A Proof of Expression in Equation (6)

In this section, we restate the results provided in [3, 16] in order to obtain eq.(6). We follow the well-known Lagrangean duality approach as in Appendix A in [14].

The following result is provided in [3, 16]. Let $\Omega(\mu, \Sigma)$ be the family of all distributions with mean $\mu$ and covariance $\Sigma$. For a fixed weight vector $\mathbf{a}$ and constant $b$, we have:

\[
\sup_{z \in \Omega(\mu, \Sigma)} \mathbb{P}_{z \sim z}[\mathbf{a}^T z \geq b] = \frac{1}{1 + \frac{d^2}{2}}
\]

where $d^2 = \inf_{\mathbf{a}^T z \geq b} \mathbf{a}^T \Sigma^{-1} \mathbf{a}$

Let $\mathbf{a} = -\mathbf{w}$ and $b = 0$. We have:

\[
\sup_{z \in \Omega(\mu, \Sigma)} \mathbb{P}_{z \sim z}[\mathbf{w}^T z \leq 0] = \frac{1}{1 + \frac{d^2}{2}}
\]

where $d^2 = \inf_{\mathbf{w}^T z \leq 0} \mathbf{w}^T \Sigma^{-1} \mathbf{w}$

Note that if $\mathbf{w}^T \mu \leq 0$, then we can just take $z = \mu$ and obtain $d^2 = 0$, which is certainly the optimum because $d^2 \geq 0$ due to positive definiteness of $\Sigma$. In what follows, we assume $\mathbf{w}^T \mu > 0$, as required in eq.(6).

We are interested in the value of $d^2$. That is, we seek for a closed-form solution of the primal problem:

\[
\min_{\mathbf{w}^T z \leq 0} (\mathbf{z} - \mu)^T \Sigma^{-1} (\mathbf{z} - \mu) = (36)
\]

which has the following Lagrangian:

\[
\mathcal{L}(\mathbf{z}, \lambda) = (\mathbf{z} - \mu)^T \Sigma^{-1} (\mathbf{z} - \mu) + \lambda \mathbf{w}^T \mathbf{z}
\]

By optimality arguments (i.e. $\partial \mathcal{L}/\partial \mathbf{z} = 0$), we have that $\mathcal{L}$ is minimized at $\mathbf{z}^* = -\frac{\lambda}{2} \Sigma \mathbf{w} + \mu$. Therefore, the Lagrange dual function is given by:

\[
g(\lambda) = \inf_{\mathbf{z}} \mathcal{L}(\mathbf{z}, \lambda)
\]

\[
= \mathcal{L}(\mathbf{z}^*, \lambda)
\]

\[
= -\frac{\lambda^2}{4} \mathbf{w}^T \Sigma \mathbf{w} + \lambda \mathbf{w}^T \mu
\]

Consequently, the dual problem of eq.(36) is:

\[
\max_{\lambda \geq 0} g(\lambda)
\]

Again, by optimality arguments (i.e. $\partial g/\partial \lambda = 0$), we have that $g$ is maximized at $\lambda^* = 2 \frac{\mathbf{w}^T \mu}{\mathbf{w}^T \Sigma \mathbf{w}}$. Note that $\lambda^* \geq 0$ since $\mathbf{w}^T \mu > 0$. Finally:

\[
d^2 = \max_{\lambda \geq 0} g(\lambda)
\]

\[
= g(\lambda^*)
\]

\[
= \frac{(\mathbf{w}^T \mu)^2}{\mathbf{w}^T \Sigma \mathbf{w}}
\]

\[
= \mathcal{F}(\mathbf{w}|\mu, \Sigma)
\]

B Moment Generating Function of the Square of a Sub-Gaussian Variable

Let $s$ be a sub-Gaussian variable with parameter $\sigma_s$ and mean $\mu_s = \mathbb{E}[s]$. By sub-Gaussianity, we know that the moment generating function is bounded as follows:

\[
(\forall t \in \mathbb{R}) \quad \mathbb{E}[e^{t(s-\mu_s)}] \leq e^{\frac{1}{2}t^2\sigma_s^2}
\]

Our goal is to find a similar bound for the moment generating function of the sub-exponential variable $v = s^2$. Let $\Gamma(r)$ be the Gamma function, the moments of the sub-Gaussian variable $s$ are bounded as follows:

\[
(\forall r \geq 0) \quad \mathbb{E}[|s|^r] \leq r^{2/r} \sigma_s^r \Gamma(r/2)
\]

Let $\mu_v = \mathbb{E}[v]$. By power series expansion and since $\Gamma(r) = (r - 1)!$ for an integer $r$, we have:

\[
\mathbb{E}[e^{t(v-\mu_v)}] = 1 + t\mathbb{E}[v - \mu_v] + \sum_{r=2}^{\infty} \frac{t^r \mathbb{E}[(v - \mu_v)^r]}{r!}
\]

\[
\leq 1 + \sum_{r=2}^{\infty} \frac{t^r \mathbb{E}[|s|^{2r}]}{r!}
\]

\[
\leq 1 + \sum_{r=2}^{\infty} \frac{t^r 2r 2^r \sigma_s^{2r} \Gamma(r)}{r!}
\]

\[
= 1 + \sum_{r=2}^{\infty} \frac{t^r 2r 2^r \sigma_s^{2r}}{r!}
\]

\[
= 1 + \sum_{r=2}^{\infty} \frac{8t^r \sigma_s^{2r}}{1 - 2t \sigma_s^2}
\]

By making $|t| \leq 1/(4\sigma_s^2)$, we have $1/(1 - 2t \sigma_s^2) \leq 2$. Finally, since $\mathbb{V} \alpha 1 + \alpha \leq e^\alpha$, we have that for a sub-Gaussian variable $s$ with parameter $\sigma_s$:

\[
(\forall |t| \leq 1/(4\sigma_s^2)) \quad \mathbb{E}[e^{t(s^2-\mathbb{E}[s^2])}] \leq e^{16t^2 \sigma_s^4}
\]

(37)

Thus, we obtained a bound for the moment generating function of the sub-exponential variable $s^2$, that is similar to that of sub-Gaussian variables but holds only for a small range of $t$. 