Numeric examples. (To be moved to the Introduction Section.) In the Numerics section, we develop several examples of robust pricing problems designed to show the improvement allowed by our result. We consider cost functions based on the the european call option applied to a basket of assets and a theoretical covariance (or cross-product) contract. The first two examples use hypothetical uniform and normal marginal distributions, while the third uses real-world marginal distributions implied by option prices of individual assets at two distinct future points in time.

In this section, we solve several versions of the problem of robust pricing of a financial contract involving the prices of a group of assets at two distinct points in time. We build upon the framework of EK21[1] and analyze potential performance improvements allowed by our result on the geometry of the optimal couplings. Our tests run across various dimensions, cost functions and marginal shapes, including marginals derived from actual stock market option prices.

[NOTE: discussion about θ moved to here and developed further]

EK21 uses neural network optimization with penalization to solve the dual problem of Optimal Transport. More specific to our case, their model approximate the potential functions of our equation 2.6 with neural networks and impose condition 2.7 though penalization. The optimization objective is summarized in a sum of integrals (equation 2.3 of EK21), which we repeat here with adapted notation. We want to (min)maximize

$$\begin{aligned}\text{Loss} &= \int \varphi d\mu_0 + \int b_\gamma \left(c - \varphi\right) d\theta \\ &= \int \left(\sum_i \phi_i + \sum_i \psi_i\right) d\mu_0 + \int b_\gamma \left(c - \varphi\right) d\theta \\ &= \sum_i \int \phi_i d\mu_i + \sum_i \int \psi_i d\nu_i + \int b_\gamma \left(c - \varphi\right) d\theta\end{aligned}$$

where $\varphi\left(x, y\right) = \sum_i \phi_i\left(x_i\right) + \sum_i \psi_i\left(y_i\right) + \sum_i h_i\left(x\right) . \left(y_i - x_i\right)$. Notice that the fact that $\int \sum_i h_i\left(x\right) . \left(y_i - x_i\right) d\mu_0 = 0$, from the martingale condition, allows us to eliminate this term from the first integral. Given the separation in the first term of the RHS, μ_0 can be simply taken as the independent coupling of the marginals. Computationally, each of ϕ_i, ψ_i, h_i is replaced by some approximation $\phi_i^m, \psi_i^m, h_i^m$ implemented as a neural network with an internal size measured by m. The calibration of the neural networks is performed through gradient descent methods referenced on a loss function. The last integral adds a penalization to the loss function, imposing the appropriate pointwise inequality between φ and c.

Our main point of interest is in the sampling measure θ used for penalization. While the only requirement on θ besides is that a true solution π^* be absolutely convex with respect to it, a better accuracy is expected in general when θ is closer to a true solution – see brief discussion in sections 4.4 and 4.5 of EK21. In its most general form, θ is composed of the independent coupling of the marginals. The novelty of our method is in the improvement of θ: since we know that a solution (unique in the case of strict submodularity) is attained with monotone coupling of the X marginals $(\vec{\mu}_i)$, we determine θ^1, the X-side of θ, by those marginals. In the special case of normal marginals, we improve θ a little further by applying additional knowledge about the full coupling of the marginals, including the Y side. It can be possible to improve θ even further and in more general cases. For instance, we could achieve great improvements with the knowledge of the conditional distribution of Y given X, that is, π_x for a fixed x. This topic remains for future research. The purpose here is to run tests with and without the known improvements that we have at hand. We fix the penalization function b (see discussion in the reference) as

$$b(t) = \frac{1}{\gamma}\frac{1}{2}\left((\gamma t)^+\right)^2 \qquad\qquad \gamma = 1000$$

Example 1. (Basket option cost function, uniform marginals.) Our first cost function mimics the payoff of a contract similar to an European call option applied to a basket composed of assets. That is, a cost function with general form

$$c(X,Y) = \left(\sum_{i=1}^{d} a_i X_i - k_1\right)^+ + \left(\sum_{i=1}^{d} b_i Y_i - k_2\right)^+$$

Our marginals are uniform distributions centered at zero. Set $d = 2$ and

$$\begin{aligned} c(X_1,Y) &= (X_1 + 2X_2 - 1)^+ + (Y_1 + 2Y_2 - 1)^+ \\ X_1, X_2 &\sim U[-1,1] \\ Y_1, Y_2 &\sim U[-2,2] \end{aligned}$$

We are interested in both upper and lower bounds. It is useful to calculate p^+ and p^- as references for the numerical output [check conditions for monotone coupling on Y] . For the maximum (minimum) we take $\hat{\pi}$ with positive (negative) monotone coupling on both $\hat{\pi}^1$ and $\hat{\pi}^2$. We calculate the bounds as

$$\begin{aligned} p^+_{\hat{\pi}} &= p^+_{X,\hat{\pi}} + p^+_{Y,\hat{\pi}} \\ p^-_{\hat{\pi}} &= p^-_{X,\hat{\pi}} + p^-_{Y,\hat{\pi}} \end{aligned}$$

where, using the formula on A1 below,

$$\begin{aligned} p^+_{X,\hat{\pi}} &= \frac{1}{3}; \qquad\qquad & p^+_{Y,\hat{\pi}} &= \frac{25}{24} \\ p^-_{Y,\hat{\pi}} &= 0; & p^-_{Y,\hat{\pi}} &= \frac{1}{8} \end{aligned}$$

Thus, our theoretical reference bounds for the numerical approximation are $p^+ = \frac{1}{3} + \frac{25}{24} = \frac{11}{8} = 1.375$ and $p^- = \frac{1}{8} = 0.125$. Our first step in the computation process is to build two samples from the theoretical distributions μ_0 and θ in equation 2.3 of –EcksteinKupper21–. [NOTE: refer to the discussion about θ in the section introduction] A more elaborate version comes from using our main result, when we set θ^1 to be determined by the monotonic coupling of $(\vec{\mu}_i)$ (positive or negative according to the objective). We use both choices of θ and compare the outputs. Sample sizes are 100k in both cases. The graph below shows the convergence of the numeric value for the number of iterations in the horizontal axis. The shaded grey area covers ± 1 standard deviations of the numeric outputs.

[graph: convergence max]

[graph: convergence min]

The following table summarizes the results after 200 iterations.

Maximization.

Coupling	Independent	Positive θ^1
Target	1.375	
Dual approx.value	1.2807	1.3170
Standard deviation	0.0254	0.0288
Penalty	0.0174	0.0055

Minimization.

Coupling	Independent	Negative θ^1
Target	0.125	
Dual approx.value	0.1433	0.1311
Standard deviation	0.0230	0.0251
Penalty	0.0082	0.0075

Example 2. (Covariance cost function, normal marginals.) For our second example, we consider a cost function inspired by a theoretical contract with payoff proportional to the covariance of two assets. The general form is

$$c(X,Y) = \sum_i \sum_{j>i} a_{ij} X_i X_j + b_{ij} Y_i Y_j$$

Now our marginals follow normal distributions with parameters

$$X_i \sim N\left(0, \sigma_i^2\right)$$
$$Y_i \sim N\left(0, \rho_i^2\right)$$

We start with $d = 2$ and set

$$c(X,Y) = Y_1 Y_2$$
$$(\sigma_1, \sigma_2) = (2, 1)$$
$$(\rho_1, \rho_2) = (3, 4)$$

By symmetry, we only need to look at the maximization problem. Proposition 3.7 gives us the exact solution to be used as a reference for the maximum price

$$p^+ = 2 * 1 + \sqrt{3^2 - 2^2}\sqrt{4^2 - 1^2} \sim 10.6603$$

To avoid noisy behavior in the potential functions, we sample from normal distributions clipped at $\pm 4\sigma$. As before, we run our optimization with the most general sampling measure θ given by the independent coupling of the marginals and with an improved measure that uses our result, namely, with θ^1 determined by the monotone coupling of $(\vec{\mu}_i)$. A futher improvement is allowed by the fact that we are using normal marginals, see (REF), which gives us a third, potentially even more efficient θ where θ^1 is given as before and θ^2 follows a joint normal with parameters given by the reference (develop). The value convergence using the three sampling measures is shown below. Sample sizes are 100k.

[graph: convergence]

The convergence graph shows similar pattern between independent and monotone coupling on θ^1. (comment)

To provide a visual illustration of the coupling process, we construct examples of distributions from the potential functions generated in the numerical procedure, according to equation 2.6 of –EcksteinKupper2021– [1].

[graphs: $\hat{\pi}$ for both cases]

Results after 2000 iterations.

Cupling	Independent	Positive θ^1	Positive θ^1 and optimal θ^2
Target	10.6603		
Dual approx.value	10.3836	10.5351	10.7753
Standard deviation	0.4413	0.4465	0.3814
Penalty	0.4135	0.1684	0.0132

Higher dimensions. To compare performance in higher dimensions, we also run examples with $d = 3$ and $d = 5$ (to do: change second to $d = 10$). Our cost is simplified as

$$c(x, y) = \sum_i \sum_{j>i} b_{ij} y_i y_j$$

For $d = 3$, the cost, marginals and target maximum values are given by

$$\begin{aligned} c(x, y) &= y_1 y_2 + 2 y_1 y_3 + \frac{1}{2} y_2 y_3 \\ (\mu_1, \mu_2, \mu_3) &= (1, 2, 3) \\ (\rho_1, \rho_2, \rho_3) &= (2, 3, 8) \\ p^+ &\approx 48.855 \end{aligned}$$

Below is a graph of the convergence

[graph: convergence]

[comment]

Results after 2000 iterations. (update)

Cupling	Independent	Positive θ^1
Target	48.855	
Dual approx.value		
Standard deviation		
Penalty		

For the $d = 5$, our setting is

$$c(x, y) = \sum_i \sum_{j>i} y_i y_j$$
$$(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5) = (1, 2, 2, 3, 3)$$
$$(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5) = (2, 3, 4, 5, 6)$$
$$p^+ \approx 153.751$$

Below is a graph of the convergence
[graph: convergence]
[comment]
Results after 2000 iterations. (update)

Cupling	Independent	Positive θ^1
Target	153.751	
Dual approx.value		
Standard deviation		
Penalty		

Example 3. (Covariance cost function with real stock market marginals.) To be developed. Include graphs of dual functions and discussion.

A1. Portfolio option price – direct calculation. Let

$$f(x_1, x_2) = (a_1 x_1 + a_2 x_2 - k)^+$$
$$X_i \sim \mu_i \equiv U[-m, m], i = 1, 2$$

with $a_i \geq 0; m > 0$. We are interested in the maximum and minimum of the expected value of f over all possible couplings of μ_1 and μ_2. Denote

$$p^+ = \max_{\pi \in \Pi(\mu_1, \mu_2)} \mathbb{E}_\pi f(x)$$
$$p^- = \min_{\pi \in \Pi(\mu_1, \mu_2)} \mathbb{E}_\pi f(x)$$

By [REF], the maximum is attained at the positive-diagonal monotone coupling of μ_1, μ_2, named π^+. Since $x_1 = x_2$ in the support of π^+, we have

$$\begin{aligned}
p^+ &= \int f(x)\,\mathbb{P}(x)\,d\pi^+ \\
&= \frac{1}{2m}\int_{-m}^{m} f((t,t))\,dt \\
&= \frac{1}{2m}\int_{-m}^{\frac{k}{a_1+a_2}} f((t,t))\,dt + \frac{1}{2m}\int_{\frac{k}{a_1+a_2}}^{m} f((t,t))\,dt
\end{aligned}$$

Notice that the first integral is zero, and so is the second one if $m \leq \frac{k}{a_1+a_2}$. If $m \geq \frac{k}{a_1+a_2}$ then we have

$$\begin{aligned}
p^+ &= \frac{1}{2m}\int_{\frac{k}{a_1+a_2}}^{m} f((t,t))\,dt \\
&= \frac{1}{2m}\int_{\frac{k}{a_1+a_2}}^{m} [(a_1+a_2)\,t - k]\,dt \\
&= \frac{a_1+a_2}{4m}\left[m^2 - \left(\frac{k}{a_1+a_2}\right)^2\right] - k\left(m - \frac{k}{a_1+a_2}\right) \\
&= \frac{a_1+a_2}{4m}\left[m^2 - 2m\frac{k}{a_1+a_2} + \left(\frac{k}{a_1+a_2}\right)^2\right] \\
&= \frac{a_1+a_2}{4m}\left(m - \frac{k}{a_1+a_2}\right)^2
\end{aligned}$$

Similarly, for the minimum we consider the negative-diagonal monotone coupling π^- in whose support $x_2 = -x_1$. Thus we have

$$\begin{aligned}
p^- &= \int f(x)\,\mathbb{P}(x)\,d\pi^- \\
&= \frac{1}{2m}\int_{-m}^{m} f((t,-t))\,dt \\
&= \frac{1}{2m}\int_{-m}^{\frac{k}{a_1-a_2}} f((t,-t))\,dt + \frac{1}{2m}\int_{\frac{k}{a_1-a_2}}^{m} f((t,-t))\,dt
\end{aligned}$$

Now the second integral is zero and so is the first if $-m \geq \frac{k}{a_1-a_2}$, or $m \leq$

$\frac{k}{a_2-a_1}$. If $m \geq \frac{k}{a_2-a_1}$ then

$$
\begin{aligned}
p^- &= \frac{1}{2m} \int_{-m}^{\frac{k}{a_1-a_2}} f\left((t,-t)\right) dt \\
&= \frac{1}{2m} \int_{-m}^{\frac{k}{a_1-a_2}} \left[(a_1-a_2)\, t - k\right] dt \\
&= \frac{a_1-a_2}{4m} \left[\left(\frac{k}{a_1-a_2} \right)^2 - m^2 \right] - k \left(\frac{k}{a_1-a_2} + m \right) \\
&= \frac{a_1-a_2}{4m} \left[- \left(\frac{k}{a_1-a_2} \right)^2 - m^2 - 2km \right] \\
&= \frac{a_2-a_1}{4m} \left(m - \frac{k}{a_2-a_1} \right)^2
\end{aligned}
$$

References

[1] S. Eckstein and M. Kupper. Computation of optimal transport and related hedging problems via penalization and neural networks. *Applied Mathematics and Optimization*, April 2021.