Spread Options

and Farkas' Lemma

Vladimir Piterbarg

Barclays Capital

1 Introduction

- We are interested in building two-dimensional distributions consistent with markets in
  - All options on X
  - All options on Y
  - All options on $S = X - Y$
- Example (candidate) solutions exist: Austing [Aus10], Andersen&Piterbarg mixing solution [AP10]
- Options marks can be translated into densities – more convenient mathematically
- Take a broader view: Fully characterize 2D distributions that have given 1D marginals and the "diagonal", ie the density of $X - Y$
  - Necessary and sufficient conditions for existence?
  - Neat copula-like formula to describe them all?

- Why? Isolated examples of such distributions are not enough to assess model risk for options with "other" 2D payoffs, ie general baskets $(w_1X + w_2Y - K)^+$
  - How much uncertainty still remains in basket option prices once the marginals and the spread options are locked down?
- Appears to be a very difficult problem. Neat numerical solutions exist but complete theory is still elusive
- The talk is really a progress report (and a challenge to better mathematicians)

2 Matrix Formulation

- Start by scaling and discretizing the distributions
- Can assume that X, Y are supported on the same grid $\{i/N, i = 0, \dots, N\}$
- Marginal distributions: $r_i = \mathrm{P}(Y = (N-i)/N)$, $c_j = \mathrm{P}(X = j/N)$, $r_i, c_i \geq 0, \sum r_i = \sum c_i = 1$
- Let $S = X - Y$ be the spread variable, with the desired distribution $d_k = \mathrm{P}(S = (k-N)/N)$ for $k = 0, \dots, 2N$, $d_k \geq 0$ and $\sum d_k = 1$
- Existence problem: given r, c and d find a matrix $p = p_{i,j}$ such that

$$p_{i,j} \geq 0, \quad i, j = 0, \dots, N, \tag{1}$$

$$\sum_{j=0}^{N} p_{i,j} = r_i, \quad i = 0, \dots, N, \tag{2}$$

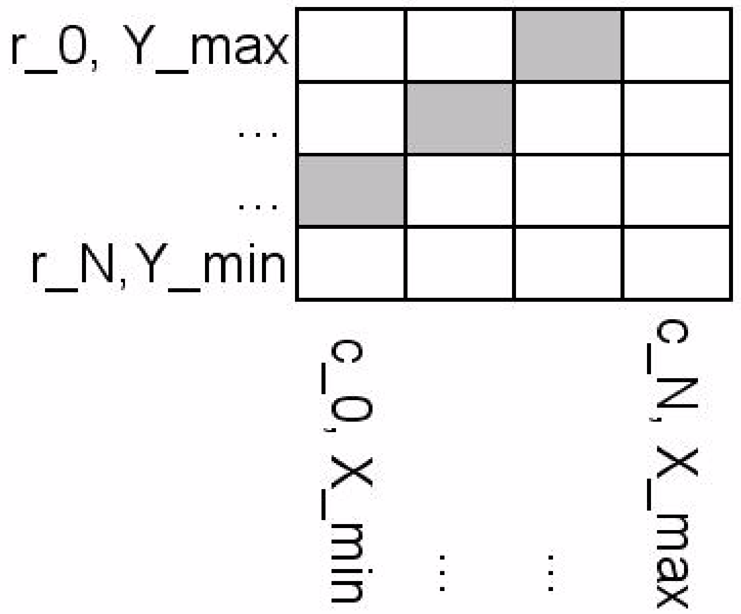
$$\sum_{i=0}^{N} p_{i,j} = c_j, \quad j = 0, \dots, N, \tag{3}$$

$$\sum_{(i,j)\in D_k} p_{i,j} = d_k, \quad k = 0, \dots, 2N, \tag{4}$$

where D_k is the k-th diagonal,

$$D_k = \{(i,j) : i + j = k, 0 \leq i, j \leq N\}, \; k = 0, \dots, 2N.$$

• Visually:



3 Existence Should Be Easy?

- Linear equations with positivity constraints. Have $(N+1)^2$ variables, and at most $(N+1)+(N+1)+(2N+1)-2 = 4N+1$ independent equations.
- For $N > 2$ under-specified – "most of the time" the solution should exists?
- An example of a potential problem. If such p exists we must have

$$\begin{aligned}\sum_{k=0}^{2N} kd_k &= \sum_{k=0}^{2N} k \sum_{(i,j)\in D_k} p_{i,j} = \sum_{k=0}^{2N} \sum_{(i,j)\in D_k} (i+j)p_{i,j} \\ &= \sum_{i,j=0}^{N} (i+j)p_{i,j} = \sum_i i \sum_j p_{i,j} + \sum_j j \sum_i p_{i,j} \\ &= \sum_i ir_i + \sum_j jc_j\end{aligned}$$

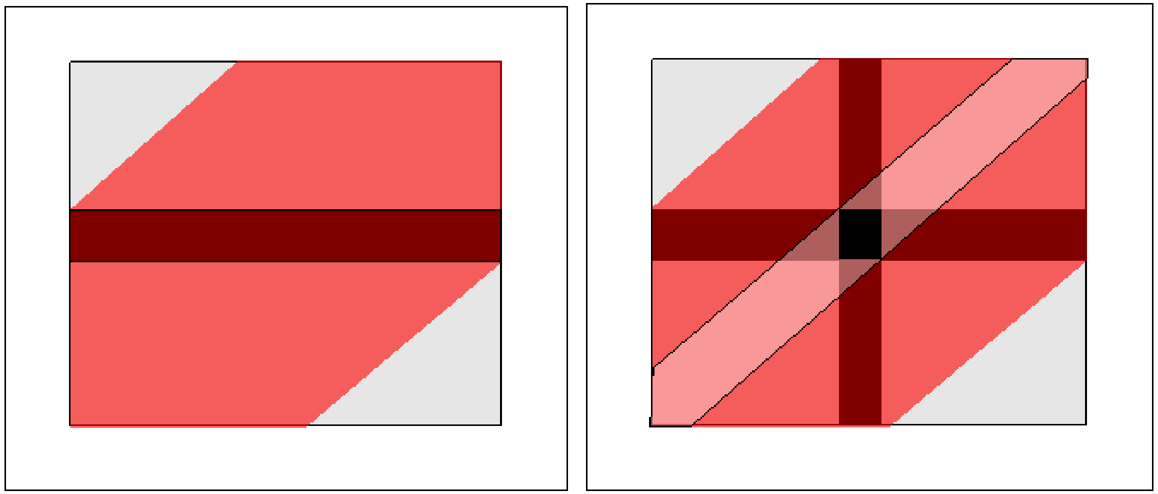
so the necessary condition for existence is

$$\sum_k kd_k = \sum_i ir_i + \sum_j jc_j.$$

- This is of course nothing but the condition that $\mathrm{E}(X-Y) = \mathrm{E}(X) - \mathrm{E}(Y)$, but e.g. have to be careful when determining the density of $S = X - Y$ and discretizing

4 Geometric necessary conditions

- There are more subtle necessary conditions



1. A row(s) can be covered by a set of diagonals, so need $\sum r_j \leq \sum' d_k$
2. More generally, a "cross" can be covered by two sets of diagonals, one for the cross and the other for the intersect (black square), so $\sum c_i + \sum r_j \leq \sum' d_k + \sum'' d_k$

- Are these "geometric" conditions sufficient? I.e. if

$$\sum_{i \in I} c_i + \sum_{j \in J} r_j \leq \sum_{k: D_k \cap (I \cup J) \neq \emptyset} d_k + \sum_{k: D_k \cap (I \cap J) \neq \emptyset} d_k$$

for all sets of columns I and rows J, then solution to (1)–(4) exists?

- NO – have a (numerical) counterexample

5 Triangle Arbitrage

- McCloud [McC11] finds "triangle arbitrage" between swaptions and CMS spread options:

$$(X - K_x)^+ - (Y - K_y)^+ \leq (S - (K_x - K_y))^+ \leq (X - K_x)^+ + (K_y - Y)^+$$

- These can be expressed in terms of c, r, d, giving another set of necessary conditions for the existence of joint distribution
- Can be expressed by the condition that spread option prices (as given by d) should lie between those given by the Frechet bounds copulas

$$V_D(K) \leq V(K) \leq V_{AD}(K), \tag{5}$$

  where the *perfect dependence* and *anti-dependence* copulas are given by

$$C_D(u, v) = \min(u, v),\ C_{AD}(u, v) = (u + v - 1)^+$$

- Sufficient? If r, c, d are such that the options satisfy Frechet bounds (5) then the joint distribution exists?
- NO – have a counterexample
- In other words there could be other arbitrage opportunities beyond those identified in [McC11]. We'll see an example

6 Spread Envelope

- Let $f(x)$ and $g(y)$ be two functions. Define their "spread envelope" by

$$\mathcal{E}_{f,g}(z) = \max_{x-y=z} \{f(x) + g(y)\}$$

- The smallest function of $x - y$ only that dominates $f(x) + g(y)$ for all x, y
- If (X, Y) have a 2D distribution, then we clearly have

$$\mathrm{E}\left(f(X)\right) + \mathrm{E}\left(g(Y)\right) \leq \mathrm{E}\left(\mathcal{E}_{f,g}(X - Y)\right)$$

  (if the payoff is dominated, the values are as well – a no-arbitrage condition)
- Note the rhs depends on the distribution of S only
- So we have necessary conditions: For any vectors f, g we must have

$$\sum_j f_j c_j + \sum_i g_i r_i \leq \sum_k \max_{(i',j')\in D_k} \{f_{i'} + g_{j'}\}\, d_k \tag{6}$$

- Both geometric and triangle arbitrage conditions are special cases of these.
  - For geometric we use indicator functions $f_i = 1_{\{i\in I\}}$ etc
  - For triangle arbitrage we use calls/puts

7 Spread Envelope and Existence Result

- If

$$\sum_j f_j c_j + \sum_i g_i r_i \leq \sum_k \max_{(i',j') \in D_k} \{f_{i'} + g_{j'}\} d_k$$

  holds for all f and g, does the joint distribution exists?

- YES! This is the so-called Farkas' lemma from the theory of linear inequalities, expressed in financial terms

- The proof is based on the separation theorem (two convex sets can be separated by a hyperplane) and can be looked up online

- Clear financial interpretation – if we cannot construct an arbitrage between marginal markets and the spread market (so called *spread envelope arbitrage*), then a joint distribution exists

- Elegant but not very useful (too many payoffs to check)

- Open question – can we find a subset of payoffs that would guarantee existence?

8 Existence by Linear Programming

- Let move from theory to practice
  - Find a solution/a family of solutions if they exist
  - If no solution find the arbitrage implied by Farkas' lemma
- Checking if $p_{i,j}$ exists such that

$$p_{i,j} \geq 0, \quad \sum_{j=0}^{N} p_{i,j} = r_i, \quad \sum_{i=0}^{N} p_{i,j} = c_j, \quad \sum_{(i,j) \in D_k} p_{i,j} = d_k, \tag{7}$$

  could be seen to be the problem of funding a *feasible solution* in *Linear Programming*.

- For $N = 100$ have about 10,000 variables – not a difficult task for the simplex method
- Reformulate using so-called *slack* variables q_i, $i = 0, \ldots, 4N + 2$ (one for each row, column and diagonal)

- Find $p_{i,j}$, q_i such that

$$\sum_{i=0}^{4N+2} q_i \to \min, \tag{8}$$

$$p_{i,j} \geq 0, \quad q_i \geq 0,$$

$$\sum_{j=0}^{N} p_{i,j} \underline{+q_i} = r_i, \quad \sum_{i=0}^{N} p_{i,j} \underline{+q_{N+1+j}} = c_j, \quad \sum_{(i,j)\in D_k} p_{i,j} \underline{+q_{2N+2+k}} = d_k.$$

- Simplex method will go vertex to vertex improving the objective function
- Can start from a feasible (ie non-optimal but satisfying the constraints) solution $p_{i,j} = 0$,

$$q_i = \begin{cases} r_i, & i \leq N, \\ c_{i-N-1}, & N+1 \leq i \leq 2N+1, \\ d_{i-2N-2}, & 2N+2 \leq i. \end{cases}$$

- If the solution to optimization problem (8) satisfies $q_i = 0$ for all i, then the original problem (7) has solution. If not, the solution to (7) does not exist.

9 Finding Multiple Solutions by LP

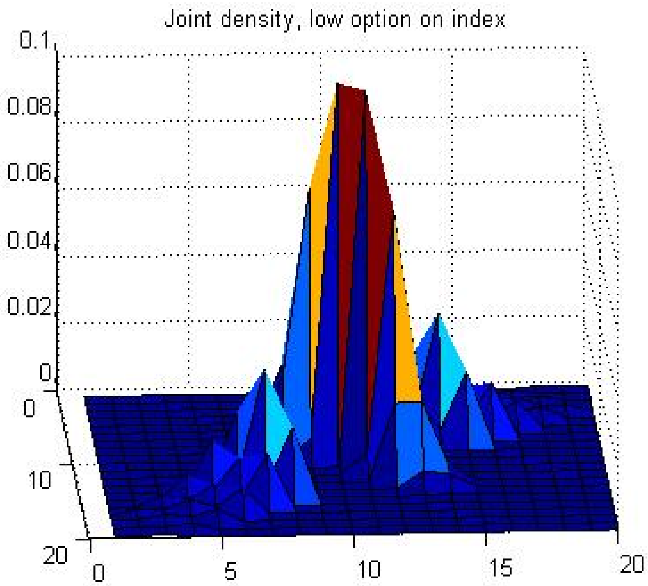
- If a solution exists there will likely be more than one
- Recall motivation: quantifying uncertainty on non-spread 2D payoffs
- Using LP we can look for the extremes
- Value of a call on an *index* $X + Y$ (strike K_{idx}) is linear in p:

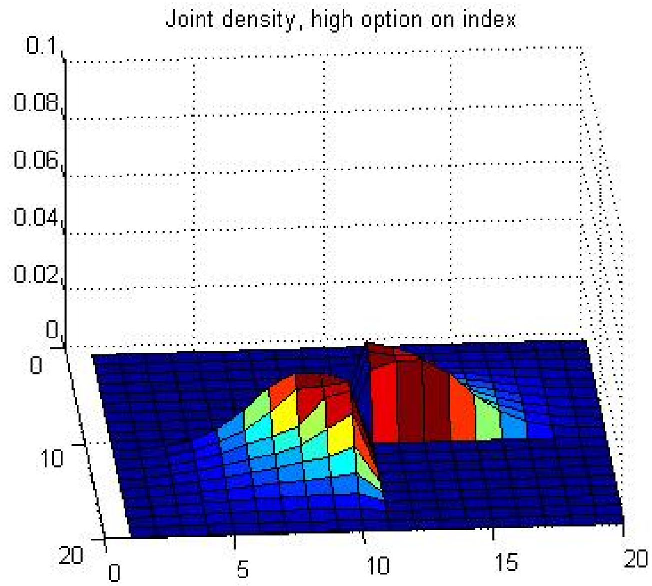
$$T(p) = \sum \left(\frac{j}{N} + \frac{N-i}{N} - K_{idx} \right)^{+} p_{i,j}.$$

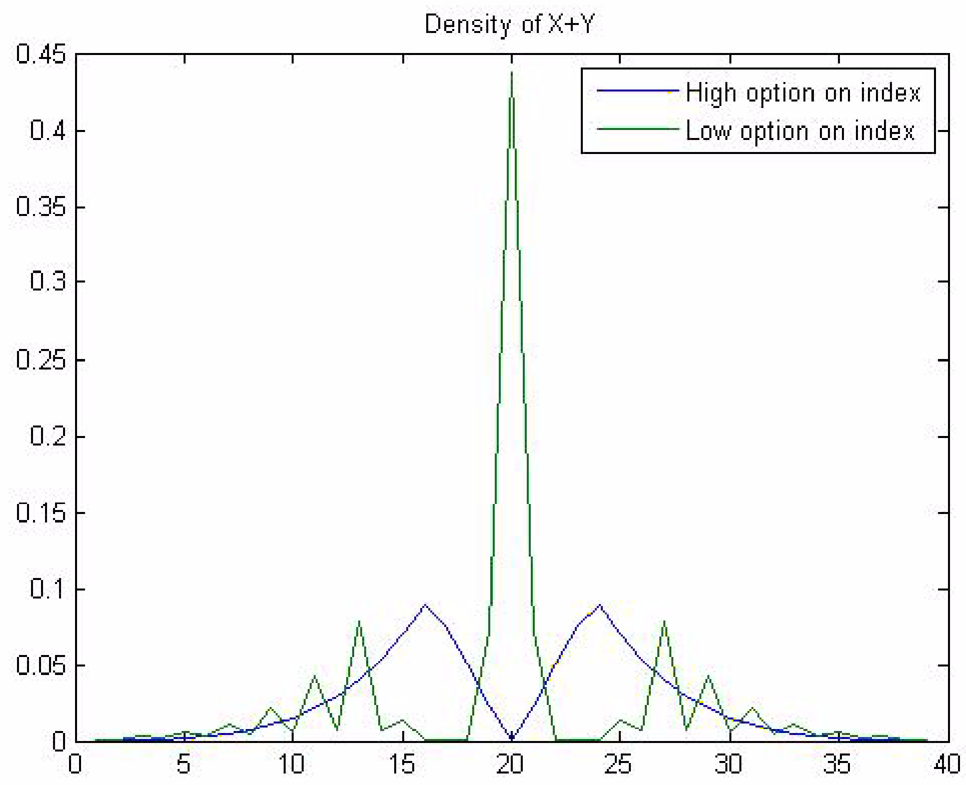
- "High option on index" solution solves $T(p) \to \max$ given constraints (7)
- "Low option on index" solution solves $T(p) \to \min$ under the same constraints
- Use simplex method starting from the feasible solution previously found
- In some sense "all the other" solutions can be found as the weighted average of the two
- Different basket options will give different results but in our experience only slightly
- The problem is solved for practical purposes

10 Results, Gaussian Case

- $\sigma_X = \sigma_Y = 1\%$, $\rho_{XY} = 70\%$. We generate marginal and spread option densities using these parameters and then apply our machinery
- Maximizing the value of the ATM index (i.e. on $X + Y$) option
- Two extremes below
- Implied Gaussian copula correlation for index option in $[9\%, > 100\%]$

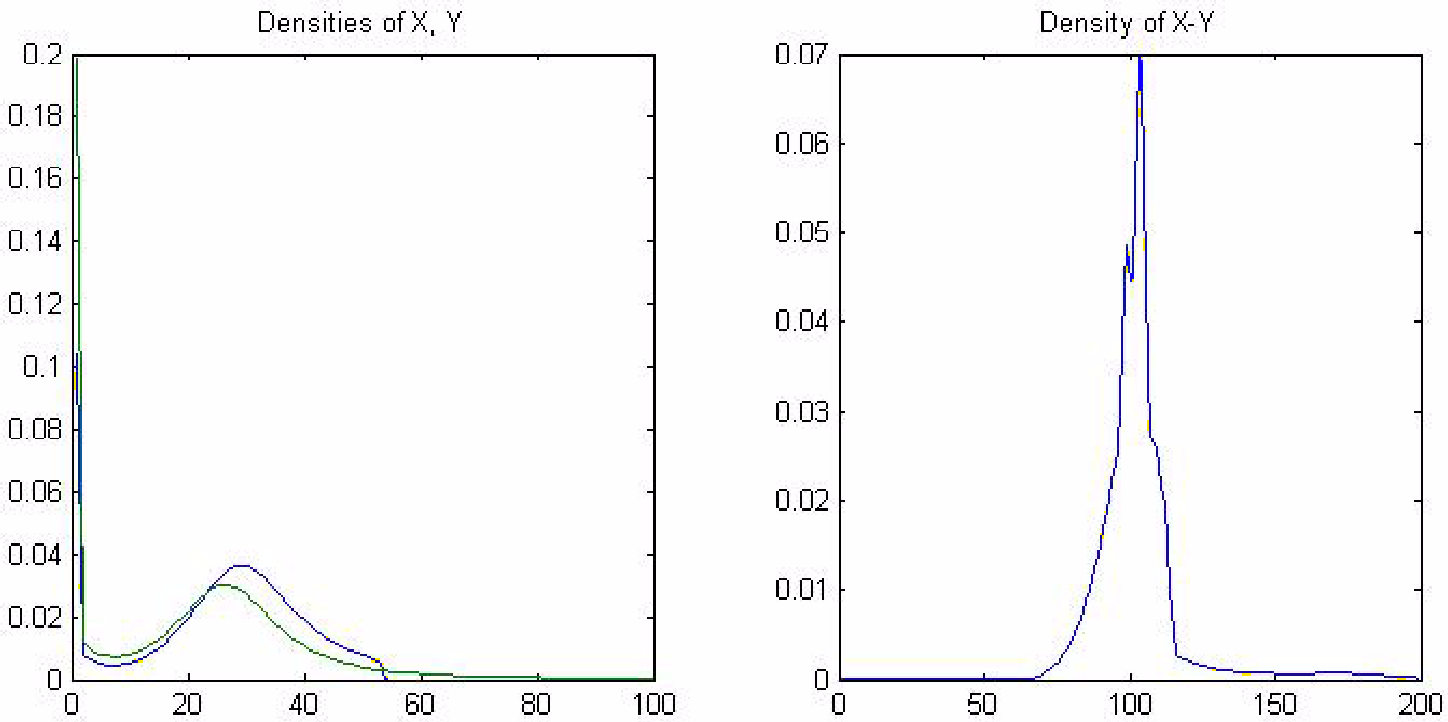




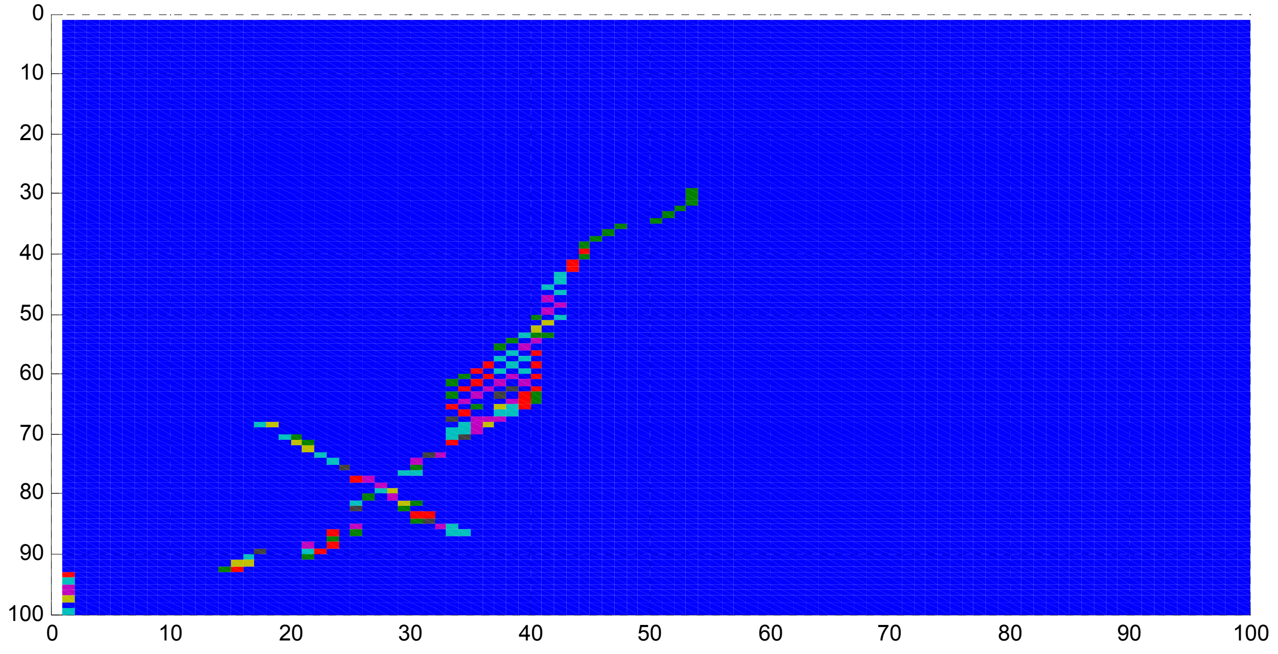

Density of X+Y
High option on index
Low option on index
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
35
40

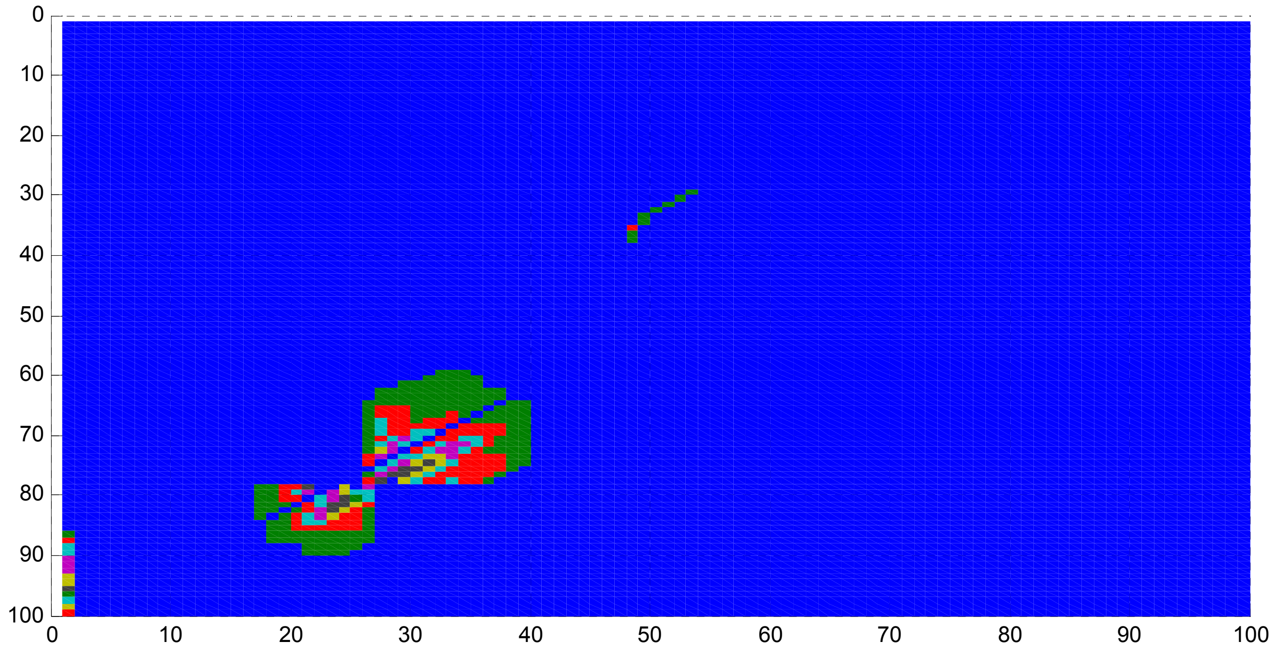
11 Results, CMS Spread Options

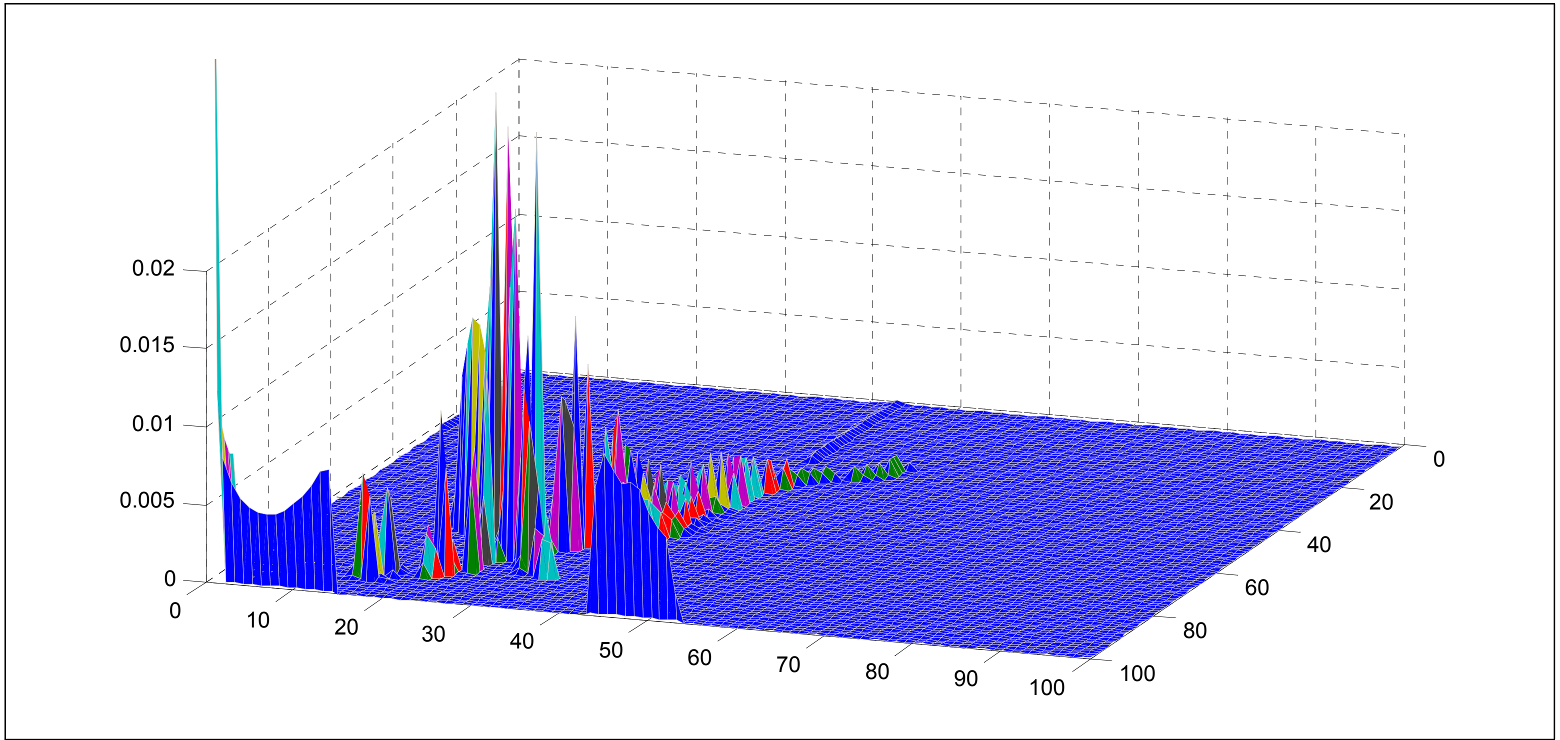
- 15y option on 30y-2y CMS spread in EUR, November 2010
- Here are marginal and spread densities

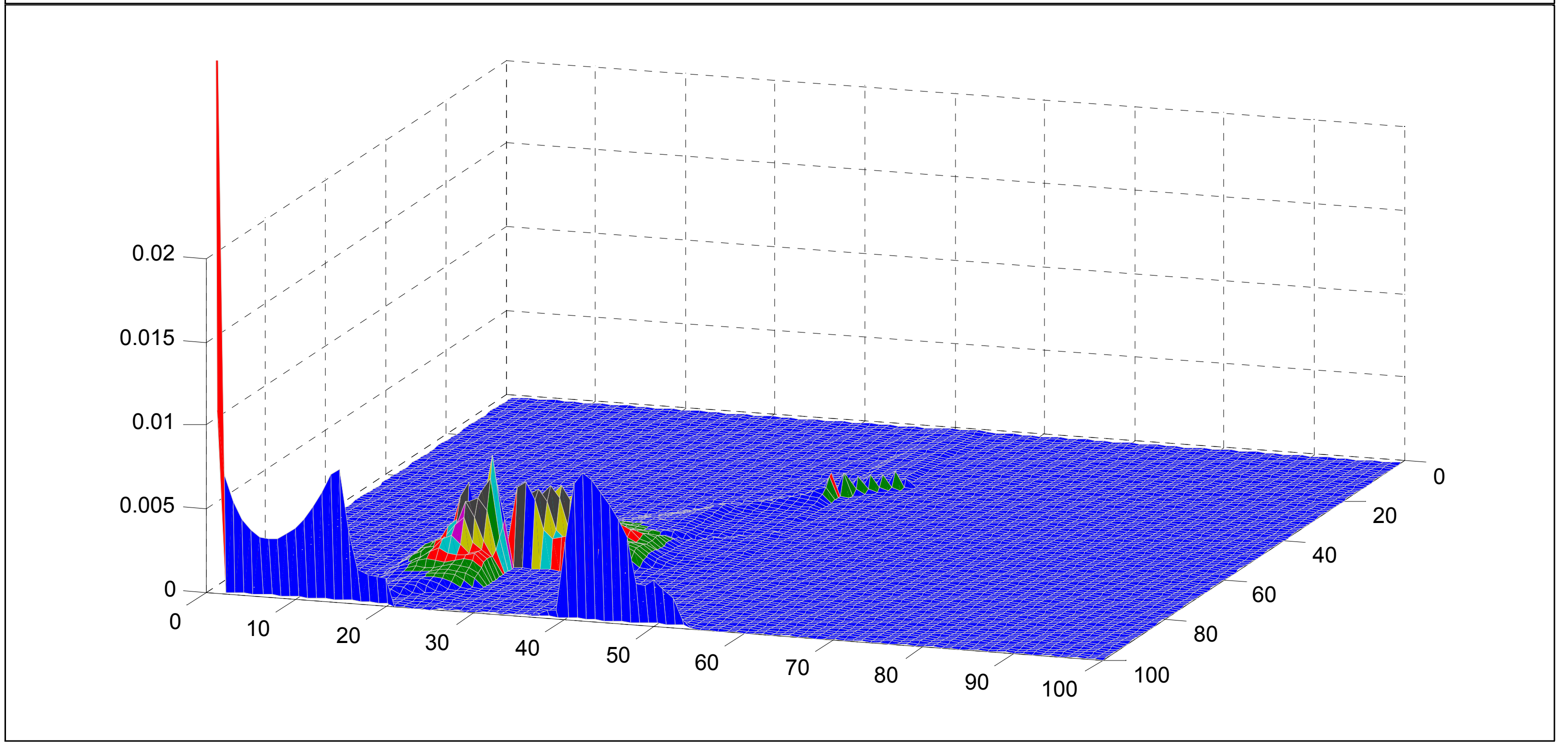


- Two extremes on next slides
- Implied Gaussian copula correlation for index option is in $[16\%, 49\%]$

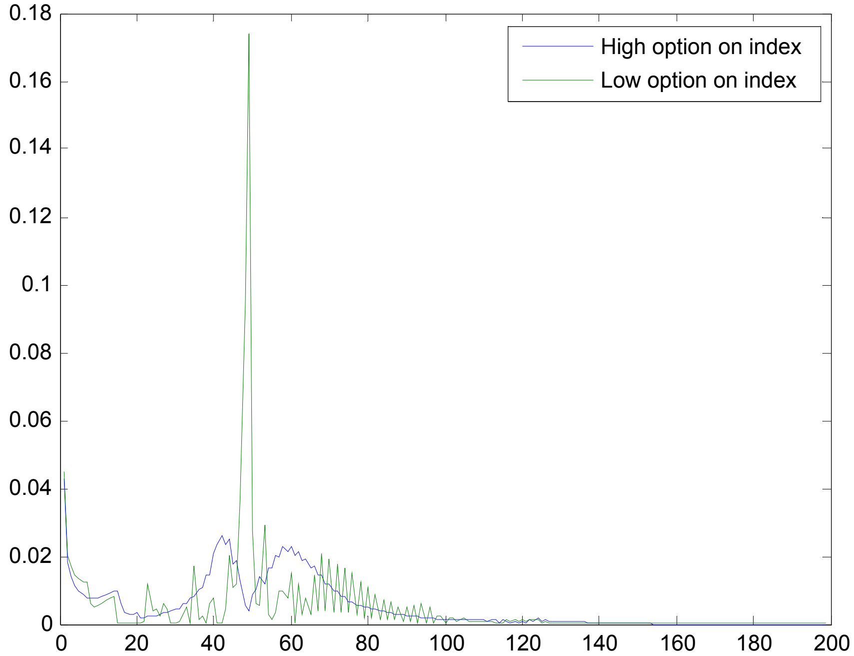

0
10
20
30
40
50
60
70
80
90
100
0
10
20
30
40
50
60
70
80
90
100


0
10
20
30
40
50
60
70
80
90
100
0
10
20
30
40
50
60
70
80
90
100


0.02
0.015
0.01
0.005
0
0
10
20
30
40
50
60
70
80
90
100
0
20
40
60
80
100


0.02
0.015
0.01
0.005
0
0
10
20
30
40
50
60
70
80
90
100
0
20
40
60
80
100

- Density of $X + Y$

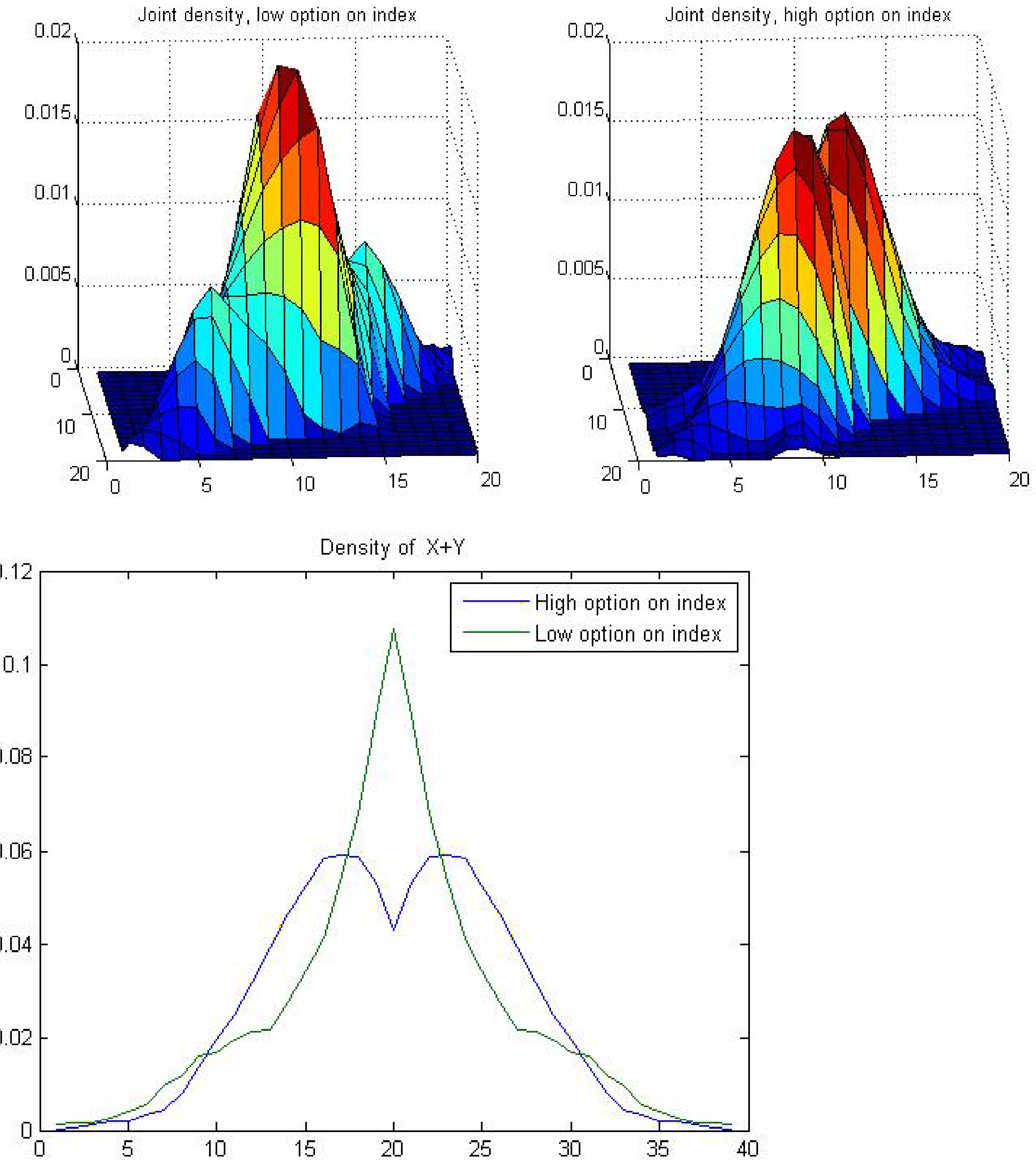


Smooth Density

- The simplex algorithm will assign non-zero values to as few $p_{i,j}$'s as it can get away with – on the order $4N$. The remaining $\approx N^2 - 4N$ points will be zero
- The densities will not be recognized as "valid" densities of financial underlyings
- Can solve this by moving from LP to QP – Quadratic Programming.
- Amend the target by smoothness conditions, i.e.

$$T(p) + w_{smooth} \sum_i \sum_j (p_{i,j} - p_{i,j+1})^2 + \cdots \to \min$$

  or quadratic deviation from a prior (smooth) density
- Results for Gaussian case. Implied Gaussian copula correlation is in the range [49%,82%]


Joint density, low option on index
Joint density, high option on index
Density of X+Y
High option on index
Low option on index

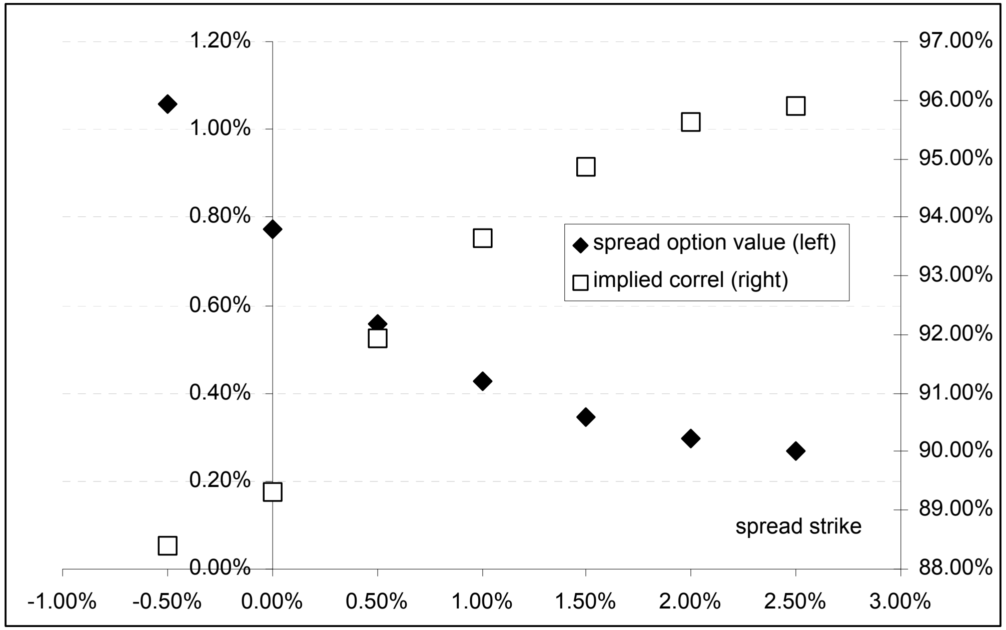
13 Coarse Constraints

- Interpolating spread volatility smile to get the whole smile density from (very) few points may introduce problems/arbitrage
- Convenient for theoretical discussions but unnecessary for numerical LP/QP formulation
- Spread options *for observable strikes only* can be incorporated as constraints into the LP/QP formulation (still linear functions of p)
- The QP algorithm will then effectively interpolate the spread option smile (arbitrage free):

$$\|d - d_0\|^2 \to \min$$
$$\text{subject to repricing chosen spread options}$$
$$\text{subject to Frechet min/max bounds}$$

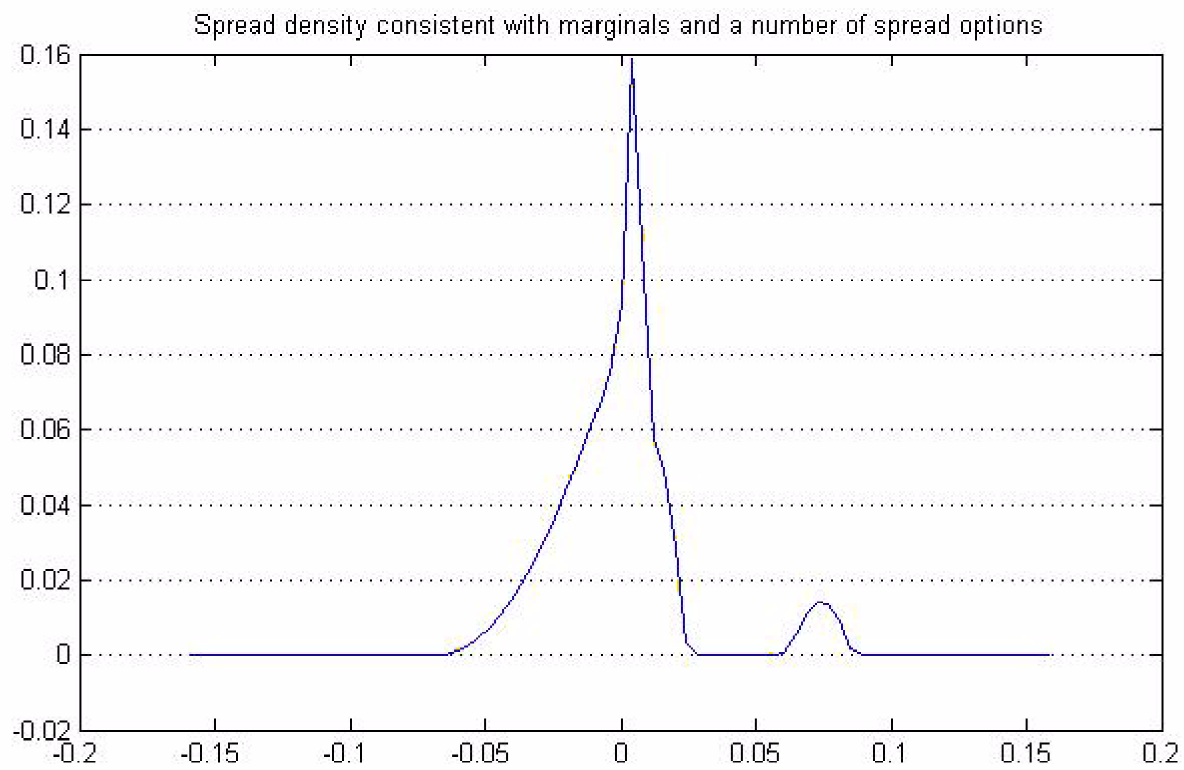
- Constraints are linear. Here p_0 is a smooth prior. Min/max constraints come from D/AD copulas and marginals

- Let us look at only a few strikes



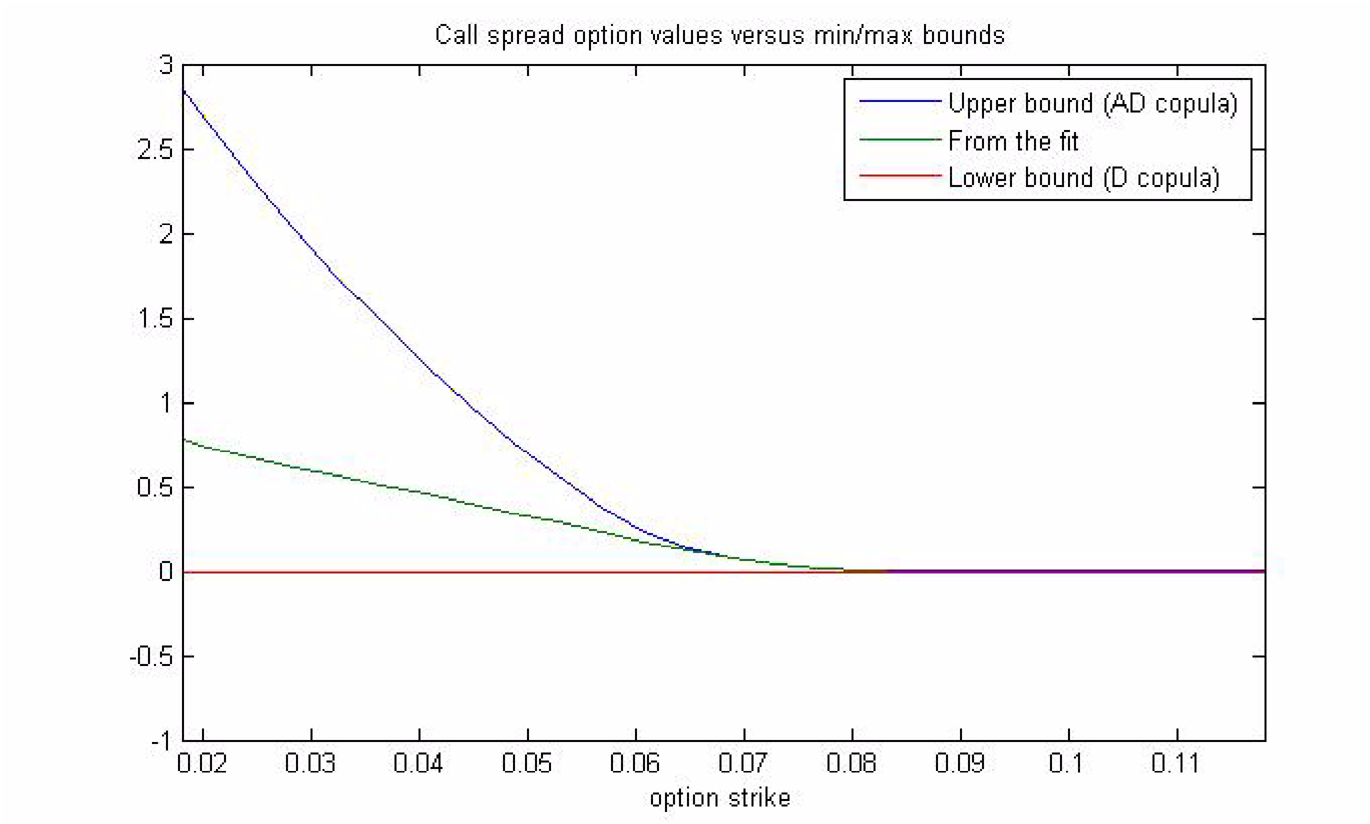
- By inspection it seems we should be able to construct a nice smooth density consistent with these observations

- It turns out however that we get something strange



- Note two humps and an area of zero probability at $\approx [2.5\%, 6\%]$
- These features are persistent even if we try to penalize them by modifying the QP problem accordingly

- What is going on?



- The spread option prices for the fitted density bump against the upper bound
- The second density hump is there to reprice strike=2.5% spread option
- It cannot be moved to the "left" – this will reduce 2.5% option price. So it has to be "smeared out".
- But we cannot move probability mass to the right because of the upper bound!

15 Finding Arbitrage

- In fact the same features persist even if we try just 4 spread options with strikes -0.5%, 0%, 0.5% and 1.0%
- Conclusion: presence of bounds makes marking even a small number of strikes for spread options potentially tricky
- Is there an arbitrage with such spread density?
- Can we just sell the digital call spread on S for $[2.5\%, 6\%]$?
  - Not an arbitrage as we have not locked it in
  - Need to use marginals somehow
  - But not a triangle arbitrage (as the spread density satisfies min/max conditions by construction)
- First test – see if the feasible solution to the joint density problem (7) exists
  - It does not
  - (8) returns a solution with not all slack variables at 0

- Recall Farkas' lemma that says that if the joint density does not exist then there exist vectors f, g such that

$$\sum_j f_j c_j + \sum_i g_i r_i > \sum_k \max_{(i',j')\in D_k} \left\{ f_{i'} + g_{j'} \right\} d_k. \tag{9}$$

- Finding such vectors would give us an arbitrage strategy as we can sell payoffs f and g on X and Y, buy the spread envelope payoff on $X - Y$, receive money upfront (by (9)) but have a non-negative payoff at expiry *under all conditions*

- Finding these vectors – another LP problem! Find f, g, h such that

$$\sum f_j c_j + \sum g_i r_i - \sum h_k d_k \to \max \tag{10}$$

subject to

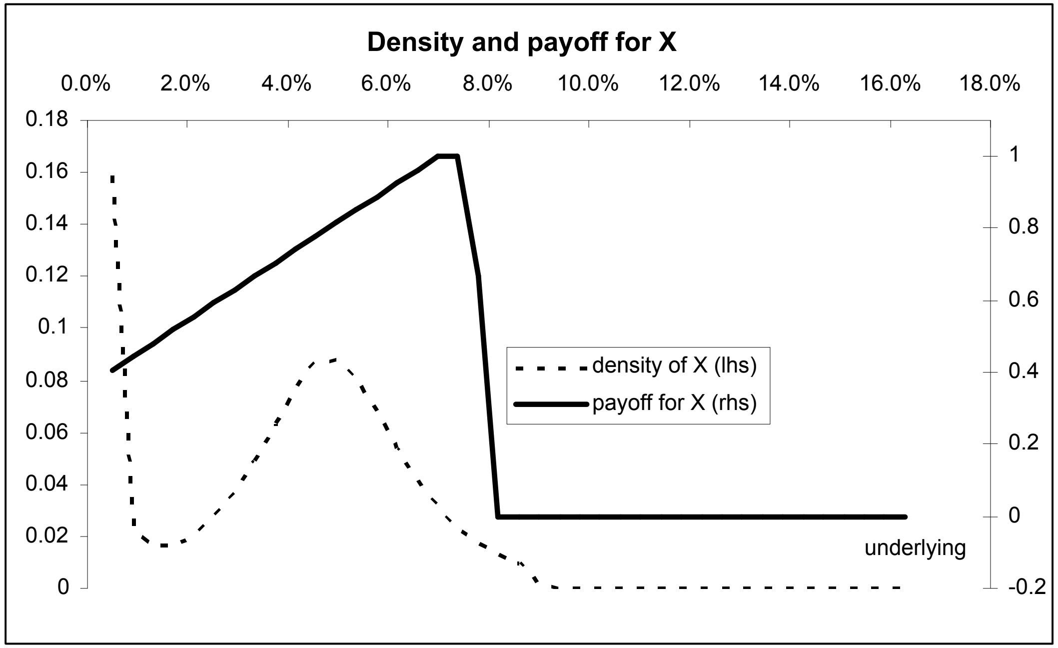
$$h_k \geq f_{i'} + g_{j'} \text{ for all } (i', j') \in D_k$$

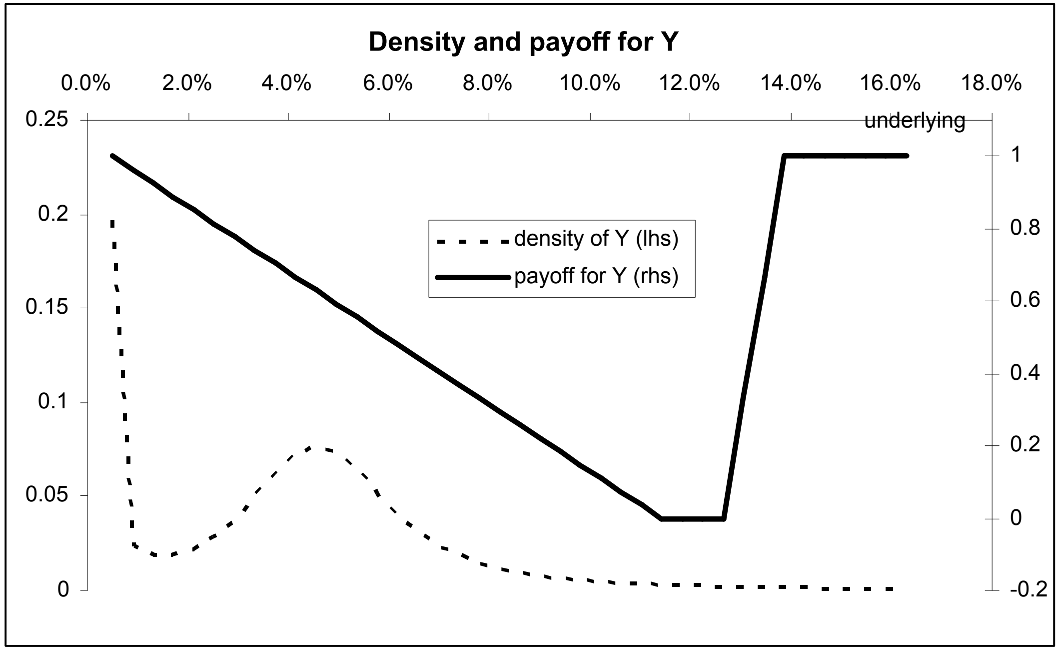
for all k. Also need constraints on the absolute size to avoid infinite solutions when exist, e.g. $|f_i|, |g_j| \leq 1$.

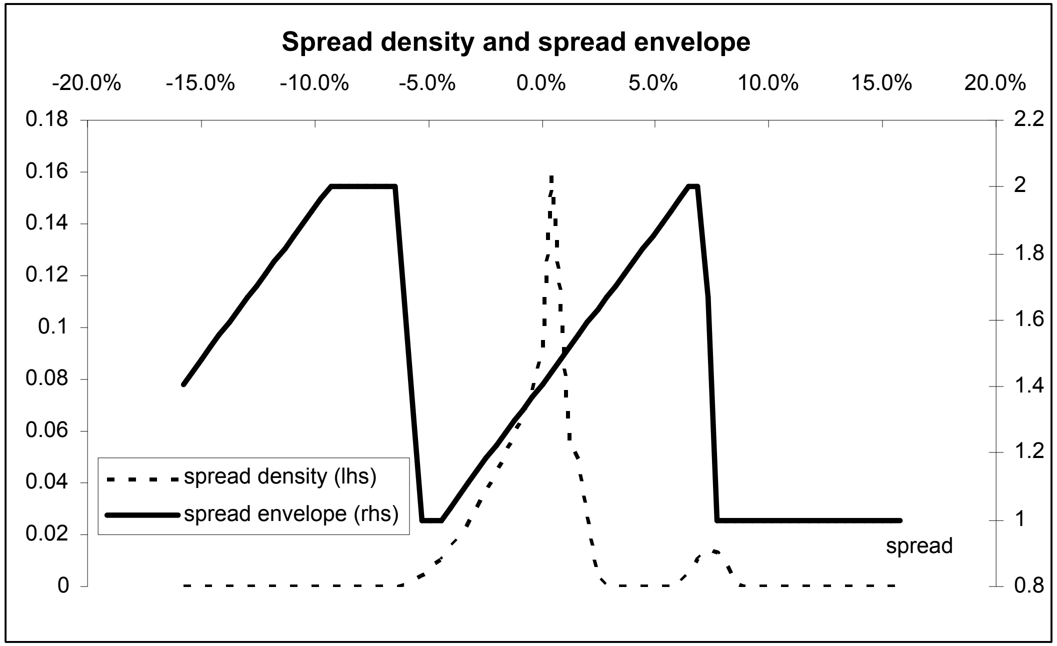
- Arbitrage is possible if the objective function in (10) is positive

- This is a dual problem to the existence problem (7)

- Here is the result







- Scale: with the notionals as displayed, the riskless profit is about 0.01; this translates into a few basis points of the notional
- This is an earlier-promised example of marginals/spread densities satisfying min/max conditions (and hence absence of triangle arbitrage) that nonetheless exhibit arbitrage

17 Conclusions

- Problem of existence of a joint distribution consistent with marginals and the spread can be solved theoretically using Farkas' lemma
  - The solution has clear financial interpretation: absence of spread envelope arbitrage implies joint distribution existence
  - Practically not very useful as too many payoffs to check
- From a practical standpoint, existence and "extreme" solutions can be determined in an efficient way numerically using Linear Programming techniques
- Same techniques can be used to find spread envelope arbitrage
- Marking spread options even for a few strikes can be tricky given effects from Frechet lower/upper bounds
- Spread envelope arbitrage can exist even if triangle arbitrage (or violation of Frechet bounds) does not

References

[AP10] Leif B.G. Andersen and Vladimir V. Piterbarg. *Interest Rate Modeling, in Three Volumes*. Atlantic Financial Press, 2010.

[Aus10] Peter Austing. Valuing multi-asset options on foreign exchange: A joint density repricing the cross smile. Submitted to Risk, 2010.

[McC11] Paul McCloud. The CMS triangle arbitrage. *Risk*, 1:126–131, 2011.