Proof detail of our $GAN_p(G, D_p)$ and $GAN_n(G, D_n)$.

Similar to GAN network,

$$GAN_p(G, D_p) = E_{x \sim p_s(x)} \left( w_p \log \left( D_p(x) \right) \right) + E_{x \sim p_t(x)} \left( \log \left( 1 - D_p(x) \right) \right)$$

$$= \int_x p_s(x) w_p \log \left( D_p(x) \right) dx + \int_x p_t(x) \log \left( 1 - D_p(x) \right) dx$$

$$= \int_x p_s(x) w_p \log \left( D_p(x) \right) + p_t(x) \log \left( 1 - D_p(x) \right) dx$$

Given $x$, the optimal $D_p^*$ is maximizing

$$p_s(x) w_p \log \left( D_p(x) \right) + p_t(x) \log \left( 1 - D_p(x) \right)$$

Find $D_p^*$ maximizing:

$$f(D_p^*) = p_s(x) w_p \log \left( D_p(x) \right) + p_t(x) \log \left( 1 - D_p(x) \right)$$

$$\frac{df(D_p)}{dD_p} = \frac{p_s(x) w_p}{D_p(x)} - \frac{p_t(x)}{1 - D_p(x)}$$

Hence,

$$D_p^* = \frac{p_s(x) w_p}{p_s(x) w_p + p_t(x)}$$

$$V(G, D_p^*) = E_{x \sim p_s(x)} \left( w_p \log \left( \frac{p_s(x) w_p}{p_s(x) w_p + p_t(x)} \right) \right) + E_{x \sim p_t(x)} \left( \log \left( \frac{p_t(x)}{p_s(x) w_p + p_t(x)} \right) \right)$$

$$= \int_x w_p p_s(x) \log \left( \frac{p_s(x) w_p}{p_s(x) w_p + p_t(x)} \right) + p_t(x) \log \left( \frac{p_t(x)}{p_s(x) w_p + p_t(x)} \right) dx$$

$$= -2 \log 2 + KL(p_s(x) w_p \| \frac{p_s(x) w_p + p_t(x)}{2}) + KL(p_t(x) \| \frac{p_s(x) w_p + p_t(x)}{2})$$

$$= -2 \log 2 + 2JSDF(p_s(x) w_p \| p_t(x))$$

In $GAN_p$, we try to minimize the JSD between $p_s(x) w_p$ and $p_t(x)$.

Hence, when $p_s(x) w_p = p_t(x)$, the positive domain adaptation can narrow the gap between $p_{sp}$ and $p_t$. 
When it comes to \( GAN_n(G, D_n) \)
\[
GAN_n(G, D_n) = E_{x \sim p_n(x)}(w_n \log(D_n(x))) + E_{x \sim p_d(x)}(w_p \log(1 - D_n(x)))
\]
\[
+ E_{x \sim p_t(x)}(\log(1 - D_n(x)))
\]
\[
= \int x p_s(x) w_n \log(D_n(x)) \, dx + \int x p_s(x) w_p \log(1 - D_n(x))
\]
\[
+ \int x p_t(x) \log(1 - D_n(x)) \, dx
\]

Given \( x \), the optimal \( D_n^* \) is maximizing
\[
p_s(x) w_n \log(D_n(x)) + p_s(x) w_p \log(1 - D_n(x)) + p_t(x) \log(1 - D_n(x))
\]

Find \( D_n^* \) maximizing:
\[
\frac{df(D_n)}{dD_n} = \frac{p_s(x) w_n}{D_n(x)} - \frac{p_t(x) + p_s(x) w_p}{1 - D_n(x)}
\]

Hence,
\[
D_n^* = \frac{p_s(x) w_n}{p_s(x) + p_t(x)}
\]

\[
V(G, D_n^*) = E_{x \sim p_d(x)} \left( w_n \log \left( \frac{p_s(x) w_n}{p_s(x) + p_t(x)} \right) \right) + E_{x \sim p_s(x)} \left( w_p \log \left( \frac{p_s(x) w_p + p_t(x)}{p_s(x) + p_t(x)} \right) \right)
\]
\[
+ E_{x \sim p_t(x)} \left( \log \left( \frac{p_s(x) w_p + p_t(x)}{p_s(x) + p_t(x)} \right) \right)
\]
\[
= \int x p_s(x) w_n \log \left( \frac{p_s(x) w_n}{p_s(x) + p_t(x)} \right) + (p_s(x) w_p + p_t(x)) \log \left( \frac{p_s(x) w_p + p_t(x)}{p_s(x) w_p + p_t(x)} \right) \, dx
\]
\[
= -2 \log 2 + KL(p_s(x) w_n \| p_s(x) + p_t(x))
\]
\[
+ KL(p_s(x) w_p + p_t(x) \| p_s(x) w_p + p_t(x))
\]
\[
= -2 \log 2 + 2 JS\text{(JSD)}(p_s(x) w_n \| p_s(x) + p_t(x))
\]

In \( GAN_n \), we try to maximize the JSD between \( p_s(x) w_n \) and \( p_s(x) w_p + p_t(x) \).

Hence, when \( p_s(x) w_n \neq p_s(x) w_p + p_t(x) \), the negative domain adaptation can zoom out the gap between \( p_{sn} \) and the set of \( p_t \) and \( p_{sp} \).

Therefore, our framework peaks its optimal solution if \( p_s(x) w_p = p_t(x) \) and \( p_s(x) w_n \neq p_s(x) w_p + p_t(x) \). The proof reveals that approaching to Nash equilibrium is equivalent to jointly minimize \( JSD(p_s(x) w_p \| p_t(x)) \) and maximize \( JSD(p_s(x) w_n \| p_s(x) w_p + p_t(x)) \).