A Further Intuition

To motivate this additional "+1" in the KO+ scheme, we consider the FDP for the KO+ pipeline with threshold $\hat{t}_+$ defined in (11):

$$\begin{aligned}\mathrm{FDP}(\hat{t}_+) &= \frac{\#\big\{(i,j):(i,j)\notin\mathcal{E},\widehat{W}_{ij}\geq\hat{t}_+\big\}}{\#\{(i,j):\widehat{W}_{ij}\geq\hat{t}_+\}\vee 1}\\ &\leq \frac{\#\{(i,j):(i,j)\notin\mathcal{E},\widehat{W}_{ij}\geq\hat{t}_+\}}{1+\#\{(i,j):(i,j)\notin\mathcal{E},\widehat{W}_{ij}\leq-\hat{t}_+\}}\\ &\qquad\cdot\frac{1+\#\{(i,j):\widehat{W}_{ij}\leq-\hat{t}_+\}}{\#\{(i,j):\widehat{W}_{ij}\geq\hat{t}_+\}\vee 1}\\ &\leq \frac{\#\{(i,j):(i,j)\notin\mathcal{E},\widehat{W}_{ij}\geq\hat{t}_+\}}{1+\#\{(i,j):(i,j)\notin\mathcal{E},\widehat{W}_{ij}\leq-\hat{t}_+\}}\cdot q\,.\end{aligned}$$

The first inequality follows from

$$\#\{(i,j):(i,j)\notin\mathcal{E},\widehat{W}_{ij}\geq\hat{t}_+\}\leq\#\{(i,j):\widehat{W}_{ij}\geq\hat{t}_+\}\,,$$

and the second inequality follows from the definition of $\hat{t}_+$. Using martingale theory, we prove in Appendix B that

$$\mathbb{E}\left[\frac{\#\{(i,j):(i,j)\notin\mathcal{E}',\widehat{W}_{ij}\geq\hat{t}_+\}}{1+\#\{(i,j):(i,j)\notin\mathcal{E}',\widehat{W}_{ij}\leq-\hat{t}_+\}}\right]\leq 1\,.$$

B Proofs

The agenda of this section is to establish proofs for Theorems 3.1 and 3.2. For this, we define the notion of swapping and study the matrix-valued test statistic $\widehat{W}\in\mathbb{R}^{p\times p}$. We write $\widehat{W}$ as $\widehat{W}(R,R^\circ)$ to emphasize that $\widehat{W}$ is a function of R and R°.

The basis for the proofs is the idea of swapping.

Definition B.1 (Swapping)**.** *Given an edge set $\mathcal{S}\subset\mathcal{V}\times\mathcal{V}$ and a matrix $M\in\mathbb{R}^{p\times p}$, we define the substitution operator* $\mathrm{Sub}_{\mathcal{S},M}:\mathbb{R}^{p\times p}\to\mathbb{R}^{p\times p}$ *as*

$$A\mapsto\mathrm{Sub}_{\mathcal{S},M}(A):=\begin{cases}M_{ij} & \text{if } (i,j)\in\mathcal{S}\\ A_{ij} & \text{if } (i,j)\notin\mathcal{S}\,.\end{cases}$$

We then define the corresponding swapped test matrix as

$$\widehat{W}_{\mathcal{S}}:=\widehat{W}\big(\mathrm{Sub}_{\mathcal{S},R^\circ}(R),\mathrm{Sub}_{\mathcal{S},R}(R^\circ)\big)\,.$$

Given an edge set $\mathcal{S}$ and a matrix M, the operator $\mathrm{Sub}_{\mathcal{S},M}(A)$ substitutes the elements of A that have indexes in $\mathcal{S}$ by the corresponding elements of M. Hence, as compared to the original test matrix $\widehat{W}$, the new test matrix $\widehat{W}_{\mathcal{S}}\equiv\widehat{W}_{\mathcal{S}}(R,R^\circ)$ has the entries of R and R° that have indexes in $\mathcal{S}$ swapped. We will see that the elements of $\widehat{W}$ and $\widehat{W}_{\mathcal{S}}$ that correspond to a zero-valued edge have the same distribution, while the distributions of other elements can differ. This gives us leverage for assessing the number of zero-valued edges in a given set $\mathcal{S}$.

The swapped test statistics still has an explicit formulation. By definition of the original test matrix in (9), we find

$$(\widehat{W}_{\mathcal{S}})_{ij}=\begin{cases}(\widehat{T}^\circ_{ij}\vee\widehat{T}_{ij})\cdot\mathrm{sign}(\widehat{T}^\circ_{ij}-\widehat{T}_{ij}) & \text{if } (i,j)\in\mathcal{S}\\ (\widehat{T}_{ij}\vee\widehat{T}^\circ_{ij})\cdot\mathrm{sign}(\widehat{T}_{ij}-\widehat{T}^\circ_{ij}) & \text{if } i\neq j \text{ and } (i,j)\notin\mathcal{S}\\ 0 & \text{if } i=j\,.\end{cases}\tag{12}$$

This means that $\widehat{W}_{\mathcal{S}}$ is an "antisymmetric" version of $\widehat{W}$:

Lemma B.1 (Antisymmetry)**.** *For every edge set* $\mathcal{S} \subset \{(k,l) \in \mathcal{V} \times \mathcal{V} : k \neq l\}$*, it holds that*

$$(\widehat{W}_{\mathcal{S}})_{ij} = \widehat{W}_{ij} \cdot \begin{cases} +1 & (i,j) \notin \mathcal{S} \\ -1 & (i,j) \in \mathcal{S}\,. \end{cases}$$

Hence, swapping two entries R_{ij}, R°_{ij} effects in switching signs in $\widehat{W}_{ij}$.

Proof of Lemma B.1. This follows directly from comparing Displays (9) and (12). □

Now, we show that the coordinates of $\widehat{W}$ and $\widehat{W}_{\mathcal{S}}$ that correspond to a zero-valued edge are equal in distribution.

Lemma B.2 (Exchangeability)**.** *For every zero-valued edge* $(i,j) \in \{(k,l) \in \mathcal{V} \times \mathcal{V} : k \neq l, \Sigma^{-1}_{kl} = 0, x_k \perp x_l\}$*, it holds that*

$$(\widehat{W}_{\mathcal{S}})_{ij} =_d \widehat{W}_{ij}\,,$$

where $\mathcal{S} \subset \{(k,l) \in \mathcal{V} \times \mathcal{V} : k \neq l\}$ *is an arbitrary set of edges and* $=_d$ *means equality in distribution.*

Proof of Lemma B.2. Our construction of the knock-offs in (4) ensures that the sample partial correlation of a zero-valued edge (i,j) and the corresponding knock-off version have the same distribution: $R^{\circ}_{ij} =_d R_{ij}$. This equality in distribution remains true under elementwise thresholding, so that also the corresponding elements of $\widehat{T}$ and $\widehat{T}^{\circ}$ in (7) and (8), respectively, are equal in distribution: $\widehat{T}_{ij} =_d \widehat{T}^{\circ}_{ij}$. This implies that $\operatorname{sign}(\widehat{T}_{ij} - \widehat{T}^{\circ}_{ij}) =_d \operatorname{sign}(\widehat{T}^{\circ}_{ij} - \widehat{T}_{ij})$ (and $\widehat{T}_{ij} \vee \widehat{T}^{\circ}_{ij} = \widehat{T}^{\circ}_{ij} \vee \widehat{T}_{ij}$ anyway). Hence, in view of the definitions of the test statistics $\widehat{W}$ and $\widehat{W}_{\mathcal{S}}$ in (9) and (12), respectively, we find $(\widehat{W}_{\mathcal{S}})_{ij} =_d \widehat{W}_{ij}$, as desired. □

We are now ready to discuss the signs of $\widehat{W}_{ij}$. The below result will be used in the proofs of Theorems 3.1 and 3.2.

Lemma B.3 (Sign-Flip)**.** *For every zero-valued edge* $(i,j) \in \{(k,l) \in \mathcal{V} \times \mathcal{V} : k \neq l, \Sigma^{-1}_{kl} = 0, x_k \perp x_l\}$*, it holds that*

$$\widehat{W}_{ij} =_d -\widehat{W}_{ij}\,.$$

This lemma justifies our previous statement that

$$\#\Big\{(i,j) : (i,j) \notin \mathcal{E}', \widehat{W}_{ij} \leq -t\Big\} =_d \#\Big\{(i,j) : (i,j) \notin \mathcal{E}', \widehat{W}_{ij} \geq t\Big\}\,.$$

Proof of Lemma B.3. Define $\mathcal{S}$ as the set that only contains the zero-valued edge in question: $\mathcal{S} := \{(i,j)\}$. Lemma B.1 then yields

$$\big(\widehat{W}_{\mathcal{S}}\big)_{ij} = \widehat{W}_{ij} \cdot (-1)\,,$$

while Lemma B.2 yields

$$(\widehat{W}_{\mathcal{S}})_{ij} =_d \widehat{W}_{ij}\,.$$

Combining these two identities concludes the proof. □

We now prove Theorems 3.1 and 3.2. For this, we start with two sequential hypothesis testing procedures, together with the theoretical results for FDR control. Then, we relate these two procedures to KO and KO+ to prove Theorems 3.1 and 3.2.

We first introduce the two selective sequential hypothesis testing procedures. Consider a sequence of null hypotheses $\mathscr{H}_1, \dots, \mathscr{H}_N$ and corresponding "p-values" $p_1, \dots, p_N$. The values $p_1, \dots, p_N$ are not necessarily p-values in a traditional sense, but we will still refer them like that, because they play the same role as p-values here; in particular, they will need to stochastically dominate a standard uniform random variable, that is, $\Pr(p_l \leq u) \leq u$ for any $u \in [0,1]$, which is a typical assumption on traditional p-values—see (Ferreira and Zwinderman, 2006, Page 1828).

We say that a p-value p_l is a null p-value if the null hypothesis $\mathscr{H}_l$ is true, and we say p_l is a non-null p-value if $\mathscr{H}_l$ is false with $l \in \{1, \ldots, N\}$.

Selective Sequential Hypothesis Testing I: For the threshold value 1/2 and any subset $\mathcal{K} \subset \{1, \ldots, N\}$, define

$$\hat{k} := \max\left\{k \in \mathcal{K} : \frac{\#\big\{l \in \{1, \ldots, k\} : p_l > 1/2\big\}}{\#\big\{l \in \{1, \ldots, k\} : p_l \leq 1/2\big\} \vee 1} \leq q\right\}. \tag{13}$$

Set $\hat{k} := 0$ if the above set is empty. We reject $\mathscr{H}_k$ for all $k \leq \hat{k}$ with $p_k \leq 1/2$. We will see that this procedure achieves the approximate FDR control. Moreover, the KO scheme can be framed as this procedure.

Selective Sequential Hypothesis Testing II: For the threshold value 1/2 and any subset $\mathcal{K} \subset \{1, 2, \ldots, N\}$, define

$$\hat{k}_+ := \max\left\{k \in \mathcal{K} : \frac{1 + \#\big\{l \in \{1, \ldots, k\} : p_l > 1/2\big\}}{\#\big\{l \in \{1, \ldots, k\} : p_l \leq 1/2\big\} \vee 1} \leq q\right\}. \tag{14}$$

Set $\hat{k}_+ := 0$ if the corresponding set is empty. We reject $\mathscr{H}_k$ for all $k \leq \hat{k}_+$ with $p_k \leq 1/2$. We will also see that this procedure achieves the exact FDR control. Moreover, the KO+ scheme can be cast as this procedure.

Our next result guarantees FDR control over the Selective Sequential Hypothesis Testing I and II.

Lemma B.4 (FDR Control Over the Hypothesis Testing I and II)**.** *Consider the two selective sequential procedures described above, and suppose that the null p-values are i.i.d., satisfy* $\Pr(p_l \leq u) \leq u$ *for any* $u \in [0, 1]$*, and are independent from the non-null p-values. Let* V, V_+ *be the numbers of false discoveries of the two procedures, that is,*

$$V := \#\big\{l \in \{1, \ldots, \hat{k}\} : p_l \textit{ is null and } p_l \leq 1/2\big\}$$
$$V_+ := \#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \textit{ is null and } p_l \leq 1/2\big\},$$

and R, R_+ *be the total number of discoveries of the two procedures, that is,*

$$R := \#\big\{l \in \{1, \ldots, \hat{k}\} : p_l \leq 1/2\big\}$$
$$R_+ := \#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \leq 1/2\big\}.$$

Define $R := V := 0$ *if* $\hat{k} = 0$*, and define* $R_+ := V_+ := 0$ *if* $\hat{k}_+ = 0$*. Then, it holds that*

$$\mathbb{E}\left[\frac{V}{R + q^{-1}}\right] \leq q \quad \textit{and} \quad \mathbb{E}\left[\frac{V_+}{R_+ \vee 1}\right] \leq q\,.$$

This lemma ensures that the Selective Sequential Hypothesis Testing I controls a quantity close to the FDR and that the Selective Sequential Hypothesis Testing II achieves exact FDR control. These guarantees will be transferred to the KO and KO+ schemes later by showing that these schemes can be formulated as Selective Sequential Hypothesis Testing I and II also.

Proof of Lemma B.4. We start with the Selective Sequential Hypothesis Testing I. The number of total discoveries is always at least as large as the number of false discoveries: $R \geq V$. Hence, $R = 0$ implies $V = 0$, and then it's easy to see that the desired inequalities hold (and are actually equalities). We can thus assume without loss of generality that $R > 0$ in the following.

Using the definition of V as the number of false discoveries, the definition of R as the total number of discoveries, and expanding the fraction, we find

$$\begin{aligned}
&\mathbb{E}\left[\frac{V}{R + q^{-1}}\right] \\
=&\mathbb{E}\Bigg[\frac{\#\big\{l \in \{1, \ldots, \hat{k}\} : p_l \text{ is null and } p_l \leq 1/2\big\}}{1 + \#\big\{1 \leq l \leq \hat{k} : p_l \text{ is null and } p_l > 1/2\big\}} \\
&\quad \cdot \frac{1 + \#\big\{l \in \{1, \ldots, \hat{k}\} : p_l \text{ is null and } p_l > 1/2\big\}}{R + q^{-1}}\Bigg].
\end{aligned}$$

The number of falsly rejected hypothesis is at most as large as the total number of rejected hypotheses

$$\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l > 1/2\big\} \le \#\big\{l \in \{1,\dots,\hat{k}\} : p_l > 1/2\big\}\,.$$

Moreover, since $R > 0$, the definition of $\hat{k}$ yields that

$$\#\big\{l \in \{1,\dots,\hat{k}\} : p_l > 1/2\big\} \le q \cdot R\,.$$

Combining these two results gives

$$\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l > 1/2\big\} \le q \cdot R\,.$$

Plugging this into the previous display and some rearranging provides us with

$$\begin{aligned}\mathbb{E}\bigg[\frac{V}{R+q^{-1}}\bigg] &\le \mathbb{E}\Bigg[\frac{\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l \le 1/2\big\}}{1+\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l > 1/2\big\}}\Bigg] \cdot \frac{1+q\cdot R}{R+q^{-1}}\\ &= \mathbb{E}\Bigg[\frac{\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l \le 1/2\big\}}{1+\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l > 1/2\big\}}\Bigg] \cdot q\,.\end{aligned}$$

Inequality (A.1) of Lemma 1 (martingale process) in the supplement to Barber and Candès (2015) gives (set $c = 1/2$)

$$\mathbb{E}\Bigg[\frac{\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l \le 1/2\big\}}{1+\#\big\{l \in \{1,\dots,\hat{k}\} : p_l \text{ is null and } p_l > 1/2\big\}}\Bigg] \le 1\,.$$

(Here, we have used the assumptions on the p-values.) Combining this with the previous display gives

$$\mathbb{E}\bigg[\frac{V}{R+q^{-1}}\bigg] \le q\,,$$

as desired.

We now prove the FDR control over Selective Sequential Hypothesis Testing II. By definitions of the total discoveries V_+ and false discoveries R_+, it holds that $V_+ = R_+ = 0$ when $\hat{k}_+ = 0$. We then find that

$$\mathbb{E}\bigg[\frac{V_+}{R_+ \vee 1} \cdot \mathbb{1}(0 = \hat{k}_+)\bigg] = 0\,,$$

which implies

$$\mathbb{E}\bigg[\frac{V_+}{R_+ \vee 1}\bigg] = \mathbb{E}\bigg[\frac{V_+}{R_+ \vee 1} \cdot \mathbb{1}(0 < \hat{k}_+)\bigg]\,.$$

Using the definitions of V_+ and R_+, and expanding the fraction gives

$$\begin{aligned}&\mathbb{E}\bigg[\frac{V_+}{R_+ \vee 1}\bigg]\\ =&\mathbb{E}\Bigg[\frac{\#\big\{l \in \{1,\dots,\hat{k}_+\} : p_l \text{ is null and } p_l \le 1/2\big\}}{1+\#\big\{l \in \{1,\dots,\hat{k}_+\} : p_l \text{ is null and } p_l > 1/2\big\}}\\ &\times \frac{1+\#\big\{l \in \{1,\dots,\hat{k}_+\} : p_l \text{ is null and } p_l > 1/2\big\}}{\#\big\{l \in \{1,\dots,\hat{k}_+\} : p_l \le 1/2\big\} \vee 1} \cdot \mathbb{1}(0 < \hat{k}_+)\Bigg]\,.\end{aligned}$$

The number of falsly rejected hypothesis is at most as large as the total number of rejected hypotheses

$$\#\big\{l \in \{1,\dots,\hat{k}_+\} : p_l \text{ is null and } p_l > 1/2\big\} \le \#\big\{l \in \{1,\dots,\hat{k}_+\} : p_l > 1/2\big\}\,.$$

Moreover, by definition of $\hat{k}_+$, it holds for $0 < \hat{k}_+$ that

$$\frac{1 + \#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l > 1/2\big\}}{\#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \leq 1/2\big\} \vee 1} \leq q\,.$$

Combining these two results gives

$$\frac{1 + \#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \text{ is null and } p_l > 1/2\big\}}{\#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \leq 1/2\big\} \vee 1} \leq q\,.$$

Plugging this into previous display and some rearranging yields

$$\mathbb{E}\left[\frac{V_+}{R_+ \vee 1}\right] \leq \mathbb{E}\left[\frac{\#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \text{ is null and } p_l \leq 1/2\big\}}{1 + \#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \text{ is null and } p_l > 1/2\big\}}\right] \cdot q\,.$$

Invoking Inequality of Lemma 1 (martingale process) in the supplement to Barber and Candès (2015) again (set $c = 1/2$), we find

$$\mathbb{E}\left[\frac{\#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \text{ is null and } p_l \leq 1/2\big\}}{1 + \#\big\{l \in \{1, \ldots, \hat{k}_+\} : p_l \text{ is null and } p_l > 1/2\big\}}\right] \leq 1\,.$$

Combining this with the previous display gives

$$\mathbb{E}\left[\frac{V_+}{R_+ \vee 1}\right] \leq q$$

as desired. □

We now show that the KO procedure is equivalent to the Selective Sequential Hypothesis Testing I, and KO+ procedure can be framed as the Selective Sequential Hypothesis Testing II. Then, the desired FDR control over KO and KO+ schemes follows directly from Lemma B.4.

Proof of Theorem 3.1 and Theorem 3.2. The proof has two steps: First, we arrange the elements of the matrix-valued statistics $\widehat{W}$ in decreasing absolute value and define "p-values" for each null hypothesis $\mathscr{H}_{(i,j)} : \Sigma^{-1}_{ij} = 0$ based on the corresponding $\widehat{W}_{ij}$. Second, we connect Selective Sequential Hypothesis Testing I and the KO scheme as well as Selective Sequential Hypothesis Testing II and the KO+ scheme and then apply Lemma B.4.

Define a set of index pairs by $\widehat{\mathcal{W}}^{\circlearrowleft} := \Big\{\widehat{W}_{ij} : (i,j) \in \mathcal{V} \times \mathcal{V}, \widehat{W}_{ij} \neq 0\Big\}$ and denote the cardinality of this set by $n^{\circlearrowleft} := \operatorname{card}(\widehat{\mathcal{W}}^{\circlearrowleft})$. Refer to the elements in $\widehat{\mathcal{W}}^{\circlearrowleft}$ by $\widehat{W}^1, \ldots, \widehat{W}^{n^{\circlearrowleft}}$ in a non-increasing order (all elements are non-zero by definition of $\widehat{\mathcal{W}}^{\circlearrowleft}$):

$$|\widehat{W}^1| \geq \cdots \geq |\widehat{W}^{n^{\circlearrowleft}}| > 0\,.$$

Define the set of indices $\mathcal{K} := \Big\{k \in \{1, \ldots, n^{\circlearrowleft} - 1\} : |\widehat{W}^k| > |\widehat{W}^{k+1}|\Big\} \bigcup \{n^{\circlearrowleft}\}$. We notice that $\mathcal{K}$ is the index set of unique non-zero values attained by $|\widehat{W}^l|, l \in \{1, \ldots, n^{\circlearrowleft}\}$.

Define corresponding p-values p_l, where $l \in \{1, \ldots, n^{\circlearrowleft}\}$, based on the test statistic $\widehat{W}^l$:

$$p_l := \begin{cases} \frac{1}{2} & \widehat{W}^l > 0 \\ 1 & \widehat{W}^l < 0\,. \end{cases}$$

By Lemma B.3 (sign-flip), $\widehat{W}_{ij}$ is positive and negative equally likely for all zero-valued edges $(i,j) \in \{(k,l) \in \mathcal{V} \times \mathcal{V} : k \neq l, \Sigma^{-1}_{kl} = 0\}$, that is,

$$\Pr\Big(\widehat{W}_{ij} > 0\Big) = \Pr\Big(\widehat{W}_{ij} < 0\Big) = \frac{1}{2}\,.$$

Combining this with the definition of the p-value p_l, it holds that for any null p-value p_l that

$$\Pr\biggl(p_l = \frac{1}{2}\biggr) = \Pr(p_l = 1) = \frac{1}{2}\,,$$

which implies $\Pr(p_l \leq u) \leq u$ for all $u \in [0,1]$. By Lemma B.3, we find the null p-values are i.i.d., satisfy $\Pr(p_l \leq u) \leq u$ for any $u \in [0,1]$, and are independent from the non-null p-values. By definition of the p-value p_l, it holds for any $k \in \mathcal{K}$ that

$$\#\bigl\{l \in \{1,\dots,k\} : p_l > 1/2\bigr\} = \#\Bigl\{l \in \{1,\dots,k\} : \widehat{W}^l < 0\Bigr\}\,.$$

Due to the assumed ordering $|\widehat{W}^1| \geq \cdots \geq |\widehat{W}^{n^{\circlearrowleft}}| > 0$, we have

$$-|\widehat{W}^1| \leq \cdots \leq -|\widehat{W}^{n^{\circlearrowleft}}| < 0\,.$$

So, it holds for any $\widehat{W}^l < 0$ that

$$-|\widehat{W}^1| \leq \cdots \leq -|\widehat{W}^{l-1}| \leq \widehat{W}^l \leq -|\widehat{W}^{l+1}| \leq \cdots \leq -|\widehat{W}^{n^{\circlearrowleft}}|\,,$$

which implies

$$\#\Bigl\{l \in \{1,\dots,k\} : \widehat{W}^l < 0\Bigr\} = \#\Bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \leq -|\widehat{W}^k|\Bigr\}\,.$$

Combining this with the previous display yields

$$\#\bigl\{l \in \{1,\dots,k\} : p_l > 1/2\bigr\} = \#\Bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \leq -|\widehat{W}^k|\Bigr\}\,. \tag{15}$$

By the same arguments, we obtain

$$\#\bigl\{l \in \{1,\dots,k\} : p_l \leq 1/2\bigr\} = \#\Bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \geq |\widehat{W}^k|\Bigr\}\,. \tag{16}$$

Plugging these two displays together, we find

$$\frac{\#\bigl\{l \in \{1,\dots,k\} : p_l > 1/2\bigr\}}{\#\bigl\{l \in \{1,\dots,k\} : p_l \leq 1/2\bigr\} \vee 1} = \frac{\#\Bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \leq -|\widehat{W}^k|\Bigr\}}{\#\Bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \geq |\widehat{W}^k|\Bigr\} \vee 1}\,.$$

Finding the largest $k \in \mathcal{K}$ such that the ratio on the left-hand side is below q is—in view of the non-increasing ordering of the $|\widehat{W}^k|$'s—equivalent to finding the smallest $|\widehat{W}^k|$ over $k \in \mathcal{K}$ such that the right-hand side is below q. By definition of the threshold value $\hat{k}$ of Selective Sequential Hypothesis Testing I in Display (13), this means that

$$\hat{k} = \max\Biggl\{k \in \mathcal{K} : \frac{\#\bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \leq -|\widehat{W}^k|\bigr\}}{\#\bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \geq |\widehat{W}^k|\bigr\} \vee 1} \leq q\Biggr\}\,.$$

Comparing to the definition of the KO threshold in Display (10), we find that $\widehat{W}^{\hat{k}}$ is equal to $\hat{t}$. This equality implies that the KO scheme is equivalent to the Selective Sequential Hypothesis Testing I, which gives us the desired FDR control.

Plugging (15) and (16) together, it also holds for $k \in \mathcal{K}$ that

$$\frac{1 + \#\bigl\{l \in \{1,\dots,k\} : p_l > 1/2\bigr\}}{\#\bigl\{l \in \{1,\dots,k\} : p_l \leq 1/2\bigr\} \vee 1} = \frac{1 + \#\Bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \leq -|\widehat{W}^k|\Bigr\}}{\#\Bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \geq |\widehat{W}^k|\Bigr\} \vee 1}\,.$$

By the definition of the threshold value $\hat{k}_+$ of the Selective Sequential Hypothesis Testing II in Display (14), this means that

$$\hat{k}_+ = \max\Biggl\{k \in \mathcal{K} : \frac{1 + \#\bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \leq -|\widehat{W}^k|\bigr\}}{\#\bigl\{l \in \{1,\dots,n^{\circlearrowleft}\} : \widehat{W}^l \geq |\widehat{W}^k|\bigr\} \vee 1} \leq q\Biggr\}\,.$$

Comparing to the definition of the KO+ threshold in Display (11), we find that $\widehat{W}^{\hat{k}_+}$ is equal to $\hat{t}_+$. This equality implies that the KO scheme is equivalent to the Selective Sequential Hypothesis Testing II. The desired FDR control of KO+ scheme follows from Lemma B.4.

□