Supplement to "Designing Transportable Experiments Under S-admissability"

In Section A we discuss the variance reduction for $d \geq 1$ when the sample size is finite. In Section B we show the proofs of Section 5.1. In Section C we show the proofs of Section 5.2.1. In Section D we show the proofs of Section 5.2.2. In Section E we show the proofs of Appendix A.

For a random variable R with value r, we write the expectation, variance and covariance conditioning on r as a short-hand for conditioning on $R = r$. On the other hand, the expectation, variance and covariance conditioning on R are functions of R and therefore are random variables. For example,$\mathbb{E}[\hat{\tau}_Y^T|\mathbf{X}, \mathbf{Y}]$ is a function of $\mathbf{X}$ and $\mathbf{Y}$, $\mathbb{E}[\hat{\tau}_Y^T|\mathbf{X}, \mathbf{y}] = \mathbb{E}[\hat{\tau}_Y^T|\mathbf{X}, \mathbf{Y} = \mathbf{y}]$ is a function of $\mathbf{X}$, while $\mathbb{E}[\hat{\tau}_Y^T|\mathbf{x}, \mathbf{y}] = \mathbb{E}[\hat{\tau}_Y^T|\mathbf{X} = \mathbf{x}, \mathbf{Y} = \mathbf{y}]$ is a value.

Conditioning on $\mathbf{x}$ and $\mathbf{y}$, the randomness only comes from $\mathbf{Z}$. Therefore $\text{var}_{\mathbf{Z}_\rho}(.|\mathbf{x}, \mathbf{y}), Cov_{\mathbf{Z}_\rho}(.|\mathbf{x}, \mathbf{y})$ and $\mathbb{E}_{\mathbf{Z}_\rho}(.|\mathbf{x}, \mathbf{y})$ can be written as $\text{var}_{\mathbf{Z}}(.|\mathbf{x}, \mathbf{y}, \rho = 1), Cov_{\mathbf{Z}}(.|\mathbf{x}, \mathbf{y}, \rho = 1)$ and $\mathbb{E}_{\mathbf{Z}}(.|\mathbf{x}, \mathbf{y}, \rho = 1)$ respectively. We use both notations in the proofs.

For a random variable R, we use $Cov(R)^{-1/2}$ to denote the Cholesky square root of $Cov(R)^{-1}$.

We restate the model and some notations here for convenience. Let the model be:

$$Y_i^1 = X_i^T \beta_1 + \mathcal{E}_i^1 \qquad Y_i^0 = X_i^T \beta_0 + \mathcal{E}_i^0$$

Let ϵ_i^1 and ϵ_i^0 be the values taken by random variables $\mathcal{E}_i^1$ and $\mathcal{E}_i^0$. Let $C_i = \frac{Y_i^0 + Y_i^1}{2}$, $\tilde{C}_i = W_i C_i$, $\mathbf{C} := (C_1, \cdots, C_n)$ and $\tilde{\mathbf{C}} = (\tilde{C}_1, \cdots, \tilde{C}_n)$. Let $c_i, \tilde{c}_i, \mathbf{c}$ and $\tilde{\mathbf{c}}$ be the values taken by $C_i, \tilde{C}_i, \mathbf{C}$ and $\tilde{\mathbf{C}}$. Then

$$C_i = X_i^T \beta + \mathcal{E}_i \qquad c_i = x_i^T \beta + \epsilon_i$$
$$\tilde{C}_i = \tilde{X}_i^T \beta + \tilde{\mathcal{E}}_i \qquad \tilde{c}_i = \tilde{x}_i^T \beta + \tilde{\epsilon}_i$$

where $\beta = \frac{\beta_1 + \beta_0}{2}, \mathcal{E}_i = \frac{\mathcal{E}_i^1 + \mathcal{E}_i^0}{2}, \tilde{X}_i = W_i X_i$ and $\tilde{\mathcal{E}}_i = W_i \mathcal{E}_i$. Let ϵ_i and $\tilde{\epsilon}_i = w_i \epsilon_i$ be the value taken by $\mathcal{E}_i$ and $\tilde{\mathcal{E}}_i$. Let $\tilde{\boldsymbol{\mathcal{E}}} = (\tilde{\mathcal{E}}_1, \cdots, \tilde{\mathcal{E}}_n)$.

A Additional Results: Finite Sample Size Variance Reduction for $d \geq 1$

In this section we discuss the finite sample case when X is a multivariate random variable, which is a generalization of the result in Section 5.2.1 when $d = 1$. We show that when the sample size is finite, if β points to all directions with equal probability, then a balance condition which also consider the target population and is similar to Target Balance achieves the optimal variance reduction in expectation over β. The proofs are in Appendix E.

We will use the variance decomposition in the matrix form similar to [Harshaw et al., 2019] and provide intuition about the effect of balancing on the variance. The following lemma is the general case when $d \geq 1$ of Lemma 3 in Section 5.2.1.

Lemma A.1. *For any function $\rho(\mathbf{x}, \mathbf{Z}) \in \{0, 1\}$ satisfying $\rho(\mathbf{x}, \mathbf{Z}) = \rho(\mathbf{x}, -\mathbf{Z})$:*

$$var^S_{\mathbf{Y}, \mathbf{Z}_\rho}(\hat{\tau}_Y^T|\mathbf{x}) = \beta^T Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})\beta + \frac{6}{n^2}\sigma_{\mathcal{E}}^2 \sum_{i=1}^n w_i^2,$$

for $V := \frac{2}{n}(\mathbf{w} \cdot \mathbf{x})^T \mathbf{Z} = \frac{2}{n}\tilde{\mathbf{x}}^T \mathbf{Z}$.

Since the design affects only the first term in the above expression, we focus on the the random variable V. V is now a d-dimensional vector and β is unknown.

To understand the first term, we use the same decomposition of $\beta^T Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})\beta$ as in [Harshaw et al., 2019]. Let $\mathbf{e}_1, ..., \mathbf{e}_n$ and $\lambda_1, .., \lambda_n$ be the normalized eigenvectors and corresponding eigenvalues of matrix $Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})$. Since $Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})$ is symmetric, the eigenvectors form an orthonormal basis so we can write β as a linear combination of $\mathbf{e}_1, .., \mathbf{e}_n$ and get:

$$\beta = \|\beta\| \sum_{i=1}^n \eta_i \mathbf{e}_i$$

where $\eta_i = \langle \beta, \mathbf{e}_i \rangle / \|\beta\|$ is the coefficient that captures the alignment of the weighted outcome β with respect to the eigenvector $\mathbf{e}_i$. Therefore:

$$\beta^T Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})\beta = \|\beta\|^2 \sum_{i=1}^{n} \eta_i^2 \lambda_i$$

In the worst case, β can align with the eigenvector of $Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})$ with the largest eigenvalue. Therefore a good design is one with ρ that minimize the largest eigenvalue of $Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})$. We leave this for future works. In this work we consider the average case direction - when β with norm $\|\beta\| = l$ can point in any direction with equal probability. In that case, we have

Lemma A.2.

$$\mathbb{E}_{\|\beta\|=l} \beta^T Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})\beta = \frac{l^2}{2} Trace(Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})). \tag{4}$$

We can then ask for the balance event Ω which results in minimizing the trace of $Cov_{\mathbf{Z}}(V|\mathbf{x}, \Omega)$, which is shown in the following lemma. Note that when $d = 1$, the trace of $Cov_{\mathbf{Z}}(V|\mathbf{x}, \Omega)$ is the variance $\text{var}_{\mathbf{Z}}(V|\mathbf{x}, \Omega)$, and this result is the general case of minimizing the variance of a 1-dimensional random variable in Section 5.2.1.

Lemma A.3. *Let $U \in \mathcal{R}^d$ be a random variable such that $\mathbb{E}[U] = 0$. Let u_α be such that $\mathbb{P}(\|U\|^2 < u_\alpha) = 1 - \alpha$. Let Ω be an event such that $\mathbb{P}(\Omega) \geq 1 - \alpha$ and $\mathbb{E}[U|\Omega] = 0$. Then:*

$$Trace(Cov(U| \|U\|^2 < u_\alpha) \leq Trace(Cov(U|\Omega))$$

It follows from Lemma A.1, Lemma A.2 and Lemma A.3 that we can minimize $\mathbb{E}_\beta \text{var}^S_{\mathbf{Y},\mathbf{Z}}(\hat{\tau}^T_Y|\mathbf{x}, \Omega)$ by defining the following balance condition:

Definition 2 (Alternate Target Balance)**.** *With a rejection threshold α, define the balance condition*

$$\phi'^{\alpha}_T = \begin{cases} 1, & \text{if } \|V\|^2 < a \\ 0, & \text{otherwise} \end{cases}$$

where a be such that $\mathbb{P}(\phi'^{\alpha}_T = 1|\mathbf{x}) = 1 - \alpha$.

Recall that Target Balance use the condition $\|B\|^2 < a$ where $B = V Cov_{\mathbf{Z}}(V)^{-1/2}$ is the normalized random variable of V. Note since that $V = \frac{2}{n}\tilde{\mathbf{x}}^T\mathbf{Z}$, Alternate Target Balance also considers the target population in the design phase. However Alternate Target Balance is not invariant under linear transformations of the covariates x_i's while Target Balance is.

We have the following Theorem which is a generalization of Theorem 2 in Section 5.2.1.

Theorem A.1. *Let $\|\beta\| = l$ and β points in any direction with equal probability and $n_0 = n_1 = n/2$. Let $\rho(\mathbf{X}, \mathbf{Z})$ be a function satisfying $\rho(\mathbf{X}, \mathbf{Z}) = \rho(\mathbf{X}, -\mathbf{Z})$ and $\mathbb{P}(\rho = 1|\mathbf{x}) \geq 1 - \alpha$. Then*

$$\mathbb{E}_\beta var^S_{\mathbf{Y},\mathbf{Z}_{\phi'^{\alpha}_T}}(\hat{\tau}^T_Y|\mathbf{x}) \leq \mathbb{E}_\beta var^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{x})$$

Similar to Section 5.2.1, applying Theorem 2 with ρ being the constant function $\rho(\mathbf{x}, \mathbf{Z}) = 1$ for all $\mathbf{x}, \mathbf{Z}$, we have:

Corollary A.1. *Let $\|\beta\| = l$ and β points in any direction with equal probability. When $n_0 = n_1 = n/2$, using Alternate Target Balance reduces the variance compared to complete randomization in expectation over β.*

$$\mathbb{E}_\beta var^S_{\mathbf{Z}_{\phi'_T},\mathbf{Y}}(\hat{\tau}^T_Y|\mathbf{x}) \leq \mathbb{E}_\beta var^S_{\mathbf{Z},\mathbf{Y}}(\hat{\tau}^T_Y|\mathbf{x})$$

Recall that the first term in the decomposition in Lemma A.1 is equal to:

$$\beta^T Cov_{\mathbf{Z}_\rho}(V|\mathbf{x})\beta = \gamma^T Cov_{\mathbf{Z}_\rho}(B|\mathbf{x})\gamma = \gamma^T Cov_{\mathbf{Z}}(B|\mathbf{x}, \rho = 1)\gamma$$

where $\gamma = \beta^T Cov_{\mathbf{Z}}(V)^{1/2}$ and $B = VCov_{\mathbf{Z}}(V)^{-1/2}$.

When the sample size is large, B converges to a standard normal distribution. Recall that Target Balance is equal to truncating $\|B\|^2 < a$. So $Cov_{\mathbf{Z}_{\phi_T}}(B|\mathbf{x})$ is the covariance of a standard normal random variable B truncated by $\|B\|^2 < a$. From Theorem 3.1 in [Morgan et al., 2012] when B is a standard normal distribution, $Cov(B|\mathbf{x}, \phi_T = 1) = vCov(B|\mathbf{x})$ for some $v < 1$, so the variance is reduced. However we do not need to go through this analysis because [Li et al., 2018] already has variance reduction results for the case when the sample size is large. In Section 5.2.2 we use the result from [Li et al., 2018] directly to show that Target Balance achieves a smaller variance than Source Balance.

B Proofs of Section 5.1

In this section we prove Theorem 1. We made use of the following lemma from Morgan et al. [2012]:

Lemma B.1 (from the proof of Theorem 2.1 in Morgan et al. [2012])**.** *Let* $\mathbf{A} := (A_1, ..., A_n)^T \in \mathcal{R}^n$. *Let* $n_1 = n_0 = n/2$. *For any function* $\rho(\mathbf{x}, \mathbf{A}) \in \{0, 1\}$ *satisfying* $\rho(\mathbf{x}, \mathbf{A}) = \rho(\mathbf{x}, 1 - \mathbf{A})$*:*

$$\mathbb{E}^S_{\mathbf{A}}[A_i|\mathbf{x}, \mathbf{y}, \rho = 1] = \frac{1}{2}$$

We also prove the following lemma in order to prove Theorem 1:

Lemma B.2. *For any function* $\rho(\mathbf{x}, \mathbf{A}) \in \{0, 1\}$ *satisfying* $\rho(\mathbf{x}, \mathbf{A}) = \rho(\mathbf{x}, 1 - \mathbf{A})$*:*

$$\mathbb{E}_{\mathbf{A}|\rho=1}[\hat{\tau}^T_Y|\mathbf{X}, \mathbf{Y}] = \frac{1}{n}\sum_{i=1}^n W_i(Y_i^1 - Y_i^0)$$

$$\mathbb{E}^S_{\mathbf{Y},\mathbf{A}|\rho=1}[\hat{\tau}^T_Y|\mathbf{X}] = \frac{1}{n}\sum_{i=1}^n W_i(\beta_1 - \beta_0)^T X_i$$

Proof. From Lemma B.1, $\mathbb{E}[A_i|\mathbf{X}, \mathbf{Y}, \rho = 1] = \mathbb{E}[A_i|\mathbf{X}, \rho = 1] = \frac{1}{2}$. Therefore:

$$\begin{aligned}
\mathbb{E}_{\mathbf{A}|\rho=1}[\hat{\tau}^T_Y|\mathbf{X}, \mathbf{Y}] &= \frac{1}{n_1}\sum_{i=1}^n \mathbb{E}_{\mathbf{A}}\left[W_i A_i Y_i^1 \middle| \mathbf{X}, \mathbf{Y}, \rho = 1\right] - \frac{1}{n_0}\sum_{i=1}^n \mathbb{E}_{\mathbf{A}}\left[W_i(1 - A_i)Y_i^0 \middle| \mathbf{X}, \mathbf{Y}, \rho = 1\right] \\
&= \frac{1}{n_1}\sum_{i=1}^n W_i Y_i^1 \mathbb{E}_{\mathbf{A}}\left[A_i \middle| \mathbf{X}, \mathbf{Y}, \rho = 1\right] - \frac{1}{n_0}\sum_{i=1}^n W_i Y_i^0 \mathbb{E}_{\mathbf{A}}\left[1 - A_i \middle| \mathbf{X}, \mathbf{Y}, \rho = 1\right] \\
&= \frac{1}{n}\sum_{i=1}^n W_i(Y_i^1 - Y_i^0) \\
\mathbb{E}^S_{\mathbf{A}|\rho=1,\mathbf{Y}}[\hat{\tau}^T_Y|\mathbf{X}] &= \mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{A}}[\hat{\tau}^T_Y|\mathbf{X}, \mathbf{Y}, \rho = 1]|\mathbf{X}\right] \\
&= \mathbb{E}^S_{\mathbf{Y}}\left[\frac{1}{n}\sum_{i=1}^n W_i(Y_i^1 - Y_i^0)|\mathbf{X}\right] \\
&= \frac{1}{n}\sum_{i=1}^n W_i(\beta_1 - \beta_0)^T X_i
\end{aligned}$$

□

Proof of Theorem 1. Let D_S and D_T be the supports of the source and target distributions. Since $p_T(X) > 0 \rightarrow$

$p_S(X) > 0$ and $p_T(Y|X) = p_S(Y|X)$, we have $D_T \subseteq D_S$. Using Lemma B.2:

$$\begin{aligned}
\mathbb{E}^S_{\mathbf{X},\mathbf{Y},\mathbf{Z}_{\phi_T}}\left[\hat{\tau}^T_Y\right] =& \mathbb{E}^S_{\mathbf{X},\mathbf{Y}}\mathbb{E}_{\mathbf{A}_{\phi_T}}[\hat{\tau}^T_Y|\mathbf{X},\mathbf{Y}] \\
=& \frac{1}{n}\sum_{i=1}^n \mathbb{E}^S_{\mathbf{X},\mathbf{Y}}\left[W_i(Y_i^1 - Y_i^0)\right] \\
=& \frac{1}{n}\sum_{i=1}^n \int_{(x,y)\in D_S} \left(\frac{p_T(x)}{p_S(x)}(y^1 - y^0)\right) p_S(x,y)dxy \\
=& \frac{1}{n}\sum_{i=1}^n \int_{(x,y)\in D_S} \left(\frac{p_T(y|x)p_T(x)}{p_S(y|x)p_S(x)}(y^1 - y^0)\right) p_S(x,y)dxy \text{ because } p_T(y|x) = p_S(y|x) \\
=& \frac{1}{n}\sum_{i=1}^n \int_{(x,y)\in D_S} \left(\frac{p_T(y,x)}{p_S(y,x)}(y^1 - y^0)\right) p_S(x,y)dxy \\
=& \frac{1}{n}\sum_{i=1}^n \int_{(x,y)\in D_S} p_T(x,y)(y^1 - y^0)dxy \\
=& \frac{1}{n}\sum_{i=1}^n \int_{(x,y)\in D_T} p_T(x,y)(y^1 - y^0)dxy \text{ because } D_T \subseteq D_S \\
=& \tau^T_Y
\end{aligned}$$

□

C Proofs of Section 5.2.1

In this section we prove Lemma 1, Lemma 2, Lemma 3, Theorem 2 and Corollary 1. Note that the results in this section are the special case when $d = 1$ of the results in Section A. Lemma 2 is a special case when $d = 1$ of Lemma E.1. Lemma 3 is a special case of Lemma A.1 and Theorem 2 is a special case of Theorem A.1. However in this section we state the full proofs for the case $d = 1$ so that the readers do not need to read the proofs of Section A in order to understand Section 5.2.1 in the main paper.

Proof of Lemma 1. By law of total variance:

$$\text{var}^S_{\mathbf{Z}_\rho,\mathbf{X},\mathbf{Y}}(\hat{\tau}^T_Y) = \mathbb{E}^S_{\mathbf{X}}\text{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{X}) + \text{var}^S_{\mathbf{X}}\left(\mathbb{E}^S_{\mathbf{Y},\mathbf{Z}_\rho}[\hat{\tau}^T_Y|\mathbf{X}]\right)$$

Since $\rho(\mathbf{x},\mathbf{Z}) = \rho(\mathbf{x},-\mathbf{Z})$, from Lemma B.2:

$$\mathbb{E}^S_{\mathbf{Y},\mathbf{Z}_\rho}[\hat{\tau}^T_Y|\mathbf{X}] = \frac{1}{n}\sum_{i=1}^n W_i(\beta_1 - \beta_0)^T X_i$$

Therefore:

$$\text{var}^S_{\mathbf{X}}\left(\mathbb{E}^S_{\mathbf{Y},\mathbf{Z}_\rho}[\hat{\tau}^T_Y|\mathbf{X}]\right) = \text{var}^S_{\mathbf{X}}\left(\frac{1}{n}\sum_{i=1}^n W_i(\beta_1 - \beta_0)^T X_i\right)$$

□

Proof of Lemma 2. By definition:

$$\text{var}_{\mathbf{Z}}(\hat{\tau}^T_Y|\mathbf{x},\mathbf{y},\rho = 1) = \mathbb{E}_{\mathbf{Z}}\left[(\hat{\tau}^T_Y - \mathbb{E}_{\mathbf{Z}}[\hat{\tau}^T_Y|\mathbf{x},\mathbf{y},\rho = 1])^2|\mathbf{x},\mathbf{y},\rho = 1\right]$$

From Lemma B.2

$$\mathbb{E}_{\mathbf{Z}}[\hat{\tau}^T_Y|\mathbf{x},\mathbf{y},\rho = 1] = \frac{1}{n}\left(\sum_{i=1}^n w_i y_i^1 - \sum_{i=1}^n w_i y_i^0\right)$$

On the other hand conditioning on $\mathbf{X} = \mathbf{x}$ and $\mathbf{Y} = \mathbf{y}$ and let y_i^* denote the observed outcome of sample i:

$$\begin{aligned}\hat{\tau}_Y^T &= \frac{2}{n}(\sum_{Z_i=1} w_i y_i^* - \sum_{Z_i=-1} w_i y_i^*) \\ &= \frac{2}{n}\sum_{i=1}^n w_i A_i y_i^1 - \frac{2}{n}\sum_{i=1}^n w_i(1-A_i)y_i^0\end{aligned}$$

Therefore:

$$\begin{aligned}\text{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{y},\rho=1) &= \mathbb{E}_{\mathbf{Z}}\left[\left(\frac{2}{n}(\sum_{i=1}^n w_i A_i y_i^1 - \sum_{i=1}^n w_i(1-A_i)y_i^0) - \frac{1}{n}\sum_{i=1}^n w_i(y_i^1 - y_i^0)\right)^2 \middle| \mathbf{x},\mathbf{y},\rho=1\right] \\ &= \mathbb{E}_{\mathbf{Z}}\left[\left(\frac{1}{n}(\sum_{i=1}^n w_i(2A_i-1)y_i^1 + \frac{1}{n}\sum_{i=1}^n w_i(2A_i-1)y_i^0)\right)^2 \middle| \mathbf{x},\mathbf{y},\rho=1\right] \\ &= \frac{4}{n^2}\mathbb{E}_{\mathbf{Z}}\left[\left(\sum_{i=1}^n w_i Z_i \frac{y_i^1+y_i^0}{2}\right)^2 \middle| \mathbf{x},\mathbf{y},\rho=1\right] \\ &= \frac{4}{n^2}\mathbb{E}_{\mathbf{Z}}\left[\left(\sum_{i=1}^n Z_i w_i c_i\right)^2 \middle| \mathbf{x},\mathbf{y},\rho=1\right]\end{aligned}$$

where $c_i = \frac{y_i^1+y_i^0}{2}$. ∎

Proof of Lemma 3. By law of total variance:

$$\begin{aligned}\text{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}_Y^T|\mathbf{x}) &= \mathbb{E}^S_{\mathbf{Y}}\left[\text{var}_{\mathbf{Z}_\rho}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y})|\mathbf{x}\right] + \text{var}^S_{\mathbf{Y}}(\mathbb{E}_{\mathbf{Z}_\rho}[\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y}]|\mathbf{x}) \\ &= \mathbb{E}^S_{\mathbf{Y}}\left[\text{var}_{\mathbf{Z}_\rho}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y})|\mathbf{x}\right] + \text{var}^S_{\mathbf{Y}}\left(\frac{1}{n}\sum_{i=1}^n w_i(Y_i^1 - Y_i^0)|\mathbf{x}\right) \\ &= \mathbb{E}^S_{\mathbf{Y}}\left[\text{var}_{\mathbf{Z}_\rho}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y})|\mathbf{x}\right] + \frac{1}{n^2}\sum_{i=1}^n w_i^2 \text{var}(\mathcal{E}_i^1 - \mathcal{E}_i^0) \\ &= \mathbb{E}^S_{\mathbf{Y}}\left[\text{var}_{\mathbf{Z}_\rho}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y})|\mathbf{x}\right] + \frac{2}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2\end{aligned}$$

Recall that $\tilde{C}_i = \beta\tilde{X}_i + \tilde{\mathcal{E}}_i$. From Lemma 2:

$$\begin{aligned}&\text{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1) \\ &= \frac{4}{n^2}\mathbb{E}_{\mathbf{Z}}\left[\left(\sum_{i=1}^n Z_i\tilde{C}_i\right)^2 \middle| \mathbf{x},\mathbf{Y},\rho=1\right] \\ &= \frac{4}{n^2}\mathbb{E}_{\mathbf{Z}}\left[\left(\mathbf{Z}^T\tilde{\mathbf{C}}\right)^2 \middle| \mathbf{x},\mathbf{Y},\rho=1\right] \\ &= \frac{4}{n^2}\mathbb{E}_{\mathbf{Z}}\left[\left(\mathbf{Z}^T\beta\tilde{\mathbf{x}} + \mathbf{Z}^T\tilde{\boldsymbol{\mathcal{E}}}\right)^2 \middle| \mathbf{x},\mathbf{Y},\rho=1\right] \\ &= \frac{4}{n^2}\beta^2\mathbb{E}_{\mathbf{Z}}\left[(\mathbf{Z}^T\tilde{\mathbf{x}})^2 \middle| \mathbf{x},\rho=1\right] + \frac{4}{n^2}\mathbb{E}_{\mathbf{Z}}\left[\left(\mathbf{Z}^T\tilde{\boldsymbol{\mathcal{E}}}\right)^2 \middle| \mathbf{x},\mathbf{Y},\rho=1\right] + \frac{4}{n^2}2\mathbb{E}_{\mathbf{Z}}\left[\tilde{\mathbf{x}}^T\mathbf{Z}\mathbf{Z}^T\tilde{\boldsymbol{\mathcal{E}}} \middle| \mathbf{x},\mathbf{Y},\rho=1\right] \quad (5)\end{aligned}$$

Now we consider $\mathbb{E}^S_{\mathbf{Y}}\left[\text{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right]$. The third term in Eq. 5 becomes:

$$\begin{aligned}\frac{4}{n^2}2\mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{Z}}\left[\tilde{\mathbf{x}}^T\mathbf{Z}\mathbf{Z}^T\tilde{\boldsymbol{\mathcal{E}}}\middle|\mathbf{x},\mathbf{Y},\rho=1\right]\middle|\mathbf{x}\right] &= \frac{8}{n^2}\mathbb{E}_{\mathbf{Z}}\left[\tilde{\mathbf{x}}^T\mathbf{Z}\mathbf{Z}^T\middle|\mathbf{x},\rho=1\right]\mathbb{E}^S_{\mathbf{Y}}[\tilde{\boldsymbol{\mathcal{E}}}|\mathbf{x}] \\ &= 0 \text{ because } \mathbb{E}^S_{\mathbf{Y}}[\tilde{\boldsymbol{\mathcal{E}}}|\mathbf{x}] = \mathbf{0}\end{aligned}$$

The second term in Eq. 5 becomes:

$$
\begin{aligned}
&\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{Z}}\left[\left(\mathbf{Z}^T\tilde{\boldsymbol{\mathcal{E}}}\right)^2\middle|\mathbf{x},\mathbf{Y},\rho=1\right]\middle|\mathbf{x}\right]\\
&=\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{Z}}\left[\left(\sum_{i=1}^n Z_i w_i \mathcal{E}_i\right)^2|\mathbf{x},\mathbf{Y},\rho=1\right]\middle|\mathbf{x}\right]\\
&=\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{Z}}\left[\sum_{i=1}^n (Z_i w_i \mathcal{E}_i)^2\big|\mathbf{x},\mathbf{Y},\rho=1\right]\middle|\mathbf{x}\right]+\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{Z}}\left[\sum_{i\neq j}(Z_i w_i \mathcal{E}_i)(Z_j w_j \mathcal{E}_j)\big|\mathbf{x},\mathbf{Y},\rho=1\right]\middle|\mathbf{x}\right]\\
&=\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{Z}}\left[\sum_{i=1}^n (Z_i w_i \mathcal{E}_i)^2\big|\mathbf{x},\mathbf{Y},\rho=1\right]\middle|\mathbf{x}\right]+\frac{4}{n^2}\sum_{i\neq j}\mathbb{E}_{\mathbf{Z}}[Z_i Z_j|\mathbf{x},\rho=1]w_i w_j \mathbb{E}^S_{\mathbf{Y}}\left[\mathcal{E}_i\mathcal{E}_j|\mathbf{x}\right]\\
&=\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\mathbb{E}_{\mathbf{Z}}\left[\sum_{i=1}^n (Z_i w_i \mathcal{E}_i)^2\big|\mathbf{x},\mathbf{Y},\rho=1\right]\middle|\mathbf{x}\right]+0 \text{ because } \mathbb{E}^S_{\mathbf{Y}}[\mathcal{E}_i\mathcal{E}_j|\mathbf{x}]=\mathbb{E}^S_{\mathbf{Y}}[\mathcal{E}_i|\mathbf{x}]\mathbb{E}^S_{\mathbf{Y}}[\mathcal{E}_j|\mathbf{x}]=0\\
&=\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\sum_{i=1}^n (w_i\mathcal{E}_i)^2]\big|\mathbf{x}\right] \text{ because } Z_i^2=1\\
&=\frac{4}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2
\end{aligned}
$$

The first term in Eq. 5 becomes:

$$
\begin{aligned}
\frac{4}{n^2}\mathbb{E}^S_{\mathbf{Y}}\left[\beta^2\mathbb{E}_{\mathbf{Z}}\left[\left(\mathbf{Z}^T\tilde{\mathbf{x}}\right)^2\middle|\mathbf{x},\rho=1\right]\middle|\mathbf{x}\right]&=\frac{4}{n^2}\beta^2\mathbb{E}_{\mathbf{Z}}\left[\left(\mathbf{Z}^T\tilde{\mathbf{x}}\right)^2\middle|\mathbf{x},\rho=1\right]\\
&=\frac{4}{n^2}\beta^2\mathbb{E}_{\mathbf{Z}}\left[(\sum_{i=1}^n Z_i w_i x_i)^2\big|\mathbf{x},\rho=1\right]
\end{aligned}
$$

Putting all 3 terms together:

$$
\mathbb{E}^S_{\mathbf{Y}}\left[\text{var}_{\mathbf{Z}}(\hat{\tau}^T_Y|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right]=\frac{4}{n^2}\beta^2\mathbb{E}_{\mathbf{Z}}\left[(\sum_{i=1}^n Z_i w_i x_i)^2\big|\mathbf{x},\rho=1\right]+\frac{4}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2
$$

Therefore:

$$
\begin{aligned}
\text{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{x})&=\mathbb{E}^S_{\mathbf{Y}}\left[\text{var}_{\mathbf{Z}}(\hat{\tau}^T_Y|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right]+\frac{2}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2\\
&=\frac{4}{n^2}\beta^2\mathbb{E}_{\mathbf{Z}}\left[(\sum_{i=1}^n Z_i w_i x_i)^2\big|\mathbf{x},\rho=1\right]+\frac{6}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2
\end{aligned}
$$

□

In order to prove Theorem 2, we will show that for a random variable U with $\mathbb{E}[U]=0$, among events Ω preserve the expectation $\mathbb{E}[U|\Omega]=0$, truncating the tail results in the smallest variance. Note that if $\rho(\mathbf{x},\mathbf{Z})=\rho(\mathbf{x},-\mathbf{Z})$ it follows from Lemma B.1 that $\mathbb{E}[\frac{2}{n}\tilde{\mathbf{x}}^T\mathbf{Z}|\rho=1]=\mathbb{E}[\frac{2}{n}\tilde{\mathbf{x}}^T\mathbf{Z}]=0$.

In order to prove Theorem 1 we show how to minimize the variance of a random variable:

Lemma C.1. *Let $U\in\mathcal{R}$ be a random variable such that $\mathbb{E}[U]=0$. Let u_α be such that $\mathbb{P}(U^2<u_\alpha)=1-\alpha$. Let Ω be an event such that $\mathbb{P}(\Omega)\geq 1-\alpha$ and $\mathbb{E}[U|\Omega]=0$. Then:*

$$
\mathbb{E}(U^2|U^2<u_\alpha)\leq\mathbb{E}(U^2|\Omega)
$$

Proof. Let $p(u)$ be the pdf of U. Define $f(u)$ as follow:

$$
f(u)=p(U=u,\Omega)
$$

then:

$$p(u|\Omega) = \frac{p(U = u, \Omega)}{\mathbb{P}(\Omega)} = \frac{f(u)}{1-\alpha} \ .$$

Therefore:

$$\mathbb{E}[U^2|\Omega] = \int_u u^2 \frac{f(u)}{1-\alpha} du \ .$$

We want to minimize $\mathbb{E}(U^2|\Omega)$:

$$\int_u u^2 \frac{f(u)}{1-\alpha} du$$

subject to:

$$0 \leq f(u) \leq p(u) \ \forall u$$
$$\mathbb{P}(\Omega) = \int_u f(u) du = 1 - \alpha$$

This can be done by maximize $f(u)$ so that $f(u) = p(u)$ for the smallest u^2, which is equal to set Ω to be the event $U^2 < u_\alpha$. □

Proof of Theorem 2. Let $V := \frac{2}{n}\sum_i w_i x_i Z_i$ and $B = V \text{var}(V)^{-1/2}$. From Lemma 3:

$$\begin{aligned}\text{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{x}) &= \beta^2 \mathbb{E}_\mathbf{Z}\left[V^2 \middle| \mathbf{x}, \rho = 1\right] + \frac{6}{n^2}\sigma^2_\mathcal{E}\sum_{i=1}^n w_i^2. \\ &= \beta^2 \text{var}(V)\mathbb{E}_\mathbf{Z}\left[B^2 \middle| \mathbf{x}, \rho = 1\right] + \frac{6}{n^2}\sigma^2_\mathcal{E}\sum_{i=1}^n w_i^2.\end{aligned}$$

Since $\rho(\mathbf{x}, \mathbf{Z}) = \rho(\mathbf{x}, -\mathbf{Z})$, from Lemma B.1 we have $\mathbb{E}_\mathbf{Z}[B|\mathbf{x}, \rho = 1] = 0$, which satisfies the criteria in Lemma C.1. Let $\eta := 1 - \mathbb{P}(\rho = 1|\mathbf{x})$. Then $\eta \leq \alpha$. Let b_η be such that $\mathbb{P}(B^2 < b_\eta|\mathbf{x}) = 1 - \eta$ and b_α be such that $\mathbb{P}(B^2 < b_\alpha|\mathbf{x}) = 1 - \alpha$. From Lemma C.1:

$$\begin{aligned}\mathbb{E}_\mathbf{Z}\left[B^2 \middle| \mathbf{x}, \rho = 1\right] &\geq \mathbb{E}_\mathbf{Z}\left[B^2 \middle| \mathbf{x}, B^2 < b_\eta\right] \\ &\geq \mathbb{E}_\mathbf{Z}\left[B^2 \middle| \mathbf{x}, B^2 < b_\alpha\right] \text{ because } b_\eta \geq b_\alpha \\ &\geq \mathbb{E}_\mathbf{Z}\left[B^2 \middle| \mathbf{x}, \phi^\alpha_T = 1\right]\end{aligned}$$

□

Proof of Corollary 1. Let ρ being the constant function $\rho(\mathbf{x}, \mathbf{Z}) = 1$ for all $\mathbf{x}, \mathbf{Z}$. Then:

$$\text{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{x}) = \text{var}^S_{\mathbf{Y},\mathbf{Z}}(\hat{\tau}^T_Y|\mathbf{x})$$

From Theorem 2 we have:

$$\text{var}^S_{\mathbf{Y},\mathbf{Z}_{\phi^\alpha_T}}(\hat{\tau}^T_Y|\mathbf{x}) \leq \text{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{x}) = \text{var}^S_{\mathbf{Y},\mathbf{Z}}(\hat{\tau}^T_Y|\mathbf{x})$$

□

D Discussion on Section 5.2.2

Proof of Lemma 4. By law of total variance:

$$\text{var}^S_{\mathbf{X},\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y) = \mathbb{E}^S_{\mathbf{X},\mathbf{Y}}\left[\text{var}_{\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{X}, \mathbf{Y})\right] + \text{var}^S_{\mathbf{X},\mathbf{Y}}\left(\mathbb{E}_{\mathbf{Z}_\rho}[\hat{\tau}^T_Y|\mathbf{X}, \mathbf{Y}]\right)$$

Since $\rho(\mathbf{x}, \mathbf{Z}) = \rho(\mathbf{x}, -\mathbf{Z})$, from Lemma B.2:

$$\mathbb{E}_{\mathbf{Z}}[\hat{\tau}_Y^T | \mathbf{X}, \mathbf{Y}, \rho = 1] = \frac{1}{n} \sum_{i=1}^{n} W_i (Y_i^1 - Y_i^0)$$

Therefore:

$$\text{var}_{\mathbf{X},\mathbf{Y}}^S \left(\mathbb{E}_{\mathbf{Z}}[\hat{\tau}_Y^T | \mathbf{X}, \mathbf{Y}, \rho = 1] \right) = \text{var}_{\mathbf{X},\mathbf{Y}}^S \left(\frac{1}{n} \sum_{i=1}^{n} W_i (Y_i^1 - Y_i^0) \right)$$

□

We now prove Lemma 5. We use the following result in Harshaw et al. [2019] to prove Lemma 5.

Lemma D.1 (Lemma A1 in Harshaw et al. [2019])**.** *Let y_i^* denote the observed outcome of sample i:*

$$\frac{2}{n} (\sum_{z_i = 1} y_i^* - \sum_{z_i = -1} y_i^*) - \frac{1}{n} \sum_{i=1}^{n} (y_i^1 - y_i^0) = \frac{2}{n} \mathbf{c}^T \mathbf{z}$$

where $c_i = \frac{y_i^1 + y_i^0}{2}$ and $\mathbf{c} := (c_1, \cdots, c_n)$.

We will also use the following lemma:

Lemma D.2. *Let $Q := \frac{n-1}{n} \mathbb{E}[\mathbf{Z}\mathbf{Z}^T]$. Let $\mathbf{I}_n$ denote the $n \times n$ identity matrix and $\mathbf{1}$ denote the n dimensional vector of 1. Then:*

$$\begin{aligned} Q &= \mathbf{I}_n - \frac{1}{n} \mathbf{1}\mathbf{1}^T. \\ Q &= Q^T \\ Q &= Q^2 = Q^T Q = Q Q^T. \end{aligned}$$

Let $\mathbf{s} \in \mathcal{R}^{n \times d}$ be a matrix. Then

$$Q\mathbf{s} = \mathbf{s} - \mathit{avg}(\mathbf{s})$$

where $\mathit{avg}(\mathbf{s}) \in \mathcal{R}^d$ is the average of rows of $\mathbf{s}$.

Proof. First we will show that:

$$\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] = \frac{n}{n-1} \left(\mathbf{I}_n - \frac{1}{n} \mathbf{1}\mathbf{1}^T \right)$$

by showing that $\mathbb{E}[Z_i^2] = 1$ and $\mathbb{E}[Z_i Z_j] = -\frac{1}{n-1}$ when $i \neq j$. First we have that $\mathbb{E}[Z_i^2] = 1$ because $Z_i^2 = 1$. Since there are exactly $n/2$ samples with value $Z_i = 1$ and $n/2$ samples with values $Z_i = -1$, note that $(\sum_{i=1}^n Z_i)^2 = 0$ and:

$$\mathbb{E}[(\sum_{i=1}^{n} Z_i)^2] = \mathbb{E}[\sum_{i=1}^{n} Z_i^2] + \sum_{i \neq j} \mathbb{E}[Z_i Z_j] \; .$$

Since all pairs (i, j) where $i \neq j$ have equal roles and there are $n(n-1)$ such pairs:

$$\begin{aligned} \mathbb{E}[Z_i Z_j] &= \frac{\mathbb{E}[(\sum_{i=1}^n Z_i)^2] - \mathbb{E}[\sum_{i=1}^n Z_i^2]}{n(n-1)} \\ &= \frac{0 - n}{n(n-1)} \\ &= \frac{-1}{n-1} \end{aligned}$$

Since Q is symmetric, $Q = Q^T$. We will show that $Q = Q^2$:

$$\begin{aligned} Q^2 &= (\mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T)(\mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T) \\ &= \mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbf{I}_n - \frac{1}{n}\mathbf{I}_n\mathbf{1}\mathbf{1}^T + \frac{1}{n^2}\mathbf{1}\mathbf{1}^T\mathbf{1}\mathbf{1}^T \\ &= \mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T = Q \end{aligned}$$

Since $Q = Q^T$, we have $Q = Q^2 = QQ^T = Q^TQ$. For the last property:

$$Q\mathbf{s} = \mathbf{I}_n\mathbf{s} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbf{s} = \mathbf{s} - \text{avg}(\mathbf{s})$$

because $\mathbf{I}_n\mathbf{s} = \mathbf{s}$ and $\frac{1}{n}\mathbf{1}\mathbf{1}^T\mathbf{s} = \text{avg}(\mathbf{s})$ □

Proof of Lemma 5. For any matrix $\mathbf{s} \in \mathcal{R}^{n\times d}$ we will compute $R^2_{\mathbf{s}} := Corr(\hat{\tau}_Y^T, \frac{2}{n}\mathbf{Z}^T\mathbf{s})$ where for any $Y \in \mathcal{R}, X \in \mathcal{R}^d$, $Corr(Y, X)$ is defined as:

$$\begin{aligned} Corr(Y, X) &= Corr(Y, X^T\beta^{*)} \\ &= \frac{Cov(Y, X^T\beta^*)}{\sqrt{\text{var}(Y)}\sqrt{\text{var}(X^T\beta^*)}} \end{aligned}$$

where $\beta^* = \arg\min_{\hat{\beta}} \mathbb{E}\|Y - X^T\hat{\beta}\|^2$. Substituting $\mathbf{s} = \mathbf{x}$ and $\mathbf{s} = \tilde{\mathbf{x}}$ will give us $R^2_{\mathbf{x}}$ and $R^2_{\tilde{\mathbf{x}}}$.

Let $\tilde{\delta}_i = \tilde{y}_i^1 - \tilde{y}_i^0$ and $\tilde{\boldsymbol{\delta}} := (\tilde{\delta}_1, \cdots, \tilde{\delta}_n)$. From Lemma D.1, we have:

$$\hat{\tau}_Y^T = \frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{c}} + \frac{1}{n}\mathbf{1}^T\tilde{\boldsymbol{\delta}}$$

where $\mathbf{1} \in \mathcal{R}^n$ is a vector of 1.

We note that conditioning on $\mathbf{y}$, $\mathbf{1}^T\tilde{\boldsymbol{\delta}}$ is a constant independent of $\mathbf{Z}$. Let $Q := \frac{n-1}{n}\mathbb{E}[\mathbf{Z}\mathbf{Z}^T]$ and note that $Q = Q^T$ and $Q = Q^2$. First, let us compute $\beta^* = \arg\min_{\hat{\beta}} \mathbb{E}_{\mathbf{Z}}\|\hat{\tau}_Y^T - \frac{2}{n}\mathbf{Z}^T\mathbf{s}\hat{\beta}\|^2$. We have,

$$\begin{aligned} \beta^* &= \arg\min_{\hat{\beta}} \mathbb{E}_{\mathbf{Z}}\|\hat{\tau}_Y^T - \frac{2}{n}\mathbf{Z}^T\mathbf{s}\hat{\beta}\|^2 \\ &= \arg\min_{\hat{\beta}} \mathbb{E}_{\mathbf{Z}}\|\frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{c}} + \frac{1}{n}\mathbf{1}^T\tilde{\boldsymbol{\delta}} - \frac{2}{n}\mathbf{Z}^T\mathbf{s}\hat{\beta}\|^2 \\ &= \arg\min_{\hat{\beta}} \mathbb{E}_{\mathbf{Z}}\|\mathbf{Z}^T\tilde{\mathbf{c}} - \mathbf{Z}^T\mathbf{s}\hat{\beta}\|^2 \\ &= \arg\min_{\hat{\beta}}(\tilde{\mathbf{c}} - \mathbf{s}\hat{\beta})^T\mathbb{E}[\mathbf{Z}\mathbf{Z}^T](\tilde{\mathbf{c}} - \mathbf{s}\hat{\beta}) \\ &= \arg\min_{\hat{\beta}}(\tilde{\mathbf{c}} - \mathbf{s}\hat{\beta})^TQ(\tilde{\mathbf{c}} - \mathbf{s}\hat{\beta}) \\ &= \arg\min_{\hat{\beta}}(\tilde{\mathbf{c}} - \mathbf{s}\hat{\beta})^TQ^TQ(\tilde{\mathbf{c}} - \mathbf{s}\hat{\beta}) \\ &= \arg\min_{\hat{\beta}} \|Q\tilde{\mathbf{c}} - Q\mathbf{s}\beta\|^2. \end{aligned}$$

Using the fact that $Q = Q^T Q$, we have $\beta^* = (\mathbf{s}^T Q \mathbf{s})^{-1}\mathbf{s}^T Q\tilde{\mathbf{c}}$. By definition, we have

$$
\begin{aligned}
Corr(\hat{\tau}_Y^T, \frac{2}{n}\mathbf{Z}^T\mathbf{s}) &= \frac{\mathbb{E}_{\mathbf{Z}}\left[\hat{\tau}_Y^T \frac{2}{n}\mathbf{Z}^T\mathbf{s}\beta^*\right] - \mathbb{E}_{\mathbf{Z}}\left[\hat{\tau}_Y^T\right]\mathbb{E}_{\mathbf{Z}}\left[\frac{2}{n}\mathbf{Z}^T\mathbf{s}\beta^*\right]}{\sqrt{\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T)\mathrm{var}_{\mathbf{Z}}(\frac{2}{n}\mathbf{Z}^T\mathbf{s}\beta^*)}} \\
&= \frac{\mathbb{E}_{\mathbf{Z}}\left[\hat{\tau}_Y^T \mathbf{Z}^T\mathbf{s}\beta^*\right]}{\sqrt{\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T)\mathrm{var}_{\mathbf{Z}}(\mathbf{Z}^T\mathbf{s}\beta^*)}} \text{ because } \mathbb{E}[\mathbf{Z}] = 0 \\
&= \frac{\mathbb{E}_{\mathbf{Z}}\left[\left(\frac{2}{n}\tilde{\mathbf{c}}^T\mathbf{Z} + \frac{1}{n}\mathbf{1}^T\tilde{\boldsymbol{\delta}}\right)\mathbf{Z}^T\mathbf{s}\beta^*\right]}{\sqrt{\mathrm{var}_{\mathbf{Z}}\left(\frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{c}} + \frac{1}{n}\mathbf{1}^T\tilde{\boldsymbol{\delta}}\right)\mathrm{var}_{\mathbf{Z}}(\mathbf{Z}^T\mathbf{s}\beta^*)}} \\
&= \frac{\mathbb{E}_{\mathbf{Z}}\left[\left(\frac{2}{n}\tilde{\mathbf{c}}^T\mathbf{Z}\right)\mathbf{Z}^T\mathbf{s}\beta^*\right]}{\sqrt{\mathrm{var}_{\mathbf{Z}}\left(\frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{c}}\right)\mathrm{var}_{\mathbf{Z}}(\mathbf{Z}^T\mathbf{s}\beta^*)}} \\
&= \frac{\mathbb{E}_{\mathbf{Z}}\left[\tilde{\mathbf{c}}^T\mathbf{Z}\mathbf{Z}^T\mathbf{s}\beta^*\right]}{\sqrt{\mathrm{var}_{\mathbf{Z}}\left(\mathbf{Z}^T\tilde{\mathbf{c}}\right)\mathrm{var}_{\mathbf{Z}}(\mathbf{Z}^T\mathbf{s}\beta^*)}}
\end{aligned}
$$

For the numerator we have:

$$
\begin{aligned}
\mathbb{E}_{\mathbf{Z}}\left[\tilde{\mathbf{c}}^T\mathbf{Z}\mathbf{Z}^T\mathbf{s}\beta^*\right] &= \tilde{\mathbf{c}}^T Q\mathbf{s}\beta^* \\
&= \frac{n}{n-1}\tilde{\mathbf{c}}^T Q\mathbf{s}(\mathbf{s}^T Q\mathbf{s})^{-1}\mathbf{s}^T Q\tilde{\mathbf{c}} \\
&= \frac{n}{n-1}\tilde{\mathbf{c}}^T Q\mathbf{s}(\mathbf{s}^T Q\mathbf{s})^{-1}\mathbf{s}^T Q\mathbf{s}(\mathbf{s}^T Q\mathbf{s})^{-1}\mathbf{s}^T Q\tilde{\mathbf{c}} \\
&= \frac{n}{n-1}\left(\tilde{\mathbf{c}}^T Q\mathbf{s}(\mathbf{s}^T Q\mathbf{s})^{-1}\mathbf{s}^T Q\right)\left(Q\mathbf{s}(\mathbf{s}^T Q\mathbf{s})^{-1}\mathbf{s}^T Q\tilde{\mathbf{c}}\right) \\
&= \frac{n}{n-1}(\beta^{*T}\mathbf{s}^T Q)(Q\mathbf{s}\beta^*) \\
&= \frac{n}{n-1}\|Q\mathbf{s}\beta^*\|^2
\end{aligned}
$$

Let $u = Q\mathbf{s}\beta$ and $v = Q\tilde{\mathbf{c}} - Q\mathbf{s}\beta$. We will show that u and v are orthogonal, therefore $\|Q\mathbf{s}\beta^*\|^2 = \|Q\tilde{\mathbf{c}}\|^2 - \|Q\tilde{\mathbf{c}} - Q\mathbf{s}\beta\|^2$:

$$
\begin{aligned}
u^T v &= (Q\tilde{\mathbf{c}} - Q\mathbf{s}\beta^*)^T(Q\mathbf{s}\beta^*) \\
&= \tilde{\mathbf{c}}^T Q\mathbf{s}\beta^* - \beta^{*T}\mathbf{s}^T Q\mathbf{s}\beta^* \\
&= \tilde{\mathbf{c}}^T Q\mathbf{s}\beta^* - \|Q\mathbf{s}\beta^*\|^2 \\
&= 0\,.
\end{aligned}
$$

Therefore $\|Q\mathbf{s}\beta^*\|^2 = \|Q\tilde{\mathbf{c}}\|^2 - \|Q\tilde{\mathbf{c}} - Q\mathbf{s}\beta\|^2$.

For the denominator, since $\mathbb{E}[\mathbf{Z}] = 0$ we have:

$$
\begin{aligned}
\mathrm{var}_{\mathbf{Z}}\left(\mathbf{Z}^T\tilde{\mathbf{c}}\right)\mathrm{var}_{\mathbf{Z}}(\mathbf{Z}^T\mathbf{s}\beta^*) &= \mathbb{E}_{\mathbf{Z}}[\tilde{\mathbf{c}}^T\mathbf{Z}\mathbf{Z}^T\tilde{\mathbf{c}}]\mathbb{E}_{\mathbf{Z}}[\beta^{*T}\mathbf{s}^T\mathbf{Z}\mathbf{Z}^T\mathbf{s}\beta^*] \\
&= \frac{n^2}{(n-1)^2}(\tilde{\mathbf{c}}^T Q\tilde{\mathbf{c}})(\beta^{*T}\mathbf{s}^T Q\mathbf{s}\beta^*) \\
&= \frac{n^2}{(n-1)^2}(\tilde{\mathbf{c}}^T Q^T Q\tilde{\mathbf{c}})(\beta^{*T}\mathbf{s}^T Q^T Q\mathbf{s}\beta^*) \\
&= \frac{n^2}{(n-1)^2}\|Q\tilde{\mathbf{c}}\|^2\|Q\mathbf{s}\beta^*\|^2
\end{aligned}
$$

Putting the numerator and denominator together we have:

$$\begin{aligned} R^2_{\mathbf{s}} &= Corr(\hat{\tau}_Y^T, \frac{2}{n}\mathbf{Z}^T\mathbf{s}) \\ &= \frac{\|Q\mathbf{s}\beta^*\|^2}{\|Q\tilde{\mathbf{c}}\|\|Q\mathbf{s}\beta^*\|} \\ &= \frac{\|Q\mathbf{s}\beta^*\|}{\|Q\tilde{\mathbf{c}}\|} \\ &= \frac{\sqrt{\|Q\tilde{\mathbf{c}}\|^2 - \|Q\tilde{\mathbf{c}} - Q\mathbf{s}\beta\|^2}}{\|Q\tilde{\mathbf{c}}\|} \end{aligned}$$

Substituting $\mathbf{s} = \mathbf{x}$ and $\mathbf{s} = \tilde{\mathbf{x}}$ gives us the expression for $R^2_{\mathbf{x}}$ and $R^2_{\tilde{\mathbf{x}}}$. □

Proof of Theorem 4. We have

$$\tilde{C} = \tilde{X}^T\beta + \tilde{\mathcal{E}}$$

where $C = \frac{Y^0+Y^1}{2}$, $\mathcal{E} = \frac{\mathcal{E}_0+\mathcal{E}_1}{2}$, $\beta = \frac{\beta_0+\beta_1}{2}$, $\tilde{C} = \frac{p_T(X)}{p_S(X)}C, \tilde{X} = \frac{p_T(X)}{p_S(X)}X$ and $\tilde{\mathcal{E}} = \frac{p_T(X)}{p_S(X)}\mathcal{E}$. Since Y_i, X_i and W_i have finite 8th moment, $\tilde{C}_i$ and $\tilde{X}_i$ have finite 4th moment using Cauchy-Schwartz inequality. Let $S \in \mathcal{R}^d$ be a random variable independent of $\mathcal{E}_i$ and with finite 4th moment. Let $\mathbf{S} \in \mathcal{R}^{n\times d}$ be n samples $S_1, \cdots, S_n$ of S. By the definition of R^2,

$$R^2_{\mathbf{S}} = \frac{\|Q\tilde{\mathbf{C}}\|^2 - \min_{\hat{\beta}} \|Q\tilde{\mathbf{C}} - Q\mathbf{S}\hat{\beta}\|^2}{\|Q\tilde{\mathbf{C}}\|^2}.$$

We will show that $\lim_{n\to\infty} R^2_{\tilde{\mathbf{X}}} \geq \lim_{n\to\infty} R^2_{\mathbf{S}}$ almost surely for any S. It is sufficient to show $\lim \min_{\hat{\beta}} \|Q\tilde{\mathbf{C}} - Q\tilde{\mathbf{X}}\hat{\beta}\|^2 \leq \lim \min_{\hat{\beta}} \|Q\tilde{\mathbf{C}} - Q\mathbf{S}\hat{\beta}\|^2$ almost surely. From Lemma D.2, note that for any matrix $\mathbf{s} \in \mathcal{R}^{n\times d}$ with n rows, $\frac{n-1}{n}Q\mathbf{s} = \mathbf{s} - \text{avg}(\mathbf{s})$ where $\text{avg}(\mathbf{s}) \in \mathcal{R}^d$ is the average of rows of $\mathbf{s}$. Let $\beta^* = \arg\min_{\hat{\beta}} \lim_{n\to\infty} \frac{1}{n}\|Q\tilde{\mathbf{C}} - Q\mathbf{S}\hat{\beta}\|^2$ and $\tilde{\beta} = \arg\min_{\hat{\beta}} \frac{1}{n}\|Q\tilde{\mathbf{C}} - Q\mathbf{S}\hat{\beta}\|^2$. If S_i and $\tilde{C}_i$ have finite 4th moment, by strong law of large number

$\lim_{n\to\infty} \tilde{\beta} = \beta^*$ almost surely. We have:

$$\begin{aligned}
&\frac{1}{n}\lim_{n\to\infty}\min_{\hat{\beta}} \|Q\tilde{\mathbf{C}} - Q\mathbf{S}\hat{\beta}\|^2 \\
&= \frac{1}{n}\lim_{n\to\infty} \|Q\tilde{\mathbf{C}} - Q\mathbf{S}\beta^*\|^2 + 2(Q\tilde{\mathbf{C}} - Q\mathbf{S}\beta^*)^T(Q\mathbf{S}\beta^* - Q\mathbf{S}\tilde{\beta}) + \left\|Q\mathbf{S}\beta^* - Q\mathbf{S}\tilde{\beta}\right\|^2 \\
&= \frac{1}{n}\lim_{n\to\infty} \|Q\tilde{\mathbf{C}} - Q\mathbf{S}\beta^*\|^2 + 2\lim_{n\to\infty}(Q\tilde{\mathbf{C}} - Q\mathbf{S}\beta^*)^T Q\mathbf{S}\lim_{n\to\infty}(\beta^* - \tilde{\beta}) + \lim_{n\to\infty}(\beta^* - \tilde{\beta})^T \lim_{n\to\infty}\mathbf{S}^T Q\mathbf{S}\lim_{n\to\infty}(\beta^* - \tilde{\beta}) \\
&\text{because } \tilde{C}_i \text{ and } S_i \text{ having finite 4th moment implies } \lim_{n\to\infty}(Q\tilde{\mathbf{C}} - Q\mathbf{S}\beta^*)^T Q\mathbf{S} \text{ and } \lim_{n\to\infty}\mathbf{S}^T Q\mathbf{S} \text{ are finite} \\
&= \lim_{n\to\infty}\frac{1}{n}\|Q\tilde{\mathbf{C}} - Q\mathbf{S}\beta^*\|^2 \text{ almost surely} \\
&= \min_{\hat{\beta}}\lim_{n\to\infty}\frac{1}{n}\|Q\tilde{\mathbf{C}} - Q\mathbf{S}\hat{\beta}\|^2 \\
&= \min_{\hat{\beta}}\lim_{n\to\infty}\frac{1}{n}\frac{n^2}{(n-1)^2}\left\|\frac{n-1}{n}Q\tilde{\mathbf{C}} - \frac{n-1}{n}Q\mathbf{S}\hat{\beta}\right\|^2 \\
&= \min_{\hat{\beta}}\lim_{n\to\infty}\frac{1}{n}\|(\tilde{\mathbf{C}} - \mathbf{S}\hat{\beta}) - (\text{avg}(\tilde{\mathbf{C}}) - \text{avg}(\mathbf{S})\hat{\beta})\|^2 \\
&= \min_{\hat{\beta}} \text{var}(\tilde{C} - S^T\hat{\beta}) \text{ almost surely if } \tilde{C}_i \text{ and } S_i \text{ have finite 4th moment} \\
&= \min_{\hat{\beta}} \mathbb{E}[(\tilde{C} - S^T\hat{\beta})^2] - \left(\mathbb{E}[\tilde{C} - S^T\hat{\beta}]\right)^2 \\
&= \min_{\hat{\beta}} \mathbb{E}[\tilde{X}^T\beta - S^T\hat{\beta}]^2 + \mathbb{E}[\tilde{\mathcal{E}}^2] - \left(\mathbb{E}[\tilde{X}^T\beta - S^T\hat{\beta}]\right)^2 \text{ because } \mathbb{E}[\tilde{\mathcal{E}}] = 0 \text{ and } \mathcal{E} \text{ is independent of } \tilde{X} \text{ and } S \\
&= \min_{\hat{\beta}} \text{var}(\tilde{X}^T\beta - S^T\hat{\beta}) + \mathbb{E}[\tilde{\mathcal{E}}^2] \geq \mathbb{E}[\tilde{\mathcal{E}}^2]
\end{aligned}$$

When $S = \tilde{X}$, this is minimized, therefore:

$$\lim_{n\to\infty} R^2_{\tilde{\mathbf{X}}} \geq \lim_{n\to\infty} R^2_{\mathbf{S}} \text{ almost surely.}$$

Substituting $\mathbf{S} = \mathbf{X}$:

$$\lim_{n\to\infty} R^2_{\tilde{\mathbf{X}}} \geq \lim_{n\to\infty} R^2_{\mathbf{X}} \text{ almost surely.}$$

Recall that:

$$\text{as-var}_{\mathbf{Z}}\left(\hat{\tau}_Y^T|\mathbf{x},\mathbf{y}, M\left(\frac{2}{n}\mathbf{Z}^T\mathbf{s}\right) \leq a\right) = \lim_{n\to\infty}\text{var}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{y})(1 - (1 - v_{d,a})R^2_{\mathbf{s}}),$$

where as-var is the variance of the asymptotic sampling distribution. Let $s(a)$ denote the rejection probability $\mathbb{P}(\phi_S = 0|\mathbf{x}) = 0$ when using threshold a in Source Balance, and $t(a)$ denote the rejection probability $\mathbb{P}(\phi_T = 0|\mathbf{x}) = 0$ when using threshold a in Target Balance. We have:

$$\begin{aligned}
\text{as-var}_{\mathbf{Z}}\left(\hat{\tau}_Y^T|\mathbf{X},\mathbf{Y},\phi_S^{s(a)} = 1\right) &= \text{as-var}_{\mathbf{Z}}\left(\hat{\tau}_Y^T|\mathbf{X},\mathbf{Y}, M\left(\frac{2}{n}\mathbf{Z}^T\mathbf{X}\right) \leq a\right) \\
&= \lim_{n\to\infty}\text{var}(\hat{\tau}_Y^T|\mathbf{X},\mathbf{Y})(1 - (1 - v_{d,a})R^2_{\mathbf{X}}) \\
&\geq \lim_{n\to\infty}\text{var}(\hat{\tau}_Y^T|\mathbf{X},\mathbf{Y})(1 - (1 - v_{d,a})R^2_{\tilde{\mathbf{X}}}) \text{ almost surely} \\
&= \text{as-var}_{\mathbf{Z}}\left(\hat{\tau}_Y^T|\mathbf{X},\mathbf{Y}, M\left(\frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{X}}\right) \leq a\right) \\
&= \text{as-var}_{\mathbf{Z}}\left(\hat{\tau}_Y^T|\mathbf{X},\mathbf{Y},\phi_T^{t(a)} = 1\right)
\end{aligned}$$

Now we will show that for any $\mathbf{x}$ and $\tilde{\mathbf{x}}$, $\lim_{n\to\infty} s(a) = \lim_{n\to\infty} t(a)$. Let $U \in \mathcal{R}^d$ be a standard multivariate random variable. We have:

$$\begin{aligned}
\lim_{n\to\infty} s(a) &= \lim_{n\to\infty} \mathbb{P}(M\left(\frac{2}{n}\mathbf{Z}^T\mathbf{x}\right) \leq a) \\
&= \lim_{n\to\infty} \mathbb{P}(\|B_S\|^2 < a) \text{ where } B_S = \frac{2}{n}\mathbf{Z}^T\mathbf{x}Cov(\frac{2}{n}\mathbf{Z}^T\mathbf{x})^{-1/2} \\
&= \mathbb{P}(\|U\|^2 < a) \text{ because } B_S \text{ converges in distribution to } U \text{ by finite central limit theorem}
\end{aligned}$$

Similarly we have:

$$\begin{aligned}
\lim_{n\to\infty} t(a) &= \lim_{n\to\infty} \mathbb{P}(M\left(\frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{x}}\right) \leq a) \\
&= \lim_{n\to\infty} \mathbb{P}(\|B_T\|^2 < a) \text{ where } B_T := \frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{x}}Cov(\frac{2}{n}\mathbf{Z}^T\tilde{\mathbf{x}})^{-1/2} \\
&= \mathbb{P}(\|U\|^2 < a) \text{ because } B_T \text{ converges in distribution to } U \text{ by finite central limit theorem}
\end{aligned}$$

Therefore $\lim_{n\to\infty} t(a) = \lim_{n\to\infty} s(a)$. When the sample size is large, with the same rejection probability, using Target Balance results in a smaller asymptotic variance than Source Balance . □

E Proofs of Section A

In this Section we present the proof of Lemma A.1, Lemma A.2, Lemma A.3, Theorem A.1 and Corollary A.1.

In order to prove Lemma A.1, we first prove the following lemma.

Lemma E.1 (minor changes to Lemma 1 in [Harshaw et al., 2019]). *Let $\tilde{\epsilon}_i = \tilde{c}_i - \beta^T\tilde{x}_i$ and $\tilde{\boldsymbol{\epsilon}} = (\tilde{\epsilon}_1, \cdots, \tilde{\epsilon}_n)$. For any function $\rho(\mathbf{x}, \mathbf{Z}) \in \{0, 1\}$ satisfying $\rho(\mathbf{x}, \mathbf{Z}) = \rho(\mathbf{x}, -\mathbf{Z})$:*

$$\begin{aligned}
\frac{n^2}{4} var_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x}, \mathbf{y}, \rho = 1) &= Cov(\tilde{\mathbf{c}}^T\mathbf{Z}|\rho = 1) &(6)\\
&= \beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}|\rho = 1]\beta + Cov(\tilde{\boldsymbol{\epsilon}}^T\mathbf{Z}|\rho = 1) + 2\beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}, \tilde{\boldsymbol{\epsilon}}^T\mathbf{Z}|\rho = 1) &(7)
\end{aligned}$$

Proof of Lemma E.1. By definition:

$$\text{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x}, \mathbf{y}, \rho = 1) = \mathbb{E}_{\mathbf{Z}}\left[(\hat{\tau}_Y^T - \mathbb{E}_{\mathbf{Z}}[\hat{\tau}_Y^T|\mathbf{x}, \mathbf{y}, \rho = 1])^2|\mathbf{x}, \mathbf{y}, \rho = 1\right]$$

We have:

$$\begin{aligned}
\mathbb{E}_{\mathbf{Z}}[\hat{\tau}_Y^T|\mathbf{x}, \mathbf{y}, \rho = 1] &= \frac{2}{n}\mathbb{E}_{\mathbf{Z}}\left[\sum_{Z_i=1} w_i y_i^* - \sum_{Z_i=-1} w_i y_i^* \middle| \rho = 1\right] \\
&= \frac{2}{n}\mathbb{E}\left[\sum_{i=1}^n A_i w_i y_i^1 - \sum_{i=1}^n (1 - A_i) w_i y_i^0 \middle| \rho = 1\right] \\
&= \frac{2}{n}\left(\sum_{i=1}^n \mathbb{E}[A_i|\rho = 1] w_i y_i^1 - \sum_{i=1}^n \mathbb{E}[1 - A_i|\rho = 1] w_i y_i^0\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n w_i y_i^1 - \sum_{i=1}^n w_i y_i^0\right) \text{ because } \mathbb{E}[A_i|\rho = 1] = 1/2 \text{ by Lemma B.1}
\end{aligned}$$

Therefore using Lemma D.1:

$$
\begin{aligned}
\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{y},\rho=1) &= \mathbb{E}_{\mathbf{Z}}\left[\left(\frac{2}{n}(\sum_{Z_i=1} w_i y_i^* - \sum_{Z_i=-1} w_i y_i^*) - \frac{1}{n}\sum_{i=1}^n w_i(y_i^1 - y_i^0)\right)^2 \middle| \mathbf{x},\mathbf{y},\rho=1\right] \\
&= \frac{4}{n^2}\mathbb{E}[\tilde{\mathbf{c}}^T\mathbf{Z}\mathbf{Z}^T\tilde{\mathbf{c}}|\mathbf{x},\mathbf{y},\rho=1] \\
&= \frac{4}{n^2}Cov(\tilde{\mathbf{c}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1) \text{ because } \mathbb{E}[\tilde{\mathbf{c}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1] = 0 \text{ from Lemma B.1} \\
&= \frac{4}{n^2}Cov((\tilde{\mathbf{x}}\beta+\tilde{\boldsymbol{\epsilon}})^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1) \\
&= \beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1]\beta + Cov(\tilde{\boldsymbol{\epsilon}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1) + 2\beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z},\tilde{\boldsymbol{\epsilon}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1)
\end{aligned}
$$

□

Proof of Lemma A.1. By law of total variance:

$$
\begin{aligned}
\mathrm{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}_Y^T|\mathbf{x}) &= \mathbb{E}^S_{\mathbf{Y}}\left[\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right] + \mathrm{var}^S_{\mathbf{Y}}(\mathbb{E}_{\mathbf{Z}}[\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1]|\mathbf{x}) \\
&= \mathbb{E}^S_{\mathbf{Y}}\left[\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right] + \mathrm{var}^S_{\mathbf{Y}}\left(\frac{1}{n}\sum_{i=1}^n w_i(Y_i^1 - Y_i^0)|\mathbf{x}\right) \\
&= \mathbb{E}^S_{\mathbf{Y}}\left[\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right] + \frac{1}{n^2}\sum_{i=1}^n w_i^2\mathrm{var}(\mathcal{E}_1 - \mathcal{E}_0) \\
&= \mathbb{E}^S_{\mathbf{Y}}\left[\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right] + \frac{2}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2
\end{aligned}
$$

From Lemma E.1:

$$
\begin{aligned}
\frac{n^2}{4}\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{y},\rho=1) &= \beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1]\beta + Cov(\tilde{\boldsymbol{\epsilon}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1) + 2\beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z},\tilde{\boldsymbol{\epsilon}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1) \\
&= \beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1]\beta + \tilde{\boldsymbol{\epsilon}}^T Cov(\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1)\tilde{\boldsymbol{\epsilon}} + 2\beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z},\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1)\tilde{\boldsymbol{\epsilon}}
\end{aligned}
$$

Recall that $Y_i^1 = \beta_1^T X_i + \mathcal{E}_i^1$ and $Y_i^0 = \beta_1^T X_i + \mathcal{E}_i^0$. Let $\mathcal{E}_i = \frac{\mathcal{E}_i^1+\mathcal{E}_i^0}{2}$ and $\tilde{\boldsymbol{\mathcal{E}}} = (\mathcal{E}_1,\cdots,\mathcal{E}_n)$. Since $\tilde{\boldsymbol{\epsilon}}$ is the value of $\tilde{\boldsymbol{\mathcal{E}}}$ we have:

$$
\begin{aligned}
&\frac{n^2}{4}\mathbb{E}^S_{\mathbf{Y}}\left[\mathrm{var}_{\mathbf{Z}}(\hat{\tau}_Y^T|\mathbf{x},\mathbf{Y},\rho=1)|\mathbf{x}\right] \\
&= \beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1]\beta + \mathbb{E}^S_{\mathbf{Y}}[\tilde{\boldsymbol{\mathcal{E}}}^T Cov(\mathbf{Z}|\mathbf{x},\mathbf{Y},\rho=1)\tilde{\boldsymbol{\mathcal{E}}}|\mathbf{x}] + 2\beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z},\mathbf{Z}|\mathbf{x},\mathbf{y},\rho=1)\mathbb{E}[\tilde{\boldsymbol{\mathcal{E}}}|\mathbf{x}] \\
&= \beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}|\mathbf{x},\rho=1]\beta + \mathbb{E}^S_{\mathbf{Y}}[Cov(\tilde{\boldsymbol{\mathcal{E}}}^T\mathbf{Z}|\mathbf{x},\rho=1)|\mathbf{x}] \text{ because } \mathbb{E}[\tilde{\boldsymbol{\mathcal{E}}}|\mathbf{x}] = 0
\end{aligned}
$$

The second term:

$$
\begin{aligned}
&\mathbb{E}^{S}_{\mathbf{Y}}[Cov(\tilde{\boldsymbol{\mathcal{E}}}^T\mathbf{Z}|\mathbf{x},\rho=1)|\mathbf{x}] \\
&= \mathbb{E}_{\tilde{\boldsymbol{\mathcal{E}}}}[\mathbb{E}_{\mathbf{Z}}[\tilde{\boldsymbol{\mathcal{E}}}^T\mathbf{Z}\mathbf{Z}^T\tilde{\boldsymbol{\mathcal{E}}}|\mathbf{x},\rho=1]|\mathbf{x}] \\
&= \mathbb{E}_{\tilde{\boldsymbol{\mathcal{E}}}}\left[\sum_{i=1}^n\sum_{j=1}^n \mathbb{E}_{\mathbf{Z}}[w_i\mathcal{E}_iZ_iZ_j\mathcal{E}_jw_j|\rho=1]|\mathbf{x}\right] \\
&= \mathbb{E}_{\tilde{\boldsymbol{\mathcal{E}}}}\left[\sum_{i=1}^n \mathbb{E}[w_i^2\mathcal{E}_i^2Z_i^2|\rho=1] + \sum_{i\neq j}\mathbb{E}[w_i\mathcal{E}_iZ_iZ_j\mathcal{E}_jw_j|\rho=1]|\mathbf{x}\right] \\
&= \mathbb{E}_{\tilde{\boldsymbol{\mathcal{E}}}}\left[\sum_{i=1}^n \mathbb{E}[w_i^2\mathcal{E}_i^2 1|\rho=1] + \sum_{i\neq j}\mathbb{E}[w_iZ_iZ_j\mathcal{E}_jw_j|\rho=1]\mathbb{E}[\mathcal{E}_i|\rho=1]|\mathbf{x}\right] \text{ because } Z_i^2=1 \\
&= \sum_{i=1}^n \mathbb{E}[w_i^2\mathcal{E}_i^2] \text{ because } \mathbb{E}[\mathcal{E}_i|\rho=1]=0 \\
&= \sum_{i=1}^n w_i^2\sigma^2_{\mathcal{E}}
\end{aligned}
$$

Putting all together:

$$
\begin{aligned}
\mathrm{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{x}) &= \frac{4}{n^2}\left(\beta^T Cov(\tilde{\mathbf{x}}^T\mathbf{Z}|\mathbf{x},\rho=1]\beta + \sum_{i=1}^n w_i^2\sigma^2_{\mathcal{E}}\right) + + \frac{2}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2 \\
&= \frac{4}{n^2}\beta^T Cov_{\mathbf{Z}}(\tilde{\mathbf{x}}^T\mathbf{Z}|\mathbf{x},\rho=1]\beta + \frac{6}{n^2}\sigma^2_{\mathcal{E}}\sum_{i=1}^n w_i^2
\end{aligned}
$$

□

Proof of Lemma A.2. We use the same decomposition of $\beta^T Cov_{\mathbf{Z}}(V|\mathbf{x},\Omega)\beta$ as in [Harshaw et al., 2019]. Let $\mathbf{e}_1,...,\mathbf{e}_n$ and $\lambda_1,..,\lambda_n$ be the normalized eigenvectors and corresponding eigenvalues of matrix $Cov_{\mathbf{Z}}(V|\mathbf{x},\mathcal{E})$. Since $Cov_{\mathbf{Z}}(V|\mathbf{x},\mathcal{E})$ is symmetric, the eigenvectors form an orthonormal basis so we can write β as a linear combination of $\mathbf{e}_1,..,\mathbf{e}_n$ and get:

$$\beta = \|\beta\|\sum_{i=1}^n \eta_i\mathbf{e}_i$$

where $\eta_i = \langle\beta,\mathbf{e}_i\rangle/\|\beta\|$ is the coefficient that captures the alignment of the weighted outcome β with respect to the eigenvector $\mathbf{e}_i$. Therefore:

$$\beta^T Cov_{\mathbf{Z}}(V|\mathbf{x},\Omega)\beta = \|\beta\|^2\sum_{i=1}^n \eta_i^2\lambda_i$$

Then:

$$\begin{aligned}
\mathbb{E}_\beta\left[\beta^T Cov_{\mathbf{Z}}(V|\mathbf{x},\Omega)\beta\right] &= \mathbb{E}_\beta\left[\|\beta\|^2\sum_{i=1}^n \eta_i^2\lambda_i\right] \\
&= l^2\sum_{i=1}^n \lambda_i \mathbb{E}_\beta[\eta_i^2] \\
&= l^2\sum_{i=1}^n \lambda_i \mathbb{E}_\theta cos^2(\theta) \quad \text{where } \theta \text{ is the angle between } \beta \text{ and } \mathbf{e}_i. \text{ Since } \beta \text{ points to any direction with equal probability, } \theta \text{ is uniformly distributed in } [0,2\pi]. \\
&= \frac{l^2}{2}\sum_{i=1}^n \lambda_i \\
&= \frac{l^2}{2}\text{Trace}(Cov_{\mathbf{Z}}(V|\mathbf{x},\Omega)).
\end{aligned}$$

□

Proof of Lemma A.3. Let $p(u)$ be the pdf of U. Define $f(u)$ as follow:

$$f(u) = p(U = u, \Omega)$$

Then:

$$p(u|\Omega) = \frac{p(U = u,\Omega)}{\mathbb{P}(\Omega)} = \frac{f(u)}{1-\alpha}$$

Since $\mathbb{P}(\Omega) = 1 - \alpha$ we have:

$$\int_u f(u)du = 1 - \alpha$$

We have:

$$Trace(Cov(U|\Omega)) = Trace(\mathbb{E}[UU^T|\Omega]) = Trace(\mathbb{E}[UU^T|\Omega] = Trace(\mathbb{E}[U^TU|\Omega] = \int_u u^Tu\frac{f(u)}{1-\alpha}du$$

We want to minimize $Trace(Cov(U|\Omega))$:

$$\int_u u^Tu\frac{f(u)}{1-\alpha}du$$

subject to:

$$\begin{gathered}
0 \le f(u) \le p(u)\ \forall u \\
\int_u f(u)du = 1 - \alpha
\end{gathered}$$

This can be done by maximize $f(u)$ so that $f(u) = p(u)$ for the smallest u^Tu, which is equal to set Ω to be the event $\|U\|^2 < u_\alpha$. □

Proof of Theorem A.1. Let $\eta := 1 - \mathbb{P}(\rho = 1|\mathbf{x})$. Then $\eta \le \alpha$. Let v_η be such that $\mathbb{P}(\|V\|^2 < v_\eta|\mathbf{x}) = 1 - \eta$.

From Lemma A.1:

$$\mathbb{E}_\beta \text{var}^S_{\mathbf{Y},\mathbf{Z}_\rho}(\hat{\tau}^T_Y|\mathbf{x}) = \frac{4}{n^2}\mathbb{E}_\beta \beta^T Cov(V|\mathbf{x},\rho=1)\beta + \frac{6}{n^2}\sigma^2\sum_{i=1}^n w_i^2 \tag{8}$$

$$= \frac{4}{n^2}\frac{l^2}{2}\text{Trace}(Cov(V|\mathbf{x},\rho=1)) + \frac{6}{n^2}\sigma^2\sum_{i=1}^n w_i^2 \tag{9}$$

$$\geq \frac{4}{n^2}\frac{l^2}{2}\text{Trace}(Cov(V|\mathbf{x},\|V\|^2 < v_\eta)) + \frac{6}{n^2}\sigma^2\sum_{i=1}^n w_i^2 \tag{10}$$

$$\geq \frac{4}{n^2}\frac{l^2}{2}\text{Trace}(Cov(V|\mathbf{x},\|V\|^2 < v_\alpha)) + \frac{6}{n^2}\sigma^2\sum_{i=1}^n w_i^2 \text{ because } v_\eta \geq v_\alpha \tag{11}$$

$$\geq \frac{4}{n^2}\frac{l^2}{2}\text{Trace}(Cov(V|\mathbf{x},{\phi^\alpha_T}'=1)) + \frac{6}{n^2}\sigma^2\sum_{i=1}^n w_i^2 \tag{12}$$

$$\geq \frac{4}{n^2}\mathbb{E}_\beta \beta^T Cov(V|\mathbf{x},{\phi^\alpha_T}'=1)\beta + \frac{6}{n^2}\sigma^2\sum_{i=1}^n w_i^2 \tag{13}$$

$$\geq \mathbb{E}_\beta \text{var}^S_{\mathbf{Y},\mathbf{Z}_{{\phi^\alpha_T}'}}(\hat{\tau}^T_Y|\mathbf{x}) \tag{14}$$

□

Proof of Corollary A.1. Let ρ being the constant function $\rho(\mathbf{x},\mathbf{Z}) = 1$ for all $\mathbf{x},\mathbf{Z}$. Then:

$$\text{var}^S_{\mathbf{Z}_\rho,\mathbf{Y}}(\hat{\tau}^T_Y|\mathbf{x}) = \text{var}^S_{\mathbf{Z},\mathbf{Y}}(\hat{\tau}^T_Y|\mathbf{x})$$

From Theorem A.1 we have:

$$\mathbb{E}_\beta \text{var}^S_{\mathbf{Z}_{\phi'_T},\mathbf{Y}}(\hat{\tau}^T_Y|\mathbf{x}) \leq \mathbb{E}_\beta \text{var}^S_{\mathbf{Z}_\rho,\mathbf{Y}}(\hat{\tau}^T_Y|\mathbf{x}) = \mathbb{E}_\beta \text{var}^S_{\mathbf{Z},\mathbf{Y}}(\hat{\tau}^T_Y|\mathbf{x})$$

□