Symmetry-Maximizing Fair Values

Zachary McNulty

1 Introduction

The *central limit order book (CLOB)* is one of the most common mechanisms for pairing buyers and sellers in electronic marketplaces, including major exchanges such as NYSE, NASDAQ, and LSE. At any point in time, the CLOB records the prices and quantities at which participants are willing to buy or sell an asset. Crossing orders are matched and removed, leaving bids and asks separated by a positive *spread*. The book changes as orders are matched, added, modified, or canceled, and its dynamics provide a rich source of information about market sentiment.

The bid side of the order book is typically encoded as list of prices $\{b_i\}$ at which orders are placed, where $b_1 > b_2 > \cdots > b_n$, and a list of total quantities $\{\beta_i\}$ at each price level. Likewise, the ask side has lists $\{a_i\}, \{\alpha_i\}$ for $a_1 < a_2 < \cdots < a_m$. The highest buy offer b_1 and lowest sell offer a_1 are called the *best bid* (BB) and *best ask/offer* (BO) respectively, and together they are called the BBO. The distance $s := a_1 - b_1$ is called the *spread*, and, since orders with crossing prices are matched and removed from the order book, it is strictly positive.

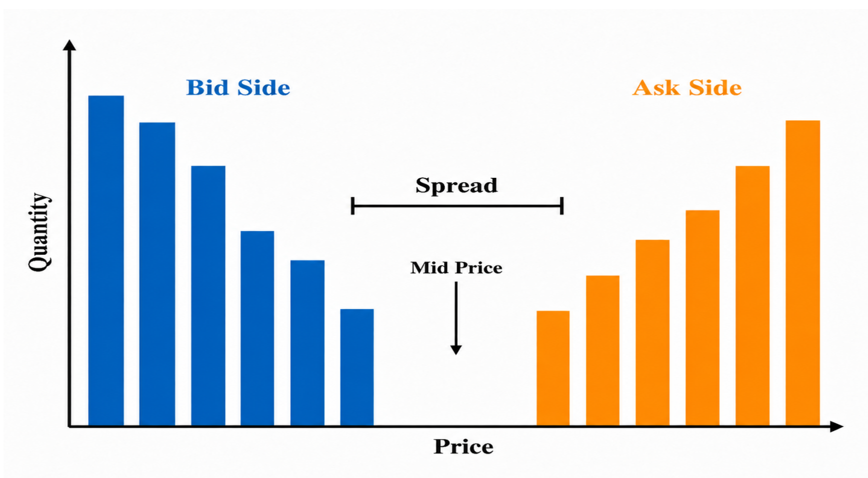


Figure 1: Limit order book.

Note in this view of the market, there is no one agreed upon price of the asset that people are willing to trade at. Instead, there is a range of prices and quantities people are willing to buy at, and a range they are willing to sell at. Nonetheless, these bids and offers still provide some information about what market participants think the asset is reasonably worth, the asset's *fair value*: no reasonable person would place a bid higher than their fair value, nor an ask lower.

Many approaches have been proposed for modeling CLOB dynamics, ranging from classical point-process and stochastic differential equation models to more recent deep-learning and agent-based methods; see [2] for a recent survey. Although these models differ substantially, they often share the goal of extracting fair-value information from displayed liquidity. This note proposes one such estimator.

A typical naive model for the fair value is just the *mid price* $v_m = (b_1 + a_1)/2$, the average of the best bid and best offer. A more nuanced measure is the *weighted mid price*:

$$v_w = \frac{b_1 \cdot \alpha_1 + a_1 \cdot \beta_1}{\beta_1 + \alpha_1} = v_m + \frac{s}{2} \cdot \frac{\beta_1 - \alpha_1}{\beta_1 + \alpha_1},$$

which incorporates information about the imbalance in quantities available at the BBO. Heuristically, when there are many more bids than asks we might expect this buying pressure to force the value of the asset upwards, and adjust our fair value accordingly.

The main downside of these two fair-value estimates is that they only use the first price level. Other heuristics, such as volume-weighted averages across several levels or market-impact estimates based on walking the book, use more of the displayed depth but still lack a clear variational motivation.

We aim to develop a fair-value estimate that captures the *shape* of the order book, has a clear variational characterization, and recovers familiar heuristics under natural assumptions.

We first view the bid side and ask side of the order book as two separate distributions μ_b, μ_a. One way of doing this is to view each quantity (or order) as a unit mass on the order book (and then possibly normalize):

$$\mu_b \propto \sum_i \beta_i \cdot \delta_{b_i}, \qquad \mu_a \propto \sum_j \alpha_j \cdot \delta_{a_j}. \tag{1.1}$$

Instead of treating each quantity identically, we can also weight the orders according to some scheme. Since it seems likely that orders near the spread are more informed, it is natural to weight these orders more. Namely, by introducing a discount $\lambda \in (0, 1]$, we could choose to add more mass to the larger bids and the smaller asks, the orders nearer to the spread:

$$\mu_b \propto \sum_i \lambda^{-b_i} \cdot \beta_i \cdot \delta_{b_i}, \qquad \mu_a \propto \sum_j \lambda^{a_j} \cdot \alpha_j \cdot \delta_{a_j}. \tag{1.2}$$

In the rest of this note, we will not assume μ_a, μ_b take any particular form, but these two examples are good to have in mind.

More generally, fix a strictly increasing homeomorphism $\phi : I \to \mathbb{R}$ from an admissible price domain I. We carry out the symmetry comparison in the transformed price coordinate $x = \phi(v)$. Namely, let

$$\mu_b^\phi := \phi_\# \mu_b, \qquad \mu_a^\phi := \phi_\# \mu_a$$

be the bid- and ask-side distributions in ϕ-space. We denote the symmetry-maximizing transformed value by $x_{s,\phi}$ and report the corresponding price-level fair value

$$v_{s,\phi} := \phi^{-1}(x_{s,\phi}).$$

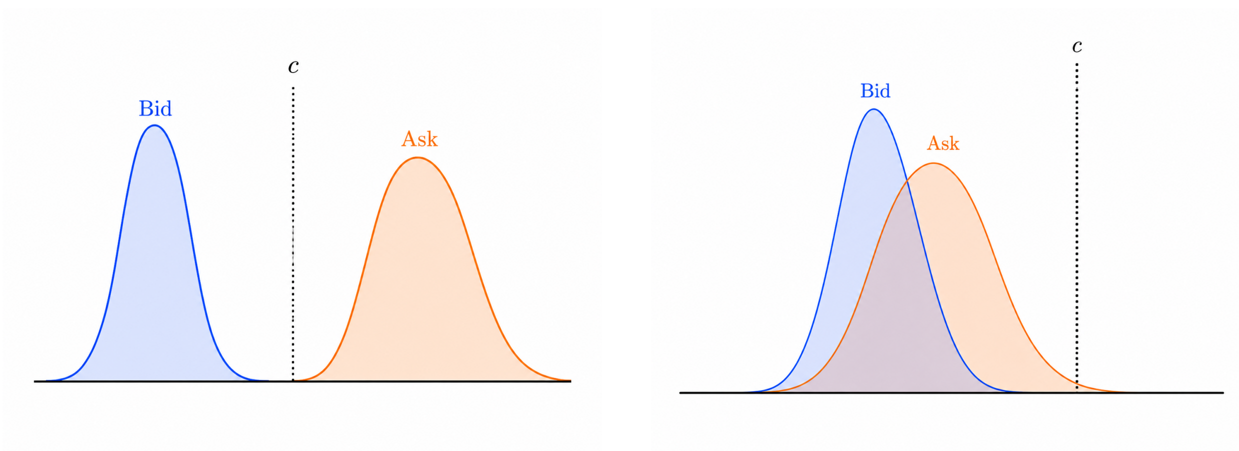
The choice $\phi(v) = v$ recovers the additive model developed below, while $\phi(v) = \log v$ gives a multiplicative model for positive prices and ensures that the resulting fair value remains positive.

To try to generalize our heuristics, consider the simple situation where the bid and ask books are perfect mirror images of each other in ϕ-space, separated possibly by some spread. In this case, buying and selling pressures are exactly equal, so the midpoint in the chosen coordinate is a natural candidate for the fair value. For the identity coordinate this is the usual mid price v_m; for a general

ϕ it is the point around which the transformed distributions are mirror images. In other words, the transformed midpoint solves:

$$x_{s,\phi} = \arg\min_{c\in\mathbb{R}} \; d\big((c-\cdot)_{\#}\mu_b^{\phi},\, (\cdot - c)_{\#}\mu_a^{\phi}\big), \qquad v_{s,\phi} = \phi^{-1}(x_{s,\phi}), \tag{1.3}$$

for any metric d on a suitable space of probability distributions. In general, we will not have such symmetry, but we can still seek to solve (1.3) for a fixed metric d. We call $v_{s,\phi}$ the *symmetry-maximizing value*; it implicitly depends on both d and the chosen price coordinate ϕ. A visualization of this procedure in transformed price space is given in Figure 2.



(a) Original order book. (b) Reflected ask book across c.

Figure 2: Visualization of the symmetry-maximizing fair value. The ask book is reflected across the candidate transformed fair value c, and the distance between the reflected ask book and the bid book is measured by some metric d. Then c is chosen to minimize this distance, and $\phi^{-1}(c)$ is reported in price units.

2 Incorporating Order Book Shape

There are many natural measures of "distance" between probability measures: total variation, Jensen-Shannon divergence, and χ^2-type distances. These are less suitable for the present application because they compare mass at fixed locations, making them sensitive to small shifts in order prices or quantities. Since financial markets are inherently noisy, we want a metric that is more robust under such perturbations, and which incorporates more of the *geometry* of the order book than these pointwise comparisons.

The Wasserstein distance is the natural candidate. In this section, we aim to study how the symmetry-maximizing fair value behaves when the underlying metric d is chosen to be this distance. In Theorem 2.1, we characterize the symmetry-maximizing fair value under the Wasserstein metric, and in Lemma 5.1 we quantify the stability of this fair-value with respect to small changes of the order book.

2.1 Wasserstein p-metric

This metric comes from the theory of optimal transport. It measures how difficult it is to convert one distribution into another by transporting it "piece by piece". The goal is to find a transportation plan that requires the least total distance the mass must be transported. Formally, it is defined as:

$$d_{W_p}(\mu,\nu) := \left(\inf_{(X,Y)\in\Pi(\mu,\nu)} \mathbb{E}|X-Y|^p\right)^{1/p},$$

where $\Pi(\mu,\nu)$ is the space of all couplings of μ,ν, joint distributions of (X,Y) that have marginals μ,ν respectively. Namely, for each $x \in \mathbb{R}$ our coupling $\pi(x,\cdot)$ specifies where in ν the mass $\mu(x)$ must be sent, and we search for a way of doing this that minimizes the total transportation distance. The below theorem shows (1.3) takes the form of a simple convex optimization problem. First, we introduce some notation for equation (1.2).

$$B_i := \frac{\sum_{j=1}^{i} \lambda^{-b_j}\cdot \beta_j}{\sum_{j=1}^{n} \lambda^{-b_j}\cdot \beta_j}, \qquad A_i := \frac{\sum_{j=1}^{i} \lambda^{a_j}\cdot \alpha_j}{\sum_{j=1}^{m} \lambda^{a_j}\cdot \alpha_j}, \qquad A_0 = B_0 = 0,$$

$$\widetilde{b}_i := \phi(b_i), \qquad \widetilde{a}_i := \phi(a_i),$$

$$z_k := \inf\big(\{A_i : A_i > z_{k-1}\} \cup \{B_i : B_i > z_{k-1}\}\big), \qquad \Delta z_\ell := z_\ell - z_{\ell-1}, \qquad z_0 := 0.$$

The A_i and B_i record the cumulative masses of the ask and bid distributions as we move away from the best offer and best bid, while $\widetilde{a}_i$ and $\widetilde{b}_i$ are the corresponding prices in ϕ-space. The z_k are the ordered values of the combined cumulative masses. Define $f(k) = i$ if $z_k \in [B_{i-1}, B_i)$ and $g(k) = j$ if $z_k \in [A_{j-1}, A_j)$. We are now ready to relate these quantities to the symmetry-maximizing fair value.

Theorem 2.1. *For any $p \geq 1$, the transformed fair value $x_{s,\phi}$ solves the convex optimization problem*

$$x_{s,\phi} = \arg\min_{c\in\mathbb{R}} \sum_{\ell=1}^{m+n} \Delta z_\ell \cdot |2c - \widetilde{b}_{f(\ell)} - \widetilde{a}_{g(\ell)}|^p, \qquad v_{s,\phi} = \phi^{-1}(x_{s,\phi}). \tag{2.1}$$

Moreover, in the case $p = 2$ the explicit solution is

$$x_{s,\phi} = \frac{1}{2}\Big(m_2(\mu_b^\phi) + m_2(\mu_a^\phi)\Big), \qquad v_{s,\phi} = \phi^{-1}(x_{s,\phi}). \tag{2.2}$$

Remark 2.2. As shown in Corollary A.4, $x_{s,\phi}$ is the mean of the W_2 barycenter of μ_b^ϕ and μ_a^ϕ. Pushing this barycenter forward through ϕ^{-1} therefore gives an estimate of the full fair-value distribution rather than only a point estimate.

3 A Simple Model For Market Makers

The following model is not intended as a literal description of order-book formation. It is a stylized mechanism showing how a symmetry-based fair value can arise when market makers place bid and ask quotes around latent fair values. In the balanced case, a paired-order model is most naturally written in the chosen price coordinate. Writing $x = \phi(v)$ for the transformed fair value, we set

$$\phi(b) = x - h, \qquad \phi(a) = x + h,$$

where h is a latent half-spread in ϕ-space, reflecting uncertainty, inventory concerns, or compensation for supplying liquidity. With imbalanced buying and selling pressure, we instead allow asymmetric quotes:

$$\phi(b) = x - \theta^{-1}h, \qquad \phi(a) = x + (1-\theta)^{-1}h,$$

where $\theta \in (0,1)$ is an exogenous imbalance parameter. As θ approaches zero, the bid is moved farther from the fair value in transformed price space; as θ approaches one, the analogous effect occurs on the ask side. Formally, we study

$$\begin{aligned} x = \phi(v) &\sim \mu \in \mathbb{P}(\mathbb{R}) && \text{(transformed fair-value distribution)} \\ h &\sim \nu \in \mathbb{P}((0,\infty)) && \text{(transformed half-spread distribution)} \\ \theta &\in (0,1) && \text{(imbalance).} \end{aligned}$$

The model implies

$$x = \theta\phi(b) + (1-\theta)\phi(a), \qquad h = \theta(1-\theta)\big(\phi(a) - \phi(b)\big),$$

and the corresponding price-level fair value is $v = \phi^{-1}(x)$.

For the identity coordinate, this is the additive model. For $\phi(v) = \log v$, the paired fair value and transformed half-spread become

$$v = b^\theta a^{1-\theta}, \qquad h = \theta(1-\theta)\log(a/b),$$

so the model is multiplicative and cannot produce a negative fair value. Thus, if we knew which bid was paired with each ask, we could aggregate these to get reasonable estimates of the fair value and half-spread distributions:

$$\hat{\mu} := \frac{1}{n}\sum_{i=1}^{n} \delta_{\theta\phi(b_i)+(1-\theta)\phi(a_i)}, \qquad \hat{\nu} := \frac{1}{n}\sum_{i=1}^{n} \delta_{\theta(1-\theta)(\phi(a_i)-\phi(b_i))}.$$

More generally, if we knew the pairings of bids and asks and were interested in the p^{th} mean $m_p(\mu) := \arg\min_c \mathbb{E}_\mu|X-c|^p$ we could approximate its transformed fair-value coordinate by

$$m_p(\hat{\mu}) = \arg\min_c \sum_{i=1}^{n} |c - \theta\phi(b_i) - (1-\theta)\phi(a_i)|^p,$$

and then map this estimate back to price space using ϕ^{-1}. As we will see later, this has the same form as our symmetric fair-value objective (4.3) once a pairing of bids and asks has been chosen. Since the identity of the issuer of each order is unknown, the main problem becomes choosing a reasonable pairing. Equivalently, the symmetry-maximizing value can be viewed as a relaxation of this latent pairing problem.

Of course, a real order book need not admit a literal one-to-one decomposition into bid/ask pairs. The paired model should instead be viewed as a finite approximation to a coupling between the bid-side and ask-side distributions.

3.1 Determining optimal bid/ask pairing

Suppose we have n bids $b_1 \geq \cdots \geq b_n$ and n asks $a_1 \leq \cdots \leq a_n$. Write $\widetilde{b}_i := \phi(b_i)$ and $\widetilde{a}_i := \phi(a_i)$. Since ϕ is increasing, the transformed quotes have the same ordering. Pairing these up amounts to choosing a permutation $\tau \in S_n$ and pairing b_i with $a_{\tau(i)}$. If we assume this is the true pairing, the above generative model specifies exactly its probability distribution. Finding the optimal way of pairing up asks and bids then amounts to solving:

$$\arg\max_{\tau \in S_n} \mathbb{P}_{x,h}\left((\widetilde{b}_i, \widetilde{a}_{\tau(i)})_{i=1}^{n}\right)$$

$$
\begin{aligned}
&= \arg\max_{\tau \in S_n} \sum_{i=1}^{n} \log\left(\mathbb{P}_{x,h}(\widetilde{b}_i, \widetilde{a}_{\tau(i)})\right) \\
&= \arg\max_{\tau \in S_n} \sum_{i=1}^{n} \log\left(\mathbb{P}_h(\theta(1-\theta)\cdot(\widetilde{a}_{\tau(i)} - \widetilde{b}_i)\right) + \log\left(\mathbb{P}_x(\theta\widetilde{b}_i + (1-\theta)\cdot\widetilde{a}_{\tau(i)})\right).
\end{aligned}
$$

If we assume the transformed half-spread and fair-value coordinates have densities $\mathbb{P}_h = e^{f(x)}\,dx$ and $\mathbb{P}_x = e^{g(x)}\,dx$ then this reduces to:

$$
= \arg\max_{\tau \in S_n} \sum_{i=1}^{n} f\Big(\theta(1-\theta)\cdot(\widetilde{a}_{\tau(i)} - \widetilde{b}_i)\Big) + g\Big(\theta\widetilde{b}_i + (1-\theta)\widetilde{a}_{\tau(i)}\Big).
$$

The next lemma gives conditions under which this score is maximized when bids and asks are paired in the natural way: from inside the spread outwards. Towards this end, we introduce the discrete differences

$$
\Delta_r f(z) := f(z+r) - f(z).
$$

Concavity of f is equivalent to $\Delta_r f$ being nonincreasing for each $r > 0$, so the second discrete derivative $\Delta_t\Delta_r f$ is nonpositive. The following identifies the optimal pairing when one of the latent distributions is sufficiently more log-concave than the other.

Lemma 3.1. *Suppose the transformed fair-value distribution is $\mu = N(0,\sigma^2)$ and the transformed half-spread distribution has density $\nu = e^{f(x)}\,dx$. If for all $z \in \mathbb{R}$ and $t, r > 0$:*

$$
\Delta_t\Delta_r f(z) \geq -\frac{1}{\theta(1-\theta)\sigma^2}\cdot rt. \tag{3.1}
$$

If $\widetilde{b}_n \leq \cdots \leq \widetilde{b}_1 \leq \widetilde{a}_1 \leq \cdots \leq \widetilde{a}_n$ then:

$$
\mathrm{id} = \arg\max_{\tau \in S_n} \sum_{i=1}^{n} f\Big(\theta(1-\theta)\cdot(\widetilde{a}_{\tau(i)} - \widetilde{b}_i)\Big) + g\Big(\theta\widetilde{b}_i + (1-\theta)\widetilde{a}_{\tau(i)}\Big).
$$

Namely, the optimal pairing is moving from inside the spread outwards. Moreover, if the reverse inequality is true the optimal pairing is from left to right.

Note, ignoring the constant depending on θ, the right hand side of equation (3.1) is just the second discrete derivative of $g(x) = -\frac{x^2}{2\sigma^2}$, the log-density of the Gaussian. Thus, the condition is saying that our fair value distribution is more log-concave than our spread distribution in a strong way.

Remark 3.2. If f is twice differentiable, then a Taylor series expansion shows $\Delta_t\Delta_r f(z) \approx rt \cdot f''(z)$. Thus the assumption of the theorem roughly says the second derivative is bounded below by some sufficiently large constant.

Remark 3.3. In practice, since bids and asks can only be placed at discrete tick-sizes we could effectively relax the assumptions to only $z, t, r \in \mathbb{Z}_{>0}$.

Example 3.4. In particular, one natural choice is $\nu = \mathrm{Gamma}(\lambda, k)$. Here:

$$
f(x) = -\lambda x + (k-1)\log(x) + C_{\lambda,k} \quad \to \quad \Delta_t\Delta_r f(x) = (k-1)\log\left(1 - \frac{rt}{(x+r)(x+t)}\right).
$$

In the exponential case ($k = 1$) or when $k < 1$ then assumption (3.1) is trivially satisfied because the second discrete derivative is nonnegative. Otherwise, if $k > 1$ then $\Delta_t\Delta_r f(x) \to -\infty$ as $x \downarrow 0$, so the assumptions are not met. In this regime, the Gamma distribution has sufficient mass away from zero, so large spreads are not as heavily penalized, and the optimal pairing is no longer necessarily from inside out.

Note that our above proof of Lemma 3.1 actually shows something slightly stronger.

Corollary 3.5. *If $\widetilde{b}_n \leq \cdots \leq \widetilde{b}_1 \leq \widetilde{a}_1 \leq \cdots \leq \widetilde{a}_n$ and for all $t, r > 0$:*

$$\Delta_{\theta t}\, \Delta_{(1-\theta)r}\, f(z) \geq \Delta_r \Delta_t g(w),$$

for $z \in [0, \widetilde{a}_n - \widetilde{b}_n]$ and $w \in [\widetilde{b}_n, \widetilde{a}_n]$, then we can see:

$$\mathrm{id} = \arg\max_{\tau \in S_n} \sum_{i=1}^{n} f\Big(\theta(1-\theta) \cdot (\widetilde{a}_{\tau(i)} - \widetilde{b}_i)\Big) + g\Big(\theta \widetilde{b}_i + (1-\theta)\widetilde{a}_{\tau(i)}\Big).$$

Namely, the optimal pairing is moving from inside the spread outwards.

Under this simple model, the optimal pairing is one of two natural pairings: "inside-out," starting from the best bid and best offer, or "left-to-right," starting from the worst bid and best offer. The Wasserstein formulation in Section 2 can be viewed as a continuous relaxation of this pairing problem.

4 Introducing Quantity Imbalances

A natural criticism of our work in Section 2 is that we lose information about imbalance between the bid and ask sides of the order book, imbalances which fair-value estimates like the weighted mid v_w discussed earlier seek to exploit. Thus, it may only be a decent estimate of the fair value in situations when the total number of bids and total number asks are roughly equal. In this section, we discuss a way of expanding our previous approach to incorporate this kind of information.

In this section, we drop our assumption that μ_a, μ_b be probability measures, so they can take unnormalized forms like in equation (1.2). One simple measure of imbalance is then just the following ratio of the total masses m_a, m_b of these measures:

$$\theta = \frac{m_a}{m_a + m_b}. \tag{4.1}$$

In the case of the unnormalized measures (1.1), this is just the ratio of the total volume on the bid book and the total volume on the ask book. Using this, we can extend (1.3) in a way that incorporates the imbalances in the bid and ask books:

$$x^\theta_{s,\phi} = \arg\min_{c \in \mathbb{R}} d\left((c - 2\theta\,\cdot)_\# \frac{\mu_b^\phi}{m_b},\ (2(1-\theta)\ \cdot - c)_\# \frac{\mu_a^\phi}{m_a}\right), \qquad v^\theta_{s,\phi} = \phi^{-1}(x^\theta_{s,\phi}). \tag{4.2}$$

We then have the corresponding extension of Theorem 2.1, which in the case $p = 2$ just says we replace the standard average in Theorem 2.1 by an average weighted by the imbalance.

Theorem 4.1. *Let $\theta \in (0,1)$. Then for any $p \geq 1$, the transformed fair value $x^\theta_{s,\phi}$ solves the convex problem*

$$x^\theta_{s,\phi} = \arg\min_{c \in \mathbb{R}} \sum_{\ell=1}^{m+n} \Delta z_\ell \cdot |2c - 2\theta \widetilde{b}_{f(\ell)} - 2(1-\theta)\widetilde{a}_{g(\ell)}|^p, \qquad v^\theta_{s,\phi} = \phi^{-1}(x^\theta_{s,\phi}). \tag{4.3}$$

Moreover, in the case $p = 2$ the explicit solution is

$$x^\theta_{s,\phi} = \theta m_2(\mu_b^\phi) + (1-\theta) m_2(\mu_a^\phi), \qquad v^\theta_{s,\phi} = \phi^{-1}(x^\theta_{s,\phi}). \tag{4.4}$$

Remark 4.2. In the case there is only one price level and (1.1), or when μ_a, μ_b only put mass on the first price level proportional to the quantity, and we choose θ by (4.1), then the identity coordinate gives $v^\theta_{s,\mathrm{id}} = v_w$, the weighted mid price. The log coordinate instead gives the weighted geometric mid $v^\theta_{s,\log} = b_1^\theta a_1^{1-\theta}$.

We can of course consider other imbalance measurements $\theta \in (0,1)$ other than (4.1). Some other reasonable alternatives might compare recent trade volumes, higher moments of μ_a, μ_b, or possibly even the age of different orders on the book. The main point is that we can easily incorporate this kind of information into our framework by simply changing the way we warp the bid and ask distributions before comparing them.

5 Stability of the Fair Value

Recall the main motivation behind using Wasserstein metrics in the first place was that they provide a natural robustness to small perturbations in the underlying distribution. In this section, we extend classical stability results for the Wasserstein metric to our symmetrized fair value. This gives the following continuity property under d_{W_2}.

Lemma 5.1. *For transformed price measures μ, ν, μ', ν', the fair-value coordinate $x^\theta_{s,\phi} := \phi(v^\theta_{s,\phi})$ under d_{W_2} in (4.4) satisfies for any $p \geq 1$:*

$$\left| x^\theta_{s,\phi}(\mu,\nu) - x^{\theta'}_{s,\phi}(\mu',\nu') \right| \leq \theta' d_{W_p}(\mu,\mu') + (1-\theta') d_{W_p}(\nu,\nu') + |\theta - \theta'| \cdot d_{W_p}(\mu,\nu). \tag{5.1}$$

In particular, if $\mu' = t^\# \mu, \nu' = s^\# \nu$ are pushforward measures then:

$$\left| x^\theta_{s,\phi}(\mu,\nu) - x^{\theta'}_{s,\phi}(\mu',\nu') \right| \leq \theta' ||t - id||_{L^p(\mu)} + (1-\theta') ||s - id||_{L^p(\nu)} + |\theta - \theta'| \cdot d_{W_p}(\mu,\nu). \tag{5.2}$$

The result is not exactly surprising: equation (4.4) shows the transformed fair value is completely determined by the means of μ, ν, so if we only perturb μ, ν by a small amount, their means should not change by much either.

Under d_{W_p} for $p \neq 2$, things get a little more interesting. Such an explicit closed-form bound is no longer possible. However, intuitively, in equation (4.3) we are solving an optimization problem which is strictly convex in c for $p > 1$ and depends smoothly on the parameters $z_\ell, \theta, \widetilde{a}_i, \widetilde{b}_j$. Hence we still expect some level of stability, which the following makes precise.

Lemma 5.2. *For any $p > 1$, if $\mu_n \xrightarrow{W_p} \mu, \nu_n \xrightarrow{W_p} \nu$, and $\theta_n \to \theta$, then*

$$x^{\theta_n}_{s,\phi}(\mu_n,\nu_n) \to x^\theta_{s,\phi}(\mu,\nu), \qquad v^{\theta_n}_{s,\phi}(\mu_n,\nu_n) \to v^\theta_{s,\phi}(\mu,\nu).$$

The difficulty in obtaining a quantitative bound analogous to Lemma 5.1 is that, for $p \neq 2$, the p^{th}-mean is no longer a linear functional of the underlying distribution, so the movement of the minimizer depends on the local curvature of the objective near its minimum rather than just on the perturbation level. The convergence in price space follows from continuity of ϕ^{-1}; a quantitative price-space bound follows whenever ϕ^{-1} is Lipschitz on the relevant range.

6 Concluding Remarks and Future Directions

In this paper, we proposed a fair-value estimate which maximizes the symmetry of the central limit order book around it in a chosen monotone price coordinate. This generalizes several common heuristics and gives a variational characterization of fair value. Under the Wasserstein distance, we derive explicit formulas and show stability under small perturbations of the order book. The identity coordinate recovers the weighted arithmetic mid in the one-level case, while the log coordinate gives its positive, multiplicative analogue.

Corollary A.4 suggests a natural extension: in the W_2 case, the transformed symmetrized fair value is the mean of the weighted barycenter of μ_a^ϕ and μ_b^ϕ, given by

$$\overline{\mu_\theta^\phi} := \arg\min_{\nu} \theta d_{W_2}(\nu, \mu_a^\phi)^2 + (1-\theta) d_{W_2}(\nu, \mu_b^\phi)^2.$$

Mapping the mean of this barycenter back through ϕ^{-1} identifies the fair value with a Wasserstein barycenter in transformed price space and suggests estimating a full fair-value distribution rather than only a point estimate. The same idea could also incorporate several order books for the same asset across different venues. Instead of merging all bids and asks directly, one could first form bid-side and ask-side barycenters across venues, reflect the ask-side barycenter, and then barycenter the two resulting distributions. This is similar in spirit to the aggregation procedure in [4]; it would be interesting to identify a market model under which such an estimator is optimal.

Another natural direction is empirical. One could simulate order books under standard CLOB models, such as the point-process or agent-based models surveyed in [2], and compare the resulting symmetry-maximizing value to the mid price, weighted mid price, and other benchmarks. Useful diagnostics would include the smoothness of the resulting time series, its visual reasonableness relative to the displayed book shape, and its usefulness as a feature for predicting short-term price movements.

Finally, the estimator studied here uses a static snapshot of the order book. A dynamic version could instead maximize symmetry across time, which may be more robust to noise. It would be useful to understand when symmetry arises in standard CLOB dynamics and how the resulting estimator performs empirically.

A Proof Appendix

We start by collecting some basic results about Wasserstein spaces which will be useful in our analysis. More details can be found in [3].

A.1 Properties of Wasserstein Spaces

Lemma A.1. *If* $\mu_n \xrightarrow{W_p} \mu$ *then* $\mathbb{E}_{\mu_n} X \to \mathbb{E}_\mu X$. *More specifically:*

$$\Big|\mathbb{E}_{\mu_n} X - \mathbb{E}_\mu X\Big| \le d_{W_p}(\mu_n, \mu).$$

Proof. Since for any coupling $\pi \in \Pi(\mu_n, \mu)$ we have for $p \ge 1$ by Jensen's inequality:

$$\left| \int_{\mathbb{R}} x \ \mu_n(dx) - \int_{\mathbb{R}} y \ \mu(dy) \right| \le \left(\int_{\mathbb{R}^2} |x-y|^p \ d\pi \right)^{1/p}.$$

Taking the infimum over such couplings provides the desired result. □

In fact, more is true: as shown in [3], convergence in W_p is equivalent to weak convergence and convergence of p^{th} moments. Below, we show for distributions on $\mathbb{R}$ that the Wasserstein distance can be easily computed using the quantile function Q_μ. Again, this result can be found in [3].

Lemma A.2. *For any* $p \ge 1$ *the map* $\mu \mapsto Q_\mu$ *from* $W_p(\mathbb{R})$ *to* $L_p([0,1], \mathrm{Leb})$ *is an isometry. Namely for* $\mu, \nu \in W_p(\mathbb{R})$*:*

$$d_{W_p}(\mu, \nu) = \left(\int_0^1 |Q_\mu(z) - Q_\nu(z)|^p \ dz \right)^{1/p}.$$

Lemma A.3. *Given* $\{\mu_i\}_{i=1}^n \in W_2(\mathbb{R})$ *its (weighted) empirical barycenter satisfies:*

$$\overline{\mu} := \operatorname*{arg\,min}_{\nu \in W_2} \sum_{i=1}^n \omega_i \cdot d_{W_2}(\nu, \mu_i)^2 \quad \to \quad Q_{\overline{\mu}} = \sum_{i=1}^n \omega_i \cdot Q_{\mu_i}.$$

In particular, $\mathbb{E}_{\overline{\mu}} X = \sum \omega_i \cdot \mathbb{E}_{\mu_i} X$.

Proof. We adapt the proof of Proposition 3.1 in [4]. Apply Lemma A.2 and the bias-variance decomposition to see $Q_{\overline{\mu}}$ must minimize:

$$\sum_{i=1}^n \omega_i \cdot ||Q_{\mu_i} - Q_{\overline{\mu}}||_{L^2}^2 = \sum_{i=1}^n \omega_i \cdot \left|\left| \sum_{i=1}^n \omega_i \cdot Q_{\mu_i} - Q_{\overline{\mu}} \right|\right|_{L^2}^2 + \left|\left| \sum_{i=1}^n \omega_i \cdot Q_{\mu_i} - Q_{\mu_i} \right|\right|_{L^2}^2.$$

For the latter result, just use $\mathbb{E}_\mu X = \int_0^1 Q_\mu(z) \ dz$ and linearity of expectation. □

A.2 Main Results

Proof of Theorems 2.1 ,4.1. By Lemma A.2 it suffices to work with quantile functions in ϕ-space. If $\mu = \sum_i p_i \delta_{x_i}$ is discrete, its quantile function is an increasing step function. Namely, assuming $x_1 < x_2 < ... < x_n$ it takes the form:

$$Q_\mu(z) = x_k \text{ for } z \in \left(\sum_{i=1}^{k-1} p_i, \quad \sum_{i=1}^{k} p_i\right].$$

Thus if $\mu = \sum_i p_i \delta_{x_i}$ and $\nu = \sum_i q_i \delta_{y_i}$ then:

$$|Q_\mu(z) - Q_\nu(z)| = |x_k - y_r| \text{ for } z \in \left(\sum_{i=1}^{k-1} p_i, \quad \sum_{i=1}^{k} p_i\right] \cap \left(\sum_{i=1}^{r-1} q_i, \quad \sum_{i=1}^{r} q_i\right],$$

is piecewise constant, with pieces delineated by endpoints $\sum_{i \leq k} p_i, \sum_{j \leq k} q_j$. Namely, if we let $z_1 \leq z_2 \leq ... \leq z_{m+n}$ be the order statistics of $\{\sum_{i=1}^{k} p_i \ : \ k \leq n\} \cup \{\sum_{i=1}^{r} q_i \ : \ r \leq m\}$ we can see:

$$d_{W_p}(\mu, \nu)^p = \sum_{\ell=1}^{m+n} \Delta z_\ell \cdot |x_{f(\ell)} - y_{g(\ell)}|^p, \qquad \Delta z_\ell := z_\ell - z_{\ell-1},$$

where $f : [m+n] \to [n]$ is the map that sends ℓ to k if $z_\ell \in \left[\sum_{i=1}^{k-1} p_i, \sum_{i=1}^{k} p_i\right)$. Likewise $g : [m+n] \to [m]$ sends ℓ to k if $z_\ell \in \left[\sum_{i=1}^{k-1} q_i, \sum_{i=1}^{k} q_i\right)$.

Thus, taking μ to be the transformed bid distribution and ν the transformed ask distribution, (4.2) gives:

$$d_{W_p}\left((c - 2\theta\,\cdot)_\# \mu,\ (2(1-\theta)\ \cdot -c)_\# \nu\right)^p = \sum_{\ell=1}^{m+n} \Delta z_\ell \cdot \left|2c - 2\theta \cdot x_{f(\ell)} - 2(1-\theta) \cdot y_{g(\ell)}\right|^p, \tag{A.1}$$

where $x_1 < ... < x_n$ and $y_1 > ... > y_n$. Note the reversal in the order of the y_i (and hence their corresponding q_i). Note that the z_ℓ depend only on p_i, q_i, not on c. In the case of W_2, by taking the gradient we can explicitly solve for the optimal c in (1.3):

$$\begin{aligned} x^\theta_{s,\phi} &= \sum_{\ell=1}^{m+n} \Delta z_\ell \cdot \left(\theta x_{f(\ell)} + (1-\theta) y_{g(\ell)}\right) \\ &= \theta \sum_{k=1}^{n} p_k x_k + (1-\theta) \sum_{k=1}^{m} q_k y_k \\ &= \theta \cdot m_2(\mu) + (1-\theta) \cdot m_2(\nu). \end{aligned}$$

where the second line follows by grouping terms where $f(\ell) = k$ and terms where $g(\ell) = k$ respectively. Taking $x_i = \widetilde{b}_i$ and $y_i = \widetilde{a}_i$, with masses determined by (1.2), yields the formula for $x^\theta_{s,\phi}$ in Theorem 4.1; applying ϕ^{-1} gives the stated price-level fair value. Theorem 2.1 is just the special "balanced" case when $\theta = \frac{1}{2}$. $\square$

We can re-state the above proof purely in terms of the quantile functions. Note if X has quantile function Q_X, then the transformed random variables $Y = c - 2\theta X$ and $Y' = 2(1-\theta)X - c$ have quantile functions (up to Lebesgue null sets) given by:

$$Q_Y(u) = c - 2\theta Q_X(1-u), \quad Q_{Y'}(u) = 2(1-\theta) Q_X(u) - c,$$

respectively. Using Lemma A.2, then (4.2) becomes:

$$d_{W_p}\left((c-2\theta\,\cdot)_{\#}\mu,\ (2(1-\theta)\ \cdot -c)_{\#}\nu\right)^p = 2^p \int_0^1 |c - \theta Q_\mu(1-u) - (1-\theta)Q_\nu(u)|^p \ du.$$

Defining:

$$\rho(u) := \theta Q_\mu(1-u) + (1-\theta)Q_\nu(u),$$

up to the factor 2^p, the objective is $\int_0^1 |c-\rho(u)|^p\, du$. Hence we can view $x^\theta_{s,\phi}$ in (4.2) as the p^{th} mean of the random variable $Z = \rho(U)$ for $U \sim \mathrm{Unif}[0,1]$. This naturally extends beyond discrete distributions as well.

In the special case of $p = 2$, the p^{th} mean is just the standard mean. By linearity

$$\mathbb{E}Z = \theta\mathbb{E}Q_\mu(1-U) + (1-\theta)\mathbb{E}Q_\nu(U) = \theta\mathbb{E}Q_\mu(U) + (1-\theta)\mathbb{E}Q_\nu(U) = \mathbb{E}Q_{\bar{\mu}_\theta}(U),$$

where $\bar{\mu}_\theta$ is the W_2 barycenter of μ, ν with weights $\theta, 1-\theta$, as discussed in Lemma A.3. Thus, in the case of $p = 2$, our transformed symmetrized fair value is just the mean of the (imbalance-weighted) W_2 barycenter of μ, ν. Namely

Corollary A.4. *For any transformed price measures $\mu, \nu \in W_2(\mathbb{R})$:*

$$x^\theta_{s,\phi} = \operatorname*{arg\,min}_{c\in\mathbb{R}} d_{W_2}\left((c-2\theta\,\cdot)_{\#}\mu,\ (2(1-\theta)\ \cdot -c)_{\#}\nu\right) = \int_{\mathbb{R}} x\,\bar{\mu}_\theta(dx), \qquad v^\theta_{s,\phi} = \phi^{-1}(x^\theta_{s,\phi}).$$

For $p \neq 2$, we lose the linearity of the p^{th} mean. However, we can still interpret ρ as a measure of the fair value at different "spreads/levels" u, and Z encodes the distribution these fair values can take on. The p^{th} mean then just aims to quantify the typical fair value. For example, when μ, ν are mirror images, then ρ is just some constant c, so every level u agrees on the fair value. In this case, the p^{th} mean is c for every $p > 1$.

Proof of Lemma 3.1. The proof is analogous to Lemma 4.1 of [1], which shows log-concavity is the crucial condition. Call the sum to be maximized $S(\tau)$ and the summand $V(\widetilde{b}_i, \widetilde{a}_{\tau(i)})$. For brevity within this proof, write $b_i := \widetilde{b}_i$ and $a_j := \widetilde{a}_j$. If $i < j$ but $\tau(i) =: u > v := \tau(j)$ we will show $S(\tau') \geq S(\tau)$ for $\tau' = (uv)\tau$. Namely, by correcting the order of i, j in the image we can improve the objective value. Note since μ is normally distributed:

$$g(x) = -\frac{x^2}{2\sigma^2} \qquad \rightarrow \qquad \Delta_r\Delta_t g(x) = -\frac{rt}{\sigma^2}.$$

Recall that $b_j \leq b_i \leq a_v \leq a_u$. Define $r := b_i - b_j \geq 0$, $z := a_v - b_i \geq 0$, and $t := a_u - a_v \geq 0$ to be the spacing between these four values and let $w := \theta b_j + (1-\theta)a_v$ to be the weighted average of the smallest ask and bid in this group. Then we see:

$$\begin{aligned}
&S(\tau') - S(\tau) \\
&= [V(b_i, a_v) + V(b_j, a_u)] - [V(b_i, a_u) + V(b_j, a_v)] \\
&= -\Big[\Big(g(\theta b_i + (1-\theta)a_u) - g(\theta b_i + (1-\theta)a_v)\Big) \\
&\qquad\qquad - \Big(g(\theta b_j + (1-\theta)a_u) - g(\theta b_j + (1-\theta)a_v)\Big)\Big]
\end{aligned}$$

$$
\begin{aligned}
&+ \Big[\Big(f(\theta(1-\theta)(a_u - b_j)) - f(\theta(1-\theta)(a_u - b_i))\Big) \\
&\qquad\qquad - \Big(f(\theta(1-\theta)(a_v - b_j)) - f(\theta(1-\theta)(a_v - b_i))\Big)\Big] \\
&= -\Delta_{\theta r}\Delta_{(1-\theta)t} g(w) + \Delta_{\theta(1-\theta)t}\,\Delta_{\theta(1-\theta)r}\, f(\theta(1-\theta)z) \\
&= \frac{\theta(1-\theta)}{\sigma^2}\cdot rt + \Delta_{\theta(1-\theta)t}\,\Delta_{\theta(1-\theta)r}\, f(\theta(1-\theta)z).
\end{aligned}
$$

Substituting $t' = \theta(1-\theta)t$ and likewise for the other variables we get:

$$
S(\tau') - S(\tau) = \frac{1}{\theta(1-\theta)\sigma^2}\cdot r't' + \Delta_{t'}\Delta_{r'} f(z').
$$

By assumption, this is nonnegative. Because we can go from any permutation to the identity through a series of transpositions, the result follows. □

Proof of Lemma 5.1. The first inequality follows almost immediately from (4.4) and Lemma A.1. Just note

$$
\begin{aligned}
&\left|x^{\theta}_{s,\phi}(\mu,\nu) - x^{\theta'}_{s,\phi}(\mu',\nu')\right| \\
&= \left|\theta m_2(\mu) + (1-\theta)m_2(\nu) - \theta' m_2(\mu') - (1-\theta')m_2(\nu')\right| \\
&= \left|\theta'\big(m_2(\mu) - m_2(\mu')\big) + (1-\theta')\big(m_2(\nu) - m_2(\nu')\big) + (\theta-\theta')\big(m_2(\mu) - m_2(\nu)\big)\right| \\
&\le \theta' \left|m_2(\mu) - m_2(\mu')\right| + (1-\theta')\left|m_2(\nu) - m_2(\nu')\right| + |\theta-\theta'|\,|m_2(\mu) - m_2(\nu)| \\
&\le \theta' d_{W_p}(\mu,\mu') + (1-\theta')d_{W_p}(\nu,\nu') + |\theta-\theta'| d_{W_p}(\mu,\nu).
\end{aligned}
$$

Using the triangle inequality on the last term gives the desired result. For the second, note by Jensen's inequality

$$
|m_2(\mu) - m_2(\mu')| = \left|\int_{\mathbb{R}} (x - t(x))\,\mu(dx)\right| \le ||\mathrm{id} - t||_{L^p(\mu)}.
$$

so the above work gives the desired result again. □

Now, we prove the more general stability result under d_{W_p} for $p \neq 2$. The main idea is that the optimization problem in (4.3) is strictly convex in c for $p > 1$ and depends smoothly on the parameters.

Proof of Lemma 5.2. For $u \in [0,1]$, define

$$
\rho_n(u) := \theta_n Q_{\mu_n}(1-u) + (1-\theta_n)Q_{\nu_n}(u), \quad \rho(u) := \theta Q_{\mu}(1-u) + (1-\theta)Q_{\nu}(u).
$$

and the functions

$$
f_n(c) := \int_0^1 |c - \rho_n(u)|^p\, du, \quad f(c) := \int_0^1 |c - \rho(u)|^p\, du.
$$

By Lemma A.2 and the definition of $x^{\theta}_{s,\phi}$ in (4.3), we have

$$
x^{\theta_n}_{s,\phi}(\mu_n,\nu_n) = \arg\min_{c\in\mathbb{R}} f_n(c), \quad x^{\theta}_{s,\phi}(\mu,\nu) = \arg\min_{c\in\mathbb{R}} f(c).
$$

We start by showing $\rho_n \to \rho$ in $L_p([0,1])$. By Lemma A.2

$$\begin{aligned}
||\rho_n - \rho||_{L_p} &\leq ||\theta_n Q_{\mu_n}(1-\cdot) - \theta Q_\mu(1-\cdot)||_{L_p} + ||(1-\theta_n)Q_{\nu_n} - (1-\theta)Q_\nu||_{L_p} \\
&\leq |\theta_n - \theta| \cdot ||Q_\mu||_{L_p} + \theta_n ||Q_{\mu_n} - Q_\mu||_{L_p} + |\theta_n - \theta| \cdot ||Q_\nu||_{L_p} + (1-\theta_n)||Q_{\nu_n} - Q_\nu||_{L_p} \\
&= |\theta_n - \theta| \cdot ||Q_\mu||_{L_p} + \theta_n ||\mu_n - \mu||_{W_p} + |\theta_n - \theta| \cdot ||Q_\nu||_{L_p} + (1-\theta_n)||\nu_n - \nu||_{W_p}.
\end{aligned}$$

As $\mu_n \xrightarrow{W_p} \mu, \nu_n \xrightarrow{W_p} \nu$ and $\theta_n \to \theta$, this vanishes, so $\rho_n \to \rho$ in L_p. Then, using the inequality

$$||a|^p - |b|^p| \leq p|a-b| \cdot (|a|^{p-1} + |b|^{p-1}),$$

we see:

$$\begin{aligned}
\Big| |\rho_n(u) - c|^p - |\rho(u) - c|^p \Big| &\leq p \cdot |\rho_n(u) - \rho(u)| \cdot \left[(|\rho_n(u)| + c)^{p-1} + (|\rho(u)| + c)^{p-1}\right] \\
&\leq C_p \cdot |\rho_n(u) - \rho(u)| \cdot \left[c + |\rho_n(u)|^{p-1} + |\rho(u)|^{p-1}\right].
\end{aligned}$$

Hence for $R > 0$

$$\sup_{|c| \leq R} |f_n(c) - f(c)| \leq C_p \cdot ||\rho_n - \rho||_{L_1} \cdot \left[R + ||\rho_n||_{L^{p-1}}^{p-1} + ||\rho||_{L^{p-1}}^{p-1}\right].$$

Since $\rho_n \to \rho$ in L_p, for $p > 1$, we have $\sup_n ||\rho_n||_{L^{p-1}} < \infty$ and $||\rho_n - \rho||_{L_1} \to 0$, so the right hand side goes to zero. Hence $f_n \to f$ uniformly on compact sets.

Note that f_n, f are strictly convex for $p > 1$, and they achieve their minimum at a unique point. This and locally uniform convergence then imply these minimizers converge, so $x_{s,\phi}^{\theta_n}(\mu_n, \nu_n) \to x_{s,\phi}^{\theta}(\mu, \nu)$. Continuity of ϕ^{-1} gives $v_{s,\phi}^{\theta_n}(\mu_n, \nu_n) \to v_{s,\phi}^{\theta}(\mu, \nu)$. □

References

[1] Sergey G. Bobkov and Michel Ledoux. "One-Dimensional Measures, Order Statistics, and Kantorovich Transport Distances". In: *Memoirs of the American Mathematical Society* 201.945 (2009).

[2] Konark Jain et al. *Limit Order Book Simulations: A Review*. 2024. arXiv: 2402.17359 [q-fin.TR]. URL: https://arxiv.org/abs/2402.17359.

[3] Victor Panaretos and Yoav Zemel. *An Invitation to Statistics in Wasserstein Space*. Springer, 2020. ISBN: 978-3-030-38437-1. DOI: 10.1007/978-3-030-38438-8.

[4] Daniel Raban. *How Should We Aggregate Ratings? Accounting for Personal Rating Scales via Wasserstein Barycenters*. 2024. arXiv: 2410.00865 [math.ST]. URL: https://arxiv.org/abs/2410.00865.