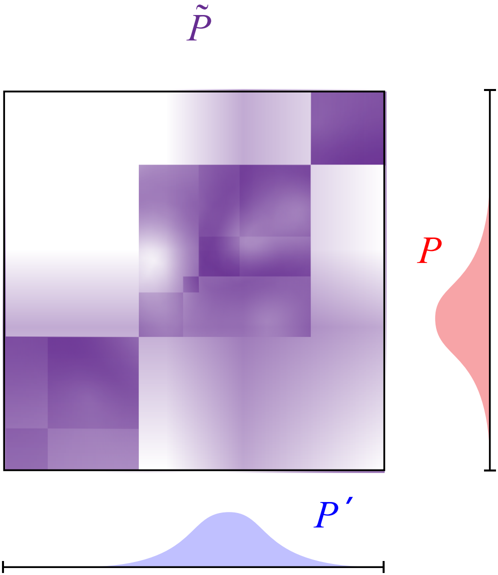


$$\sum_{n=1}^{\infty} \tilde{P}_n(\cdot \cap E_n) + \gamma M \otimes M'$$



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




$\tilde{P}$
P
P'

This shows strong dualizability for all *hypersmooth* (countable union of smooth) equivalence relations, and this covers most "reasonable" equivalence relations occurring in probability.

Thank you!

References

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

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.

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

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.