# Optimal Rebalancing under Fixed and Variable Trading Costs

Yannick Kälber · 21 Oct 2026

https://functor.network/user/3370/entry/2022

Whenever current portfolio weights differ from their target weights, either due to market movements, signal changes or other reasons, a portfolio manager in a frictionless world would immediately trade to the target portfolio. In reality, however, every adjustment incurs trading costs. Therefore a portfolio manager faces an optimization problem, where he needs to balance portfolio suboptimality and trading costs.

In this article, I derive an analytical framework for this trade-off. Starting from a simple asset-by-asset baseline, I show why variable costs prevent us from ever trading all the way to target weights and how fixed costs split the rebalancing rule into two distinct bands. Extending the model to a full return covariance matrix then reveals two effects that an isolated asset-by-asset view overlooks. Weight deviations across correlated assets can offset each other at the portfolio level, and it can be preferable to leave expensive assets untouched while trading cheaper proxies, even when doing so pushes those proxies further away from their own target weights.

# Cost Function

Say we have $n \in \mathbb{N}_{>0}$ assets. Let $\mathbf{w}^{\text{cur}} \in \mathbb{R}^n$ denote their current portfolio weights, $\mathbf{w}^{\text{tar}} \in \mathbb{R}^n$ the target weights, and $\mathbf{w}^{\text{new}} \in \mathbb{R}^n$ the new weights we choose to rebalance to. At each rebalancing date, we select $\mathbf{w}^{\text{new}}$ to minimize the total cost

$$C(\mathbf{w}^{\text{new}}) = k\mathbf{1}_{\{\mathbf{w}^{\text{new}} \neq \mathbf{w}^{\text{cur}}\}} + \mathbf{c}^\top |\mathbf{w}^{\text{new}} - \mathbf{w}^{\text{cur}}| + \delta(\mathbf{w}^{\text{new}} - \mathbf{w}^{\text{tar}})^\top \mathbf{Q}(\mathbf{w}^{\text{new}} - \mathbf{w}^{\text{ta} \tag{1}$$

$k \in \mathbb{R}_{\geq 0}$ is a fixed transaction cost (such as ticket fees or operational overhead) incurred as a lump sum whenever any trade occurs, regardless of trade size. The indicator function $\mathbf{1}_{\{\mathbf{w}^{\text{new}} \neq \mathbf{w}^{\text{cur}}\}}$ takes the value $1$ as soon as at least one portfolio weight is changed, and $0$ if the portfolio remains untouched at $\mathbf{w}^{\text{cur}}$. The vector

$\mathbf{c} = (c_1, \ldots, c_n)^\top \in \mathbb{R}^n_{>0}$ contains the asset-specific variable transaction costs (such as broker commissions), where the absolute value operator in $|\mathbf{w}^{\text{new}} - \mathbf{w}^{\text{cur}}|$ acts element-wise on the trade vector because both purchases and sales incur positive costs. The scalar $\delta \in \mathbb{R}_{>0}$ is the suboptimality-aversion parameter, translating units of quadratic deviation into units of transaction cost, and $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is a symmetric positive definite penalty matrix ($\mathbf{Q} \succ 0$) that weights the quadratic deviation from the target portfolio.

If $\mathbf{Q} = \mathbf{I}$, the quadratic term in (1) penalizes raw weight deviations equally across assets, reducing the problem to $n$ independent scalar optimization problems. If $\mathbf{Q} = \mathbf{\Sigma}$, where $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ is the positive definite return covariance matrix, the quadratic term equals the tracking-error variance $\text{Var}(r_p - r_{\text{tar}}) = (\mathbf{w}^{\text{new}} - \mathbf{w}^{\text{tar}})^\top \mathbf{\Sigma} (\mathbf{w}^{\text{new}} - \mathbf{w}^{\text{tar}})$, coupling the assets through their correlations. The latter is economically more natural because portfolio return deviations are what ultimately matter, not weight deviations.

To simplify the subsequent algebra, we shift the coordinate system so that the target portfolio $\mathbf{w}^{\text{tar}}$ lies at the origin. To do so, we define the post-trade deviation from the target as $\mathbf{x} = \mathbf{w}^{\text{new}} - \mathbf{w}^{\text{tar}} \in \mathbb{R}^n$ and the current deviation as $\mathbf{d} = \mathbf{w}^{\text{cur}} - \mathbf{w}^{\text{tar}} \in \mathbb{R}^n$. The trade vector $\mathbf{\Delta} = \mathbf{w}^{\text{new}} - \mathbf{w}^{\text{cur}}$ can then be expressed as the difference between these two deviations.

$$
\begin{aligned}
\mathbf{\Delta} &= \mathbf{w}^{\text{new}} - \mathbf{w}^{\text{cur}} && (2)\\
&= (\mathbf{w}^{\text{new}} - \mathbf{w}^{\text{tar}}) - (\mathbf{w}^{\text{cur}} - \mathbf{w}^{\text{tar}}) && (3)\\
&= \mathbf{x} - \mathbf{d} && (4)
\end{aligned}
$$

In (2) we start from the definition of the trade vector, and in (3) we subtract and add the target weight vector $\mathbf{w}^{\text{tar}}$. In (4) we substitute the definitions $\mathbf{x} = \mathbf{w}^{\text{new}} - \mathbf{w}^{\text{tar}}$ and $\mathbf{d} = \mathbf{w}^{\text{cur}} - \mathbf{w}^{\text{tar}}$. Equation (4) simply states that the trade executed in each asset is the difference between the desired post-trade deviation $\mathbf{x}$ and the pre-trade deviation $\mathbf{d}$. Since subtracting $\mathbf{w}^{\text{tar}}$ from both sides also makes the condition $\mathbf{w}^{\text{new}} \neq \mathbf{w}^{\text{cur}}$ equivalent to $\mathbf{x} \neq \mathbf{d}$, substituting (4) into (1) gives the total cost function in terms of $\mathbf{x}$.

$$
C(\mathbf{x}) = k\mathbf{1}_{\{\mathbf{x} \neq \mathbf{d}\}} + \mathbf{c}^\top |\mathbf{x} - \mathbf{d}| + \delta \mathbf{x}^\top \mathbf{Q} \mathbf{x} \tag{5}
$$

If we refrain from trading altogether by choosing $\mathbf{x} = \mathbf{d}$ (which corresponds to $\mathbf{w}^{\text{new}} = \mathbf{w}^{\text{cur}}$ and $\mathbf{\Delta} = \mathbf{0}$), the first two terms in (5) vanish.

$$
\begin{aligned}
C(\mathbf{d}) &= k\mathbf{1}_{\{\mathbf{d}\neq\mathbf{d}\}} + \mathbf{c}^\top|\mathbf{d}-\mathbf{d}| + \delta\mathbf{d}^\top\mathbf{Q}\mathbf{d} && (6)\\
&= k\cdot 0 + \mathbf{c}^\top\mathbf{0} + \delta\mathbf{d}^\top\mathbf{Q}\mathbf{d} && (7)\\
&= \delta\mathbf{d}^\top\mathbf{Q}\mathbf{d} && (8)
\end{aligned}
$$

In (6) we evaluate (5) at the current deviation $\mathbf{x}=\mathbf{d}$. In (7) we use the fact that the statement $\mathbf{d}\neq\mathbf{d}$ is false and the trade vector equals the zero vector, leaving only the quadratic penalty of the current deviation in (8). $C(\mathbf{d}) = \delta\mathbf{d}^\top\mathbf{Q}\mathbf{d}$ is the status-quo suboptimality.

## Two-Stage Problem Separation

At first glance, the jump discontinuity at $\mathbf{x}=\mathbf{d}$ in (5) complicates the analytical minimization, as the function is not continuous everywhere and therefore not globally differentiable. For every point $\mathbf{x}\neq\mathbf{d}$, however, the indicator term $k\mathbf{1}_{\{\mathbf{x}\neq\mathbf{d}\}}$ is simply an additive constant $k$. Therefore, the fixed costs affect only the extensive margin (whether to trade at all), not the intensive margin (how much to trade once a trade is initiated). Mathematically, shifting the cost surface vertically upward by a constant $k$ leaves its gradient and thus the location of its minimum unchanged. We can therefore isolate the continuous part of the cost function from (5) as $\tilde{C}(\mathbf{x}) = \mathbf{c}^\top|\mathbf{x}-\mathbf{d}| + \delta\mathbf{x}^\top\mathbf{Q}\mathbf{x}$ and determine its unique global minimum $\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathbb{R}^n}\tilde{C}(\mathbf{x})$.

Moving from the status quo deviation $\mathbf{d}$ to this optimal deviation $\mathbf{x}^*\neq\mathbf{d}$ strictly lowers total cost (5) below the status-quo holding cost (8) if and only if the following condition holds.

$$
\begin{aligned}
C(\mathbf{x}^*) &< C(\mathbf{d}) && (9)\\
k + \tilde{C}(\mathbf{x}^*) &< \delta\mathbf{d}^\top\mathbf{Q}\mathbf{d} && (10)\\
\delta\mathbf{d}^\top\mathbf{Q}\mathbf{d} - \tilde{C}(\mathbf{x}^*) &> k && (11)\\
B(\mathbf{d}) &> k && (12)
\end{aligned}
$$

In (9) we state the economic rationality condition that the total post-trade cost (including the fixed fee) must be strictly smaller than the cost of doing nothing. In (10) we substitute (5) for $\mathbf{x}^*\neq\mathbf{d}$ on the left-hand side and the holding cost (8) on the right-hand side. In (11) we subtract the continuous post-trade cost $\tilde{C}(\mathbf{x}^*)$ from both sides and read the inequality from right to left. In (12) we define the gross cost savings of the optimal rebalancing trade as $B(\mathbf{d}) = \delta\mathbf{d}^\top\mathbf{Q}\mathbf{d} - \tilde{C}(\mathbf{x}^*)$, which represents the cost reduction after paying variable transaction costs $\mathbf{c}^\top|\mathbf{x}^*-\mathbf{d}|$, but before paying the fixed fee $k$. Thus,

the rebalancing problem decouples into first computing the optimal target $\mathbf{x}^*$ under variable costs alone, and second checking whether the resulting gross savings $B(\mathbf{d})$ clear the fixed-cost hurdle $k$ in (12).

# Diagonal $\mathbf{Q}$ Case

## Stage 1

Let us first solve this two-stage problem for a diagonal positive definite penalty matrix $\mathbf{Q} = \operatorname{diag}(q_1, \ldots, q_n)$ with diagonal entries $q_i > 0$. This specification covers both the identity matrix $\mathbf{Q} = \mathbf{I}$ ($q_i = 1$) and standalone variance-weighted penalties ($q_i = \sigma_i^2$) where cross-asset correlations are ignored. Substituting the diagonal structure of $\mathbf{Q}$ and the element-wise definition of the absolute value vector into the continuous cost function $\tilde{C}(\mathbf{x})$ gives the following expression.

$$\begin{aligned} \tilde{C}(\mathbf{x}) &= \mathbf{c}^\top |\mathbf{x} - \mathbf{d}| + \delta \mathbf{x}^\top \operatorname{diag}(q_1, \ldots, q_n)\mathbf{x} && (13) \\ &= \sum_{i=1}^{n} c_i |x_i - d_i| + \delta \sum_{i=1}^{n} q_i x_i^2 && (14) \\ &= \sum_{i=1}^{n} \left( c_i |x_i - d_i| + \delta q_i x_i^2 \right) && (15) \end{aligned}$$

In (13) we write $\tilde{C}(\mathbf{x})$ with the explicit diagonal matrix. In (14) we expand the inner product $\mathbf{c}^\top |\mathbf{x} - \mathbf{d}|$ and the quadratic form into sums over the individual assets $i \in \{1, \ldots, n\}$, since all off-diagonal entries of $\mathbf{Q}$ are zero. In (15) we merge both sums into a single summation. Because each summand $\tilde{C}_i(x_i) = c_i |x_i - d_i| + \delta q_i x_i^2$ depends exclusively on the scalar variable $x_i$ of asset $i$ without any cross-terms $x_i x_j$, minimizing the portfolio cost (15) is equivalent to minimizing $\tilde{C}_i(x_i)$ for each asset independently.

### No-Trade Threshold

To minimize $\tilde{C}_i(x_i)$ with respect to $x_i$, we derive the first-order optimality condition. However, the variable transaction cost term $c_i |x_i - d_i|$ has a sharp V-shaped kink at the status-quo point $x_i = d_i$ and is not differentiable there in the classical sense. If we buy asset $i$ ($x_i > d_i$), costs increase at rate $+c_i$; if we sell asset $i$ ($x_i < d_i$), costs also increase as $x_i$ moves down, which means the slope with respect to $x_i$ is $-c_i$. Right at the point of inaction ($x_i = d_i$), the slope jumps instantaneously from $-c_i$ to $+c_i$. We handle such kinks using the subdifferential $\partial |x_i - d_i|$. Geometrically, the subdifferential consists of the

slopes of all supporting lines that touch the graph at the given point and lie entirely below it. Away from the kink ($x_i \neq d_i$), the graph is smooth, so there is only one tangent slope, namely the sign $\operatorname{sgn}(x_i - d_i) \in \{-1, +1\}$. At the bottom of the V-kink ($x_i = d_i$), however, any line with a slope $\xi_i$ between $-1$ and $+1$ stays below the V, so the subdifferential is the entire closed interval $[-1, 1]$. A convex function attains its global minimum at $x_i^*$ if and only if a flat horizontal tangent (slope $0$) is contained in the sum of the subdifferential of the cost term and the derivative $2\delta q_i x_i^*$ of the smooth suboptimality term, which we then evaluate at the status-quo point $x_i^* = d_i$ to determine when inaction is optimal.

$$0 \in c_i \partial |x_i^* - d_i| + 2\delta q_i x_i^* \tag{16}$$
$$-2\delta q_i x_i^* \in c_i \partial |x_i^* - d_i| \tag{17}$$
$$-2\delta q_i d_i \in [-c_i, c_i] \tag{18}$$
$$|2\delta q_i d_i| \leq c_i \tag{19}$$
$$|d_i| \leq \frac{c_i}{2\delta q_i} = \tau_i \tag{20}$$

In (16) we state the first-order optimality condition by adding the subdifferential of the variable cost term and the classical derivative $2\delta q_i x_i^*$ of the quadratic suboptimality term. In (17) we move the marginal suboptimality term to the left-hand side. To check under what condition it is optimal not to trade asset $i$ at all ($x_i^* = d_i$), in (18) we substitute $x_i^* = d_i$ and $\partial|0| = [-1, 1]$ into (17). In (19) we rewrite the interval inclusion as an absolute value inequality, and in (20) we divide by $2\delta q_i > 0$ to define the asset-specific no-trade threshold $\tau_i = \frac{c_i}{2\delta q_i}$. Both intermediate equation (19) and final threshold (20) have a neat economic interpretation, since $|2\delta q_i d_i|$ is the marginal suboptimality reduction of trading the very first dollar of asset $i$ back toward target, while $c_i$ is the proportional transaction cost per dollar traded. As long as marginal benefit $|2\delta q_i d_i|$ does not exceed marginal cost $c_i$, the best action is inaction. Consequently, the width of the no-trade band $\tau_i$ in (20) is directly proportional to the asset's variable cost $c_i$ and inversely proportional to both the suboptimality-aversion parameter $\delta$ and the asset's standalone penalty weight $q_i$.

## Optimal Active Trade

If the current deviation breaches this threshold, so that $|d_i| > \tau_i$, $-2\delta q_i d_i$ falls outside $[-c_i, c_i]$, and holding $x_i^* = d_i$ violates (17). Because the quadratic suboptimality penalty $\delta q_i x_i^2$ pulls the portfolio toward the origin $0$, the optimal trade moves $x_i^*$ from $d_i$ toward $0$. If the asset is overweight ($d_i > \tau_i > 0$), we sell ($x_i^* < d_i$), and if it is underweight ($d_i < -\tau_i < 0$), we buy ($x_i^* > d_i$). In either case, the trade direction $x_i^* - d_i$ has the opposite sign of the current

deviation $d_i$, so $\operatorname{sgn}(x_i^* - d_i) = -\operatorname{sgn}(d_i)$. Because $x_i^* \neq d_i$, we are on the smooth flank of the V-function where the subdifferential contains only the single slope $\{\operatorname{sgn}(x_i^* - d_i)\}$, and (16) reduces to an ordinary first-order condition.

$$
\begin{aligned}
0 &= c_i \operatorname{sgn}(x_i^* - d_i) + 2\delta q_i x_i^* && (21)\\
0 &= -c_i \operatorname{sgn}(d_i) + 2\delta q_i x_i^* && (22)\\
2\delta q_i x_i^* &= c_i \operatorname{sgn}(d_i) && (23)\\
x_i^* &= \operatorname{sgn}(d_i) \frac{c_i}{2\delta q_i} = \operatorname{sgn}(d_i)\tau_i && (24)
\end{aligned}
$$

In (21) we replace the subdifferential in (16) for an active trade $x_i^* \neq d_i$ with the sign function. In (22) we substitute the trade direction $\operatorname{sgn}(x_i^* - d_i) = -\operatorname{sgn}(d_i)$. In (23) we move the marginal transaction cost $c_i \operatorname{sgn}(d_i)$ to the right-hand side, which shows that at the optimal stopping point $x_i^*$, the remaining marginal suboptimality benefit $2\delta q_i |x_i^*|$ must equal the proportional trading cost $c_i$. In (24) we divide by $2\delta q_i$ and insert $\tau_i$ from (20). Because $|x_i^*| = \tau_i < |d_i|$, $x_i^*$ stays on the same side of zero as $d_i$ without overshooting the origin, confirming $\operatorname{sgn}(x_i^* - d_i) = -\operatorname{sgn}(d_i)$. Combining the inactive regime $x_i^* = d_i = \operatorname{sgn}(d_i)|d_i|$ for $|d_i| \leq \tau_i$ from (20) and the active regime $x_i^* = \operatorname{sgn}(d_i)\tau_i$ for $|d_i| > \tau_i$ from (24), we obtain closed-form expressions for the post-trade deviation $x_i^*$ and the optimal trade size $\Delta_i^* = x_i^* - d_i$.

$$
\begin{aligned}
x_i^* &= \operatorname{sgn}(d_i) \min(|d_i|, \tau_i) && (25)\\
\Delta_i^* &= \operatorname{sgn}(d_i) \min(|d_i|, \tau_i) - \operatorname{sgn}(d_i)|d_i| && (26)\\
&= -\operatorname{sgn}(d_i) \left(|d_i| - \min(|d_i|, \tau_i)\right) && (27)\\
&= -\operatorname{sgn}(d_i) \max(|d_i| - \tau_i, 0) = -\mathcal{S}_{\tau_i}(d_i) && (28)
\end{aligned}
$$

In (25) we unify (20) and (24) via the minimum of the current deviation magnitude and the threshold $\tau_i$. In (26) we subtract $d_i = \operatorname{sgn}(d_i)|d_i|$ to form the trade $\Delta_i^* = x_i^* - d_i$. In (27) we factor out $-\operatorname{sgn}(d_i)$, and in (28) we rewrite $|d_i| - \min(|d_i|, \tau_i)$ as $\max(|d_i| - \tau_i, 0)$, which is the negative of the soft-thresholding operator $\mathcal{S}_{\tau_i}(d_i)$ familiar from Lasso regression.

Equations (24) and (28) reveal a first important principle of trading under proportional costs, namely that **even when the current deviation is large enough to trigger a trade, you never trade all the way back to the target weight ($x_i^* = 0$).** Instead, you only shave off the excess deviation $|d_i| - \tau_i$ outside the band and stop right on the boundary $\pm\tau_i$. Mathematically, this is because quadratic suboptimality has diminishing marginal costs near zero

($2\delta q_i|x_i| \to 0$ as $x_i \to 0$), whereas variable transaction costs $c_i$ are linear and constant. Once you reach the edge $\pm\tau_i$, every additional basis point of trading costs $c_i$ but saves less than $c_i$ in suboptimality.

## Stage 2

Having solved Stage 1, we now evaluate Stage 2 (the fixed-cost trigger (12)) by computing the gross cost savings $B_i(d_i) = \delta q_i d_i^2 - \tilde{C}_i(x_i^*)$ of each asset. For untraded assets ($|d_i| \le \tau_i$), we have $x_i^* = d_i$ and zero savings. For actively traded assets ($|d_i| > \tau_i$), equations (24) and (28) imply that the post-trade residual suboptimality is $\delta q_i (x_i^*)^2 = \delta q_i \tau_i^2$ and the traded notional is $|x_i^* - d_i| = |d_i| - \tau_i$. Substituting these into the cost difference gives the gross savings for asset $i$.

$$
\begin{aligned}
B_i(d_i) &= \delta q_i d_i^2 - \left( c_i |x_i^* - d_i| + \delta q_i (x_i^*)^2 \right) && (29)\\
&= \delta q_i d_i^2 - c_i(|d_i| - \tau_i) - \delta q_i \tau_i^2 && (30)\\
&= \delta q_i d_i^2 - 2\delta q_i \tau_i (|d_i| - \tau_i) - \delta q_i \tau_i^2 && (31)\\
&= \delta q_i \left( |d_i|^2 - 2\tau_i |d_i| + 2\tau_i^2 - \tau_i^2 \right) && (32)\\
&= \delta q_i \left( |d_i| - \tau_i \right)^2 = \delta q_i (\Delta_i^*)^2 && (33)
\end{aligned}
$$

In (29) we subtract the continuous post-trade cost (variable transaction cost plus residual suboptimality penalty) from the initial holding cost $\delta q_i d_i^2$. In (30) we substitute the active destination $x_i^* = \operatorname{sgn}(d_i)\tau_i$; notice economically that the raw risk reduction $\delta q_i (d_i^2 - \tau_i^2)$ is diminished by the variable costs payment $c_i(|d_i| - \tau_i)$. In (31) we eliminate $c_i$ using the marginal-cost identity $c_i = 2\delta q_i \tau_i$ from (20). In (32) we factor out $\delta q_i$ and use $d_i^2 = |d_i|^2$, allowing us to compress the bracketed expression in (33) via the second binomial formula into the square of the optimal trade size $\Delta_i^*$ from (28). Equation (33) shows that after paying the variable costs $c_i$, the net savings $B_i(d_i)$ do not depend on the total deviation $d_i$, but only on the squared excess deviation $(|d_i| - \tau_i)^2$ that protrudes beyond the no-trade threshold. Summing (33) across all $n$ assets and inserting the sum into (12) yields the closed-form global decision rule for diagonal $\mathbf{Q}$.

$$
\mathbf{w}^{\text{new}*} = \begin{cases} \mathbf{w}^{\text{tar}} + \operatorname{sgn}(\mathbf{d}) \odot \min(|\mathbf{d}|, \boldsymbol{\tau}) & \text{if } \sum_{i=1}^n \delta q_i \left( \max(|d_i| - \tau_i, 0) \right)^2 > k, \\ \mathbf{w}^{\text{cur}} & \text{otherwise.} \end{cases} \tag{34}
$$

Here, $\odot$ denotes element-wise multiplication and $\boldsymbol{\tau} = (\tau_1, \dots, \tau_n)^\top$ is the threshold vector. Geometrically, (34) creates two nested regions around the target portfolio. An inner hyper-rectangle $\mathcal{Z}_0 = \{\mathbf{d} \in \mathbb{R}^n \mid |d_i| \le \tau_i \ \forall i\}$ where

variable costs alone prohibit trading ($B(\mathbf{d}) = 0$) is surrounded by a wider outer inaction region $\mathcal{Z}_k = \{\mathbf{d} \in \mathbb{R}^n \mid \sum_{i=1}^n \delta q_i \max(|d_i| - \tau_i, 0)^2 \leq k\}$ shaped like a rounded box. In the buffer shell between $\mathcal{Z}_0$ and $\mathcal{Z}_k$, trading would be optimal under variable costs alone, but the benefit of doing so is too small to cover the fixed fee $k$. Once $\mathbf{d}$ crosses the outer trigger boundary $\partial\mathcal{Z}_k$, we trade across the buffer shell onto the inner boundary $\partial\mathcal{Z}_0$, which prevents high-frequency chattering of tiny trades on consecutive days.

Equation (34) reveals a second important principle, which is that **combining variable and fixed transaction costs creates a two-band structure that separates the outer trigger boundary from the inner destination boundary.** Under variable costs alone ($k = 0$), both boundaries coincide at $\partial\mathcal{Z}_0$, so any deviation beyond $\pm\tau_i$ immediately triggers a trade to $\pm\tau_i$. Under fixed costs alone ($\mathbf{c} = \mathbf{0}, k > 0$), the inner band vanishes ($\mathcal{Z}_0 = \{\mathbf{0}\}$) and breaching the trigger boundary $\partial\mathcal{Z}_k$ sends the portfolio all the way to the target $\mathbf{x}^* = \mathbf{0}$. Only if both frictions act together ($k > 0, \mathbf{c} > \mathbf{0}$) does the threshold split into two distinct bands.

# Coupled $\mathbf{Q} = \mathbf{\Sigma}$ Case

## Stage 1

Let us now solve the same two-stage problem for the economically natural specification $\mathbf{Q} = \mathbf{\Sigma}$, where $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ is the symmetric positive definite return covariance matrix. Proceeding in strict analogy to (13)–(15) of the diagonal case, we substitute $\mathbf{Q} = \mathbf{\Sigma}$ into the continuous cost function $\tilde{C}(\mathbf{x})$ and expand both the inner product and the quadratic form into explicit sums over the individual assets.

$$\tilde{C}(\mathbf{x}) = \mathbf{c}^\top |\mathbf{x} - \mathbf{d}| + \delta \mathbf{x}^\top \mathbf{\Sigma} \mathbf{x} \tag{35}$$

$$= \sum_{i=1}^n c_i |x_i - d_i| + \delta \sum_{r=1}^n \sum_{s=1}^n \Sigma_{rs} x_r x_s \tag{36}$$

$$= \sum_{i=1}^n c_i |x_i - d_i| + \delta \sum_{i=1}^n \Sigma_{ii} x_i^2 + \delta \sum_{i=1}^n \sum_{j \neq i} \Sigma_{ij} x_i x_j \tag{37}$$

In (35) we write the continuous cost function for $\mathbf{Q} = \mathbf{\Sigma}$. In (36) we expand the inner product $\mathbf{c}^\top |\mathbf{x} - \mathbf{d}|$ into a single sum and the quadratic tracking-error penalty $\delta \mathbf{x}^\top \mathbf{\Sigma} \mathbf{x}$ into a double sum over row index $r \in \{1, \ldots, n\}$ and column index $s \in \{1, \ldots, n\}$. In (37) we split this double sum into the $n$ diagonal variance terms where $r = s = i$ and the $n(n-1)$ off-diagonal covariance terms

where $r = i$ and $s = j \neq i$. Comparing (37) with its diagonal counterpart (15) shows why the coupled $\mathbf{Q} = \mathbf{\Sigma}$ case is not as simple. In (15), all off-diagonal entries were zero, allowing us to minimize each asset's cost $\tilde{C}_i(x_i)$ in isolation. In (37), however, the cross-terms $\Sigma_{ij} x_i x_j$ couple all $n$ assets together, because changing the post-trade deviation $x_j$ of asset $j$ alters the tracking-error penalty of asset $i$.

## No-Trade Threshold

To derive the first-order optimality condition for asset $i$ just as we did in (16), we need the partial derivative of the smooth tracking-error penalty $\delta \mathbf{x}^\top \mathbf{\Sigma} \mathbf{x}$ with respect to $x_i$. In the double sum $\delta \sum_{r=1}^{n} \sum_{s=1}^{n} \Sigma_{rs} x_r x_s$ from (36), the specific variable $x_i$ appears on the diagonal where $r = i$ and $s = i$, in the $i$-th row where $r = i$ and $s = j \neq i$, and in the $i$-th column where $r = j \neq i$ and $s = i$, whereas all remaining terms with $r \neq i$ and $s \neq i$ do not contain $x_i$.

$$\delta \mathbf{x}^\top \mathbf{\Sigma} \mathbf{x} = \delta \Sigma_{ii} x_i^2 + \delta \sum_{j \neq i} \Sigma_{ij} x_i x_j + \delta \sum_{j \neq i} \Sigma_{ji} x_j x_i + \delta \sum_{r \neq i} \sum_{s \neq i} \Sigma_{rs} x_r x_s \tag{38}$$

$$= \delta \Sigma_{ii} x_i^2 + 2\delta \sum_{j \neq i} \Sigma_{ij} x_i x_j + \delta \sum_{r \neq i} \sum_{s \neq i} \Sigma_{rs} x_r x_s \tag{39}$$

$$\frac{\partial}{\partial x_i} \left( \delta \mathbf{x}^\top \mathbf{\Sigma} \mathbf{x} \right) = 2\delta \Sigma_{ii} x_i + 2\delta \sum_{j \neq i} \Sigma_{ij} x_j + 0 \tag{40}$$

$$= 2\delta \sum_{j=1}^{n} \Sigma_{ij} x_j \tag{41}$$

$$= 2\delta (\mathbf{\Sigma} \mathbf{x})_i \tag{42}$$

In (38) we partition the double sum from (36) into the diagonal term, the $i$-th row sum, the $i$-th column sum, and the residual double sum free of $x_i$. In (39) we use the symmetry of the covariance matrix ($\Sigma_{ji} = \Sigma_{ij}$) and scalar commutativity ($x_j x_i = x_i x_j$) to merge the row and column sums into $2\delta \sum_{j \neq i} \Sigma_{ij} x_i x_j$. In (40) we take the partial derivative with respect to $x_i$, where the residual double sum vanishes. In (41) we absorb the diagonal term $2\delta \Sigma_{ii} x_i$ back into the single sum over all $j \in \{1, \ldots, n\}$. Finally, in (42) we recognize $\sum_{j=1}^{n} \Sigma_{ij} x_j$ as the inner product of the $i$-th row of $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ with the column vector $\mathbf{x} \in \mathbb{R}^n$, which is the $i$-th scalar entry $(\mathbf{\Sigma} \mathbf{x})_i \in \mathbb{R}$ of the $n \times 1$ matrix-vector product $\mathbf{\Sigma} \mathbf{x} \in \mathbb{R}^n$.

With the partial derivative (42) in hand, we now mirror (16)–(20) step by step. The strictly convex function $\tilde{C}(\mathbf{x})$ attains its unique global minimum at $\mathbf{x}^* \in \mathbb{R}^n$ if and only if $0$ lies in the partial subdifferential of $\tilde{C}(\mathbf{x}^*)$ with respect to every coordinate $x_i$ simultaneously (*Karush-Kuhn-Tucker* (KKT) condition). We state

this condition for each asset $i$, stack all $n$ conditions into vector form, and evaluate at the status-quo point $\mathbf{x}^* = \mathbf{d}$ to determine when complete inaction is optimal.

$$0 \in c_i \partial |x_i^* - d_i| + 2\delta(\boldsymbol{\Sigma}\mathbf{x}^*)_i \quad \forall i \in \{1, \dots, n\} \tag{43}$$

$$\mathbf{0}_n \in \mathbf{c} \odot \partial |\mathbf{x}^* - \mathbf{d}| + 2\delta \boldsymbol{\Sigma}\mathbf{x}^* \tag{44}$$

$$\mathbf{0}_n = \mathbf{c} \odot \boldsymbol{\xi} + 2\delta \boldsymbol{\Sigma}\mathbf{x}^* \tag{45}$$

$$-2\delta \boldsymbol{\Sigma}\mathbf{x}^* = \mathbf{c} \odot \boldsymbol{\xi} \tag{46}$$

$$-2\delta(\boldsymbol{\Sigma}\mathbf{d})_i \in [-c_i, c_i] \quad \forall i \in \{1, \dots, n\} \tag{47}$$

$$|2\delta(\boldsymbol{\Sigma}\mathbf{d})_i| \leq c_i \quad \forall i \in \{1, \dots, n\} \tag{48}$$

$$\left| \Sigma_{ii} d_i + \sum_{j \neq i} \Sigma_{ij} d_j \right| \leq \frac{c_i}{2\delta} \quad \forall i \in \{1, \dots, n\} \tag{49}$$

In (43) we state the first-order optimality condition for coordinate $i$ by adding the variable cost subdifferential $c_i \partial |x_i^* - d_i|$ and the partial derivative $2\delta(\boldsymbol{\Sigma}\mathbf{x}^*)_i$ from (42). In (44) we stack the $n$ scalar conditions into an $n \times 1$ vector, where $\mathbf{0}_n \in \mathbb{R}^n$ and $\partial|\mathbf{x}^* - \mathbf{d}|$ acts element-wise. In (45) we write this inclusion as an equation by introducing the $n \times 1$ subgradient vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^\top \in \mathbb{R}^n$, where $\xi_i = \operatorname{sgn}(x_i^* - d_i) \in \{-1, +1\}$ if asset $i$ is traded ($x_i^* \neq d_i$) and $\xi_i \in [-1, 1]$ if asset $i$ is not traded ($x_i^* = d_i$). In (46) we move the marginal tracking-error vector to the left-hand side. To check when refraining from trading across all assets ($\mathbf{x}^* = \mathbf{d}$) is optimal, in (47) we substitute $\mathbf{x}^* = \mathbf{d}$ and $\xi_i \in [-1, 1]$ into the $i$-th row of (46). In (48) we rewrite the interval inclusion as an absolute value inequality, and in (49) we divide by $2\delta > 0$ and expand $(\boldsymbol{\Sigma}\mathbf{d})_i$ using (40)–(42).

Comparing (48)–(49) with the diagonal threshold (19)–(20) reveals a third important principle, namely that **when assets are coupled through their covariance matrix, weight deviations across correlated assets can offset each other at the portfolio level**. Say in (49) that all off-diagonal covariances $\Sigma_{ij}$ ($j \neq i$) are zero and $\Sigma_{ii} = q_i$. Then (49) reduces to $|q_i d_i| \leq \frac{c_i}{2\delta}$, which is identical to the diagonal threshold (20). When correlations are non-zero, however, asset $i$'s standalone variance term $\Sigma_{ii} d_i$ is augmented inside the absolute value operator by the cross-covariance sum $\sum_{j \neq i} \Sigma_{ij} d_j$. Economically, $(\boldsymbol{\Sigma}\mathbf{d})_i = \operatorname{Cov}(r_i, r_p - r_{\text{tar}})$ is the covariance between asset $i$'s return and the tracking error return of the entire deviated portfolio. If asset $i$ and asset $j$ are negatively correlated ($\Sigma_{ij} < 0$) and both deviate in the same direction ($d_i > 0, d_j > 0$), or if they are positively correlated ($\Sigma_{ij} > 0$) and deviate in opposite directions ($d_i > 0, d_j < 0$), then $\Sigma_{ij} d_j$ has the opposite sign of $\Sigma_{ii} d_i$ and cancels out part of asset $i$'s standalone deviation inside (49). Because the

two weight deviations naturally hedge each other's contribution to portfolio tracking error, the marginal benefit $|2\delta(\boldsymbol{\Sigma}\mathbf{d})_i|$ of trading either asset may stay below its variable cost $c_i$ even when individual weight deviations $d_i$ are quite large. Geometrically, the inner no-trade zone
$\mathcal{Z}_0 = \{\mathbf{d} \in \mathbb{R}^n \mid |2\delta(\boldsymbol{\Sigma}\mathbf{d})_i| \le c_i \ \forall i\} = \boldsymbol{\Sigma}^{-1}[-\frac{1}{2\delta}\mathbf{c}, \frac{1}{2\delta}\mathbf{c}]$ maps an axis-aligned box through $\boldsymbol{\Sigma}^{-1}$ into a sheared $n$-dimensional parallelotope that is elongated along self-hedging directions.

## Optimal Active Trade

Now let's move on to the case when $\mathbf{d}$ breaches the parallelotope ($\mathbf{d} \notin \mathcal{Z}_0$), so that $|2\delta(\boldsymbol{\Sigma}\mathbf{d})_i| > c_i$ for at least one asset. We attempt to follow the same steps as in (21)–(24) of the diagonal case by taking an actively traded asset $i$ ($x_i^* \neq d_i$, where the subgradient is the trade sign $\xi_i = s_i = \operatorname{sgn}(x_i^* - d_i) \in \{-1, +1\}$) and solving its first-order condition (43) for $x_i^*$ using the expansion from (40)–(42).

$$0 = c_i \operatorname{sgn}(x_i^* - d_i) + 2\delta(\boldsymbol{\Sigma}\mathbf{x}^*)_i \tag{50}$$

$$0 = c_i s_i + 2\delta\Sigma_{ii}x_i^* + 2\delta\sum_{j\neq i}\Sigma_{ij}x_j^* \tag{51}$$

$$2\delta\Sigma_{ii}x_i^* = -c_i s_i - 2\delta\sum_{j\neq i}\Sigma_{ij}x_j^* \tag{52}$$

$$x_i^* = -\frac{c_i s_i}{2\delta\Sigma_{ii}} - \sum_{j\neq i}\frac{\Sigma_{ij}}{\Sigma_{ii}}x_j^* \tag{53}$$

In (50) we replace the subdifferential in (43) for an active trade $x_i^* \neq d_i$ with $\operatorname{sgn}(x_i^* - d_i)$. In (51) we write $s_i = \operatorname{sgn}(x_i^* - d_i)$ and expand $2\delta(\boldsymbol{\Sigma}\mathbf{x}^*)_i$ via (40). In (52) we isolate the diagonal term $2\delta\Sigma_{ii}x_i^*$ on the left-hand side, and in (53) we divide by $2\delta\Sigma_{ii} > 0$.

Comparing (53) with its diagonal counterpart (24) ($x_i^* = \operatorname{sgn}(d_i)\frac{c_i}{2\delta q_i}$) reveals the reasons why a naive asset-by-asset solution is suboptimal. In (24), the right-hand side contained only known constants, immediately fixing $x_i^*$. In (53), however, asset $i$'s optimal post-trade deviation $x_i^*$ depends on the unknown post-trade deviations $x_j^*$ of all other assets $j \neq i$. Regarding (45) ($\mathbf{0}_n = \mathbf{c} \odot \boldsymbol{\xi} + 2\delta\boldsymbol{\Sigma}\mathbf{x}^*$), we can't solve for all $x_1^*, \ldots, x_n^*$ at once, because in general only a subset of assets is actively traded while the remaining assets are held untouched. For an actively traded asset $i$ ($x_i^* \neq d_i$), its subgradient is fixed on the smooth flank of the V-kink at $\xi_i = s_i \in \{-1, +1\}$, while its post-trade deviation $x_i^*$ is the unknown variable. For an untraded asset $k$ ($x_k^* = d_k$), it is the other way around, as its post-trade deviation is known to equal its pre-trade deviation $x_k^* = d_k$ (at the bottom of the V-kink), while its subgradient $\xi_k \in [-1, 1]$ is an unknown value inside the interval $[-1, 1]$.

To solve such a system, we must separate the two groups of rows using block matrices. Suppose that at the optimum, $m \in \{1, \dots, n\}$ assets are actively traded and $n - m$ assets are not traded. Let $\mathcal{A} = \{i \mid x_i^* \neq d_i\}$ with cardinality $|\mathcal{A}| = m$ denote the active set, and let $\mathcal{I} = \{1, \dots, n\} \setminus \mathcal{A}$ with cardinality $|\mathcal{I}| = n - m$ denote the inactive set. By reordering the asset indices so that the $m$ active assets come first and the $n - m$ inactive assets come last, we partition every $n \times 1$ vector and the $n \times n$ covariance matrix $\mathbf{\Sigma}$ into blocks with explicit dimensions. Here, $\mathbf{x}_{\mathcal{A}}^* \in \mathbb{R}^{m \times 1}$ is the column vector of unknown post-trade deviations of the $m$ active assets, and $\mathbf{x}_{\mathcal{I}}^* = \mathbf{d}_{\mathcal{I}} \in \mathbb{R}^{(n-m) \times 1}$ is the column vector of known pre-trade deviations of the $n - m$ untraded assets. Furthermore, $\mathbf{c}_{\mathcal{A}} \in \mathbb{R}_{>0}^{m \times 1}$ and $\mathbf{s}_{\mathcal{A}} = \operatorname{sgn}(\mathbf{x}_{\mathcal{A}}^* - \mathbf{d}_{\mathcal{A}}) \in \{-1, +1\}^{m \times 1}$ are the $m \times 1$ vectors of active variable costs and active trade signs (so $\boldsymbol{\xi}_{\mathcal{A}} = \mathbf{s}_{\mathcal{A}}$), while $\mathbf{c}_{\mathcal{I}} \in \mathbb{R}_{>0}^{(n-m) \times 1}$ and $\boldsymbol{\xi}_{\mathcal{I}} \in [-1, 1]^{(n-m) \times 1}$ are the $(n - m) \times 1$ vectors of inactive variable costs and unknown inactive subgradients. Finally, $\mathbf{\Sigma}_{\mathcal{A}\mathcal{A}} \in \mathbb{R}^{m \times m}$ is the covariance submatrix of the $m$ active assets with themselves, $\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}} \in \mathbb{R}^{m \times (n-m)}$ is the cross-covariance submatrix between the $m$ active assets (rows) and the $n - m$ inactive assets (columns), $\mathbf{\Sigma}_{\mathcal{I}\mathcal{A}} = \mathbf{\Sigma}_{\mathcal{A}\mathcal{I}}^\top \in \mathbb{R}^{(n-m) \times m}$ is the cross-covariance submatrix between the $n - m$ inactive assets (rows) and the $m$ active assets (columns), and $\mathbf{\Sigma}_{\mathcal{I}\mathcal{I}} \in \mathbb{R}^{(n-m) \times (n-m)}$ is the covariance submatrix of the $n - m$ inactive assets with themselves.

Substituting these partitioned vectors and submatrices into the $n \times 1$ KKT equation (45) and carrying out the block matrix-vector multiplication row by column gives the partitioned system.

$$\begin{pmatrix} \mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}} \\ \mathbf{c}_{\mathcal{I}} \odot \boldsymbol{\xi}_{\mathcal{I}} \end{pmatrix} + 2\delta \begin{pmatrix} \mathbf{\Sigma}_{\mathcal{A}\mathcal{A}} & \mathbf{\Sigma}_{\mathcal{A}\mathcal{I}} \\ \mathbf{\Sigma}_{\mathcal{I}\mathcal{A}} & \mathbf{\Sigma}_{\mathcal{I}\mathcal{I}} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{\mathcal{A}}^* \\ \mathbf{d}_{\mathcal{I}} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_m \\ \mathbf{0}_{n-m} \end{pmatrix} \tag{54}$$

$$\begin{pmatrix} \mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}} \\ \mathbf{c}_{\mathcal{I}} \odot \boldsymbol{\xi}_{\mathcal{I}} \end{pmatrix} + \begin{pmatrix} 2\delta\mathbf{\Sigma}_{\mathcal{A}\mathcal{A}}\mathbf{x}_{\mathcal{A}}^* + 2\delta\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}}\mathbf{d}_{\mathcal{I}} \\ 2\delta\mathbf{\Sigma}_{\mathcal{I}\mathcal{A}}\mathbf{x}_{\mathcal{A}}^* + 2\delta\mathbf{\Sigma}_{\mathcal{I}\mathcal{I}}\mathbf{d}_{\mathcal{I}} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_m \\ \mathbf{0}_{n-m} \end{pmatrix} \tag{55}$$

$$\begin{pmatrix} \mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}} + 2\delta\mathbf{\Sigma}_{\mathcal{A}\mathcal{A}}\mathbf{x}_{\mathcal{A}}^* + 2\delta\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}}\mathbf{d}_{\mathcal{I}} \\ \mathbf{c}_{\mathcal{I}} \odot \boldsymbol{\xi}_{\mathcal{I}} + 2\delta\mathbf{\Sigma}_{\mathcal{I}\mathcal{A}}\mathbf{x}_{\mathcal{A}}^* + 2\delta\mathbf{\Sigma}_{\mathcal{I}\mathcal{I}}\mathbf{d}_{\mathcal{I}} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_m \\ \mathbf{0}_{n-m} \end{pmatrix} \tag{56}$$

In (54) we write the $n \times 1$ KKT condition (45) in partitioned block form, where $\mathbf{0}_m \in \mathbb{R}^{m \times 1}$ and $\mathbf{0}_{n-m} \in \mathbb{R}^{(n-m) \times 1}$ denote the respective zero column vectors. In (55) we multiply the partitioned $n \times n$ covariance matrix by the partitioned $n \times 1$ deviation vector. In the top $m \times 1$ block, multiplying $\mathbf{\Sigma}_{\mathcal{A}\mathcal{A}} \in \mathbb{R}^{m \times m}$ by $\mathbf{x}_{\mathcal{A}}^* \in \mathbb{R}^{m \times 1}$ yields an $m \times 1$ vector, and multiplying $\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}} \in \mathbb{R}^{m \times (n-m)}$ by $\mathbf{d}_{\mathcal{I}} \in \mathbb{R}^{(n-m) \times 1}$ also yields an $m \times 1$ vector; analogously, both products in the bottom block have dimension $(n - m) \times 1$. In (56) we add the two stacked column vectors block by block.

The unknown inactive subgradient vector $\boldsymbol{\xi}_{\mathcal{I}} \in [-1,1]^{(n-m)\times 1}$ appears only in the bottom block row, while the top $m \times 1$ block row contains only the $m \times 1$ unknown vector $\mathbf{x}^*_{\mathcal{A}}$ alongside the known vector $\mathbf{d}_{\mathcal{I}}$ and the sign vector $\mathbf{s}_{\mathcal{A}}$. Setting the top $m \times 1$ block of (56) equal to $\mathbf{0}_m$ therefore gives a self-contained system of $m$ linear equations in $m$ unknowns.

$$
\begin{aligned}
\mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}} + 2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}} + 2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{I}} \mathbf{d}_{\mathcal{I}} &= \mathbf{0}_m && (57)\\
2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}} &= -2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{I}} \mathbf{d}_{\mathcal{I}} - \mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}} && (58)\\
\frac{1}{2\delta} \boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \left(2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}}\right) &= \frac{1}{2\delta} \boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \left(-2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{I}} \mathbf{d}_{\mathcal{I}} - \mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}}\right) && (59)\\
\left(\boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}}\right) \mathbf{x}^*_{\mathcal{A}} &= -\boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{I}} \mathbf{d}_{\mathcal{I}} - \frac{1}{2\delta} \boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} (\mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}}) && (60)\\
\mathbf{I}_m \mathbf{x}^*_{\mathcal{A}} &= -\boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{I}} \mathbf{d}_{\mathcal{I}} - \frac{1}{2\delta} \boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} (\mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}}) && (61)\\
\mathbf{x}^*_{\mathcal{A}} &= -\boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{I}} \mathbf{d}_{\mathcal{I}} - \frac{1}{2\delta} \boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} (\mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}}) && (62)
\end{aligned}
$$

In (57) we equate the top $m \times 1$ block of (56) to $\mathbf{0}_m$. In (58) we subtract $2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{I}} \mathbf{d}_{\mathcal{I}}$ and $\mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}}$ from both sides to isolate the unknown active term $2\delta \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}}$ on the left-hand side. Because $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ is symmetric positive definite, its principal submatrix $\boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}} \in \mathbb{R}^{m \times m}$ is also symmetric positive definite and therefore invertible; in (59) we left-multiply both sides by the $m \times m$ matrix $\frac{1}{2\delta} \boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}}$. In (60) we cancel the scalar $2\delta$ and group $\boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}}$ on the left-hand side. In (61) we replace $\boldsymbol{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \boldsymbol{\Sigma}_{\mathcal{A}\mathcal{A}}$ with the $m \times m$ identity matrix $\mathbf{I}_m$, and in (62) we obtain the closed-form block solution for $\mathbf{x}^*_{\mathcal{A}} \in \mathbb{R}^{m \times 1}$. Subtracting $\mathbf{d}_{\mathcal{A}} \in \mathbb{R}^{m \times 1}$ gives the optimal active trade vector $\boldsymbol{\Delta}^*_{\mathcal{A}} = \mathbf{x}^*_{\mathcal{A}} - \mathbf{d}_{\mathcal{A}} \in \mathbb{R}^{m \times 1}$.

Once $\mathbf{x}^*_{\mathcal{A}}$ is computed from the top block via (62), we substitute it into the bottom $(n-m) \times 1$ block of (56) to check that the untraded assets $\mathcal{I}$ indeed remain inside their no-trade thresholds.

$$
\begin{aligned}
\mathbf{c}_{\mathcal{I}} \odot \boldsymbol{\xi}_{\mathcal{I}} + 2\delta \boldsymbol{\Sigma}_{\mathcal{I}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}} + 2\delta \boldsymbol{\Sigma}_{\mathcal{I}\mathcal{I}} \mathbf{d}_{\mathcal{I}} &= \mathbf{0}_{n-m} && (63)\\
-2\delta \left(\boldsymbol{\Sigma}_{\mathcal{I}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}} + \boldsymbol{\Sigma}_{\mathcal{I}\mathcal{I}} \mathbf{d}_{\mathcal{I}}\right) &= \mathbf{c}_{\mathcal{I}} \odot \boldsymbol{\xi}_{\mathcal{I}} && (64)\\
-2\delta \left(\boldsymbol{\Sigma}_{\mathcal{I}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}} + \boldsymbol{\Sigma}_{\mathcal{I}\mathcal{I}} \mathbf{d}_{\mathcal{I}}\right) &\in [-\mathbf{c}_{\mathcal{I}}, \mathbf{c}_{\mathcal{I}}] && (65)\\
\left|2\delta \left(\boldsymbol{\Sigma}_{\mathcal{I}\mathcal{A}} \mathbf{x}^*_{\mathcal{A}} + \boldsymbol{\Sigma}_{\mathcal{I}\mathcal{I}} \mathbf{d}_{\mathcal{I}}\right)_k\right| &\le c_k \quad \forall k \in \mathcal{I} && (66)
\end{aligned}
$$

In (63) we equate the bottom $(n-m) \times 1$ block of (56) to $\mathbf{0}_{n-m}$. In (64) we move the post-trade marginal tracking-error vector of the inactive assets to the left-hand side. In (65) we impose the subdifferential requirement $\boldsymbol{\xi}_{\mathcal{I}} \in [-1,1]^{(n-m)\times 1}$, which we write as an element-wise absolute value inequality for each untraded asset $k \in \mathcal{I}$ in (66). Because strict convexity of

$\tilde{C}(\mathbf{x})$ guarantees a unique global minimum, a candidate regime $(\mathcal{A}, \mathcal{I}, \mathbf{s}_{\mathcal{A}})$ is the global solution if and only if $\mathbf{x}^*_{\mathcal{A}}$ from (62) satisfies the active sign consistency $\text{sgn}(\mathbf{x}^*_{\mathcal{A}} - \mathbf{d}_{\mathcal{A}}) = \mathbf{s}_{\mathcal{A}}$ and the inactive boundary check (66).

## Proxy Hedging via Regression Betas

Interestingly, the first term $-\mathbf{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}}\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}}\mathbf{d}_{\mathcal{I}}$ in (62) has the form of a multivariate ordinary least squares (OLS) regression and points to another important principle. Let's first derive this relationship explicitly. Let $\mathbf{r}_{\mathcal{A}} \in \mathbb{R}^{m \times 1}$ denote the random return vector of the $m$ active assets and $\mathbf{r}_{\mathcal{I}} \in \mathbb{R}^{(n-m) \times 1}$ denote the random return vector of the $n - m$ inactive assets. Suppose we linearly regress the $(n - m) \times 1$ inactive returns $\mathbf{r}_{\mathcal{I}}$ on the $m \times 1$ active returns $\mathbf{r}_{\mathcal{A}}$ via the multivariate linear regression model $\mathbf{r}_{\mathcal{I}} = \boldsymbol{\alpha} + \mathbf{B}^\top \mathbf{r}_{\mathcal{A}} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\alpha} \in \mathbb{R}^{(n-m) \times 1}$ is the vector of intercepts, $\mathbf{B} \in \mathbb{R}^{m \times (n-m)}$ is the $m \times (n - m)$ matrix of regression slope coefficients (whose $k$-th column contains the $m$ regression betas of inactive asset $k \in \mathcal{I}$ on the $m$ active assets, so $\mathbf{B}^\top \in \mathbb{R}^{(n-m) \times m}$), and $\boldsymbol{\varepsilon} \in \mathbb{R}^{(n-m) \times 1}$ is the zero-mean residual vector that is uncorrelated with the regressors $\mathbf{r}_{\mathcal{A}}$, meaning $\text{Cov}(\mathbf{r}_{\mathcal{A}}, \boldsymbol{\varepsilon}) = \mathbf{0}_{m \times (n-m)}$.

$$
\begin{aligned}
\mathbf{r}_{\mathcal{I}} &= \boldsymbol{\alpha} + \mathbf{B}^\top \mathbf{r}_{\mathcal{A}} + \boldsymbol{\varepsilon} && (67)\\
\text{Cov}(\mathbf{r}_{\mathcal{A}}, \mathbf{r}_{\mathcal{I}}) &= \text{Cov}\Big(\mathbf{r}_{\mathcal{A}}, \boldsymbol{\alpha} + \mathbf{B}^\top \mathbf{r}_{\mathcal{A}} + \boldsymbol{\varepsilon}\Big) && (68)\\
\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}} &= \text{Cov}(\mathbf{r}_{\mathcal{A}}, \boldsymbol{\alpha}) + \text{Cov}(\mathbf{r}_{\mathcal{A}}, \mathbf{r}_{\mathcal{A}})\mathbf{B} + \text{Cov}(\mathbf{r}_{\mathcal{A}}, \boldsymbol{\varepsilon}) && (69)\\
\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}} &= \mathbf{0}_{m \times (n-m)} + \mathbf{\Sigma}_{\mathcal{A}\mathcal{A}}\mathbf{B} + \mathbf{0}_{m \times (n-m)} && (70)\\
\mathbf{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}}\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}} &= \Big(\mathbf{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}}\mathbf{\Sigma}_{\mathcal{A}\mathcal{A}}\Big)\,\mathbf{B} && (71)\\
\mathbf{B} &= \mathbf{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}}\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}} \in \mathbb{R}^{m \times (n-m)} && (72)
\end{aligned}
$$

In (67) we state the multivariate linear regression equation of $\mathbf{r}_{\mathcal{I}} \in \mathbb{R}^{(n-m) \times 1}$ on $\mathbf{r}_{\mathcal{A}} \in \mathbb{R}^{m \times 1}$. In (68) we take the $m \times (n - m)$ cross-covariance matrix $\text{Cov}(\mathbf{r}_{\mathcal{A}}, \mathbf{r}_{\mathcal{I}})$ on both sides. In (69) we use the bilinearity of the covariance operator, where $\text{Cov}(\mathbf{r}_{\mathcal{A}}, \mathbf{B}^\top \mathbf{r}_{\mathcal{A}}) = \text{Cov}(\mathbf{r}_{\mathcal{A}}, \mathbf{r}_{\mathcal{A}})(\mathbf{B}^\top)^\top = \mathbf{\Sigma}_{\mathcal{A}\mathcal{A}}\mathbf{B}$. In (70) the covariance with the constant vector $\boldsymbol{\alpha}$ and the orthogonal residual vector $\boldsymbol{\varepsilon}$ both vanish to the $m \times (n - m)$ zero matrix $\mathbf{0}_{m \times (n-m)}$, and we substitute $\text{Cov}(\mathbf{r}_{\mathcal{A}}, \mathbf{r}_{\mathcal{A}}) = \mathbf{\Sigma}_{\mathcal{A}\mathcal{A}} \in \mathbb{R}^{m \times m}$. In (71) we left-multiply both sides by $\mathbf{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}} \in \mathbb{R}^{m \times m}$, which collapses to the identity matrix in (72) and proves that $\mathbf{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}}\mathbf{\Sigma}_{\mathcal{A}\mathcal{I}}$ is the $m \times (n - m)$ OLS beta matrix $\mathbf{B}$. Substituting (72) into (62) allows us to write the optimal active post-trade deviation in terms of $\mathbf{B}$.

$$
\mathbf{x}^*_{\mathcal{A}} = -\mathbf{B}\mathbf{d}_{\mathcal{I}} - \frac{1}{2\delta}\mathbf{\Sigma}^{-1}_{\mathcal{A}\mathcal{A}}(\mathbf{c}_{\mathcal{A}} \odot \mathbf{s}_{\mathcal{A}}) \tag{73}
$$

Multiplying the $m \times (n-m)$ regression beta matrix $-\mathbf{B}$ by the $(n-m) \times 1$ untraded deviation vector $\mathbf{d}_{\mathcal{I}}$ yields an $m \times 1$ vector of weight adjustments across the $m$ active assets. Equation (73) therefore reveals a fourth important principle, showing that **when rebalancing a coupled multi-asset portfolio, it can be optimal to leave expensive assets untouched at their deviated weights while deliberately trading cheaper proxy assets away from their own target weights to hedge the unadjusted deviation.** If every asset in the portfolio is traded ($m = n, \mathcal{I} = \emptyset$), the first term in (73) vanishes and $\mathbf{x}^* = -\frac{1}{2\delta}\boldsymbol{\Sigma}^{-1}(\mathbf{c} \odot \mathbf{s})$, which is the multivariate counterpart of stopping at the no-trade boundary $\pm\tau_i$. Whenever rebalancing is partial ($m < n, \mathcal{I} \neq \emptyset$), however, suppose an illiquid micro cap equity fund $k \in \mathcal{I}$ has deviated overweight ($d_k > 0$), while a cheap, highly liquid equity index future $i \in \mathcal{A}$ is positively correlated with asset $k$ ($B_{ik} > 0$) and currently sits right at its target weight ($d_i = 0$). A naive asset-by-asset rule would see $d_i = 0$ and do nothing in the liquid future $i$, while either paying the wide variable costs $c_k$ to sell the illiquid asset $k$ or leaving its induced tracking error unhedged. The coupled solution (73), by contrast, leaves the expensive asset $k$ untouched ($x_k^* = d_k$) and uses the liquid proxy $i \in \mathcal{A}$ to sell $-B_{ik}d_k$, deliberately pushing the liquid asset away from its own target weight ($x_i^* < 0$) because the portfolio-wide tracking-error reduction from the cross-hedge outweighs both the small trading cost $c_i$ and the standalone weight deviation created in asset $i$.

Because each of the $n$ assets can be bought, held, or sold, there are $3^n$ possible regimes, and the cross-hedging feedback in (53) means we do not know in advance which regime applies. For small portfolios (such as tactical asset allocation across $n \leq 5$ asset classes) or when all assets trade ($m = n$), we can evaluate (73) directly in closed form; for larger $n$, a numerical procedure can be used to find the active set $\mathcal{A}$.

## Stage 2

Having solved Stage 1 for $\mathbf{x}^*$, we need to evaluate Stage 2 (the fixed-cost trigger (12)) in analogy to (29)–(34) of the diagonal case by deriving a closed-form expression for the portfolio's gross cost savings $B(\mathbf{d}) = \delta\mathbf{d}^\top\boldsymbol{\Sigma}\mathbf{d} - \tilde{C}(\mathbf{x}^*)$. In the diagonal case (31), we eliminated $c_i$ from the variable transaction cost term by substituting the first-order condition $c_i = 2\delta q_i \tau_i$. In the coupled case, we perform the same elimination in vector form using the KKT condition (46) ($\mathbf{c} \odot \boldsymbol{\xi} = -2\delta\boldsymbol{\Sigma}\mathbf{x}^*$). We know from the definition of $\boldsymbol{\xi}$ in (45) that for every asset $i \in \{1, \ldots, n\}$, the product $\xi_i(x_i^* - d_i)$ equals the absolute value $|x_i^* - d_i|$.

If asset $i$ is not traded ($x_i^* = d_i$), both $|x_i^* - d_i|$ and $\xi_i(x_i^* - d_i)$ equal zero, and if asset $i$ is traded ($x_i^* \neq d_i$), then $\xi_i = \operatorname{sgn}(x_i^* - d_i)$, where multiplying any real number by its sign gives its absolute value.

$$\begin{aligned}
\mathbf{c}^\top |\mathbf{x}^* - \mathbf{d}| &= \sum_{i=1}^{n} c_i |x_i^* - d_i| && (74)\\
&= \sum_{i=1}^{n} c_i \xi_i (x_i^* - d_i) && (75)\\
&= (\mathbf{c} \odot \boldsymbol{\xi})^\top (\mathbf{x}^* - \mathbf{d}) && (76)\\
&= (-2\delta \boldsymbol{\Sigma} \mathbf{x}^*)^\top (\mathbf{x}^* - \mathbf{d}) && (77)\\
&= -2\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} (\mathbf{x}^* - \mathbf{d}) && (78)
\end{aligned}$$

In (74) we expand the inner product $\mathbf{c}^\top |\mathbf{x}^* - \mathbf{d}|$ into a sum over all $n$ assets. In (75) we substitute the identity $|x_i^* - d_i| = \xi_i (x_i^* - d_i)$ for each asset $i$, and in (76) we write the sum as the inner product of the $n \times 1$ vector $\mathbf{c} \odot \boldsymbol{\xi}$ and the $n \times 1$ trade vector $\mathbf{x}^* - \mathbf{d}$. In (77) we eliminate $\mathbf{c} \odot \boldsymbol{\xi}$ by substituting $-2\delta \boldsymbol{\Sigma} \mathbf{x}^*$ from the KKT condition (46), and in (78) we apply the matrix transpose rule $(\boldsymbol{\Sigma} \mathbf{x}^*)^\top = (\mathbf{x}^*)^\top \boldsymbol{\Sigma}^\top = (\mathbf{x}^*)^\top \boldsymbol{\Sigma}$ using the symmetry of $\boldsymbol{\Sigma}$.

Substituting the variable transaction cost (78) into the gross cost savings $B(\mathbf{d}) = \delta \mathbf{d}^\top \boldsymbol{\Sigma} \mathbf{d} - \tilde{C}(\mathbf{x}^*)$ from (11)–(12) mirrors the binomial completion of the square in (29)–(33).

$$\begin{aligned}
B(\mathbf{d}) &= \delta \mathbf{d}^\top \boldsymbol{\Sigma} \mathbf{d} - \left( \mathbf{c}^\top |\mathbf{x}^* - \mathbf{d}| + \delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^* \right) && (79)\\
&= \delta \mathbf{d}^\top \boldsymbol{\Sigma} \mathbf{d} - \left( -2\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} (\mathbf{x}^* - \mathbf{d}) + \delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^* \right) && (80)\\
&= \delta \mathbf{d}^\top \boldsymbol{\Sigma} \mathbf{d} + 2\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^* - 2\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{d} - \delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^* && (81)\\
&= \delta \left( \mathbf{d}^\top \boldsymbol{\Sigma} \mathbf{d} - 2 (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{d} + (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^* \right) && (82)\\
&= \delta (\mathbf{d} - \mathbf{x}^*)^\top \boldsymbol{\Sigma} (\mathbf{d} - \mathbf{x}^*) = \delta (\boldsymbol{\Delta}^*)^\top \boldsymbol{\Sigma} \boldsymbol{\Delta}^* && (83)
\end{aligned}$$

In (79) we subtract the continuous post-trade cost $\tilde{C}(\mathbf{x}^*)$ from the status-quo holding cost $\delta \mathbf{d}^\top \boldsymbol{\Sigma} \mathbf{d}$. In (80) we replace the variable transaction cost $\mathbf{c}^\top |\mathbf{x}^* - \mathbf{d}|$ with (78). In (81) we multiply out $-2\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} (\mathbf{x}^* - \mathbf{d})$ and dissolve the outer brackets. In (82) we combine $+2\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^*$ and $-\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^*$ into $+\delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^*$ and factor out the scalar $\delta$. Finally, in (83) we compress the quadratic form via the multivariate binomial identity $(\mathbf{d} - \mathbf{x}^*)^\top \boldsymbol{\Sigma} (\mathbf{d} - \mathbf{x}^*)$ and substitute $\boldsymbol{\Delta}^* = \mathbf{x}^* - \mathbf{d}$ from (4).

Comparing (83) with its diagonal counterpart (33) ($\sum_{i=1}^{n} \delta q_i (\Delta_i^*)^2$) shows that once variable transaction costs are deducted, the net benefit $B(\mathbf{d})$ of an optimal rebalance is not the raw difference in portfolio variance $\delta \mathbf{d}^\top \boldsymbol{\Sigma} \mathbf{d} - \delta (\mathbf{x}^*)^\top \boldsymbol{\Sigma} \mathbf{x}^*$, but $\delta$ times the tracking-error variance of the executed trade vector $\boldsymbol{\Delta}^*$ itself.

Inserting (83) into the fixed-cost trigger condition (12) yields the global decision rule for the coupled case.

$$\mathbf{w}^{\text{new}*} = \begin{cases} \mathbf{w}^{\text{tar}} + \mathbf{x}^* & \text{if } \delta(\mathbf{d}-\mathbf{x}^*)^\top \boldsymbol{\Sigma} (\mathbf{d}-\mathbf{x}^*) > k, \\ \mathbf{w}^{\text{cur}} & \text{otherwise.} \end{cases} \tag{84}$$

Geometrically, (84) generalizes the two-band structure of (34) from axis-aligned boxes to sheared parallelotopes. Around the target portfolio lies the inner parallelotope $\mathcal{Z}_0 = \{\mathbf{d} \in \mathbb{R}^n \mid |2\delta(\boldsymbol{\Sigma}\mathbf{d})_i| \le c_i \ \forall i\}$, where variable costs alone prohibit trading ($B(\mathbf{d}) = 0$). Whenever $\mathbf{d}$ lies outside $\mathcal{Z}_0$, at least one asset $i \in \mathcal{A}$ is actively traded with $|\xi_i| = 1$ while all untraded assets $k \in \mathcal{I}$ satisfy $|\xi_k| \le 1$, so the KKT condition (46) ensures that the post-trade destination $\mathbf{x}^*$ always stops right on the inner boundary $\partial\mathcal{Z}_0$, where $|2\delta(\boldsymbol{\Sigma}\mathbf{x}^*)_i| = c_i$ holds for the traded assets and $|2\delta(\boldsymbol{\Sigma}\mathbf{x}^*)_k| \le c_k$ holds for the untraded ones. Surrounding $\mathcal{Z}_0$ is the wider outer inaction region
$\mathcal{Z}_k = \{\mathbf{d} \in \mathbb{R}^n \mid \delta(\mathbf{d}-\mathbf{x}^*)^\top \boldsymbol{\Sigma} (\mathbf{d}-\mathbf{x}^*) \le k\}$, which collects all pre-trade deviations whose covariance-weighted distance to their boundary destination $\mathbf{x}^* \in \partial\mathcal{Z}_0$ does not exceed $k$, forming a sheared parallelotope with elliptically rounded corners. In the buffer shell between $\mathcal{Z}_0$ and $\mathcal{Z}_k$, trading would be optimal under variable costs alone, but the net benefit is still too small to cover the fixed fee $k$. Once $\mathbf{d}$ crosses the outer trigger boundary $\partial\mathcal{Z}_k$, we trade across the buffer shell onto the inner boundary $\partial\mathcal{Z}_0$.

# Conclusion

Thinking about rebalancing problem in this way replaces ad-hoc calendar or heuristic band rules with four more reasonable principles. First, proportional transaction costs in combination with a convex suboptimality penalty term create an inner no-trade boundary $\partial\mathcal{Z}_0$. Because quadratic suboptimality has vanishing marginal cost near the origin while variable trading costs remain constant down to the last dollar traded, an optimal trade never rebalances all the way back to the target portfolio, but stops right at the edge of the inner band. Second, adding fixed transaction costs $k$ decouples the outer trigger boundary $\partial\mathcal{Z}_k$ from the inner destination boundary $\partial\mathcal{Z}_0$, creating a buffer shell in which small deviations are tolerated until the net cost savings of trading clear the lump-sum hurdle $k$. Third, measuring portfolio suboptimality by the return covariance matrix $\boldsymbol{\Sigma}$ rather than by isolated weight deviations transforms the inner no-trade zone from a rigid box into a sheared parallelotope $\mathcal{Z}_0$, allowing offsetting deviations across correlated assets to naturally hedge each other without triggering unnecessary trades. Fourth, when deviations do breach the outer

boundary and only a subset of assets is traded, the coupled KKT solution automatically uses cheaper assets as proxies to hedge the unadjusted deviation of more expensive assets, even when doing so deliberately pushes those cheaper proxies away from their own target weights.