We first introduce some notations that are used by Hajivassiliou *et al.* (1996). For a vector η with elements $(\eta_1, \cdots, \eta_n)$, we use the notation $\eta_{<j}$ to denote the subvector $(\eta_1, \cdots, \eta_{j-1})$ that contains the first $j-1$ elements of η, and use η_{-j} to denote a subvector that excludes component j of η. Similarly, for a matrix Σ, if $\Sigma_{j.}$ is a vector that stores the j-th row of Σ, then $\Sigma_{j.<j}$ denotes a vector containing the first $j-1$ elements of the j-th row of Σ, and $\Sigma_{-j,-j}$ denotes the submatrix of Σ excluding row j and column j.

Let $\mathbf{y}$ be a $p \times 1$ random vector that is distributed as normal $N_p(\tilde{\mu}, \tilde{\Sigma})$. Let $\mathbf{r} = (r_1, \cdots, r_p)^T$ be the rank pattern of $\mathbf{y}$, and $\mathbf{ro} = (ro(1), \cdots, ro(p))^T$ be the order of the elements of $\mathbf{y}$ such that $y_{ro(1)} >, \cdots, > y_{ro(p)}$. If $\mathbf{y}$ is subjected to the rank constraint $\mathbf{r}$, we say that $\mathbf{y}$ has a rank truncated normal distribution, and the probability density function of $\mathbf{y}$ is then given by

$$f(\tilde{\mu}, \tilde{\Sigma}) I_{D^*}(\mathbf{y}) / Pr(\mathbf{y} \in D^*), \tag{1}$$

where $f(\tilde{\mu}, \tilde{\Sigma})$ is the probability density of the normal distribution with mean $\tilde{\mu}$ and variance $\tilde{\Sigma}$, $D^* = \{\mathbf{y} | y_{ro(1)} > y_{ro(2)} > \cdots > y_{ro(p)}\}$, $I_{D^*}(\mathbf{y})$ is an indicator function whose value is 1 when $\mathbf{y} \in D^*$ and 0 elsewhere, and $Pr(\mathbf{y} \in D^*)$ is the probability that $\mathbf{y}$ falls in D^*.

Now we will derive the first and second moment of $(\mathbf{y} - \mu)$ subject to the rank constraints. Without loss of generality, we consider the case $\mathbf{r} = (1, 2, \cdots, p)^T$, that is, $D^* = \{\mathbf{y} | y_1 > y_2 > \cdots y_p\}$. Let $\mathbf{v} = \mathbf{A}\mathbf{y}$ where $\mathbf{A}$ is a $p \times p$ constant matrix whose elements a_{ij} are all equal to zero except that $a_{ii} = 1$ for $i = 1, 2, \cdots, p$, and $a_{i,i+1} = -1$ for $i = 1, 2, \cdots, p-1$. Then, $\mathbf{v} \sim N(\mu, \Sigma)$, where $\mu = \mathbf{A}\tilde{\mu}$ and $\Sigma = \mathbf{A}\tilde{\Sigma}\mathbf{A}^T$. Let v_i, $i = 1, \cdots, p$ be the elements of $\mathbf{v}$, then the constraint that $\mathbf{y}$ falls in $D^* = \{\mathbf{y} | y_1 > y_2 > \cdots > y_p\}$ is equivalent to the constraint that $\mathbf{v}$ falls in D where

$$E[(\mathbf{y}-\tilde{\mu})|\mathbf{y}\in D^*] = \mathbf{A}^{-1}E[(\mathbf{v}-\mu)|\mathbf{v}\in D], \quad (2)$$

$$E[(\mathbf{y}-\tilde{\mu})(\mathbf{y}-\tilde{\mu})^T|\mathbf{y}\in D^*] = \mathbf{A}^{-1}E[(\mathbf{v}-\mu)(\mathbf{v}-\mu)^T|\mathbf{v}\in D]\mathbf{A}^{-T}. \quad (3)$$

As v_p has no constraint, given $v_1,\cdots,v_{p-1}$, its first and the second moment can be obtained in closed form by the conditional distribution of v_p given $\mathbf{v}_{-p}$:

$$v_p|\mathbf{v}_{-p} \sim N(\mu_p+\Sigma_{p.<p}\Sigma_{-p,-p}^{-1}(\mathbf{v}_{-p}-\mu_{-p}), \Sigma_{pp}-\Sigma_{p.<p}\Sigma_{-p,-p}^{-1}\Sigma_{p.<p}^T), \quad (4)$$

where μ_p and Σ_{pp} are respectively the last elements of μ and Σ. From (4), it can be shown that

$$\begin{aligned}
E[(v_p-\mu_p)|\mathbf{v}_{-p}] &= \Sigma_{p.<p}\Sigma_{-p,-p}^{-1}(\mathbf{v}_{-p}-\mu_{-p}), && (5)\\
E[(v_p-\mu_p)^2|\mathbf{v}_{-p}] &= \Sigma_{pp}-\Sigma_{p.<p}\Sigma_{-p,-p}^{-1}[\Sigma_{-p,-p} \\
&\quad -(\mathbf{v}_{-p}-\mu_{-p})(\mathbf{v}_{-p}-\mu_{-p})^T]\Sigma_{-p,-p}^{-1}\Sigma_{p.<p}^T && (6)\\
E[(v_p-\mu_p)(\mathbf{v}_{-p}-\mu_{-p})|\mathbf{v}_{-p}] &= (\mathbf{v}_{-p}-\mu_{-p})(\mathbf{v}_{-p}-\mu_{-p})^T\Sigma_{-p,-p}^{-1}\Sigma_{p.<p}^T && (7)\\
E(\mathbf{v}-\mu)=E\Big[E[(\mathbf{v}-\mu)|\mathbf{v}_{-p}]\Big] &= E((\mathbf{v}_{-p}-\mu_{-p})^T, E[(v_p-\mu_p)|\mathbf{v}_{-p}])^T; && (8)\\
E(\mathbf{v}-\mu)(\mathbf{v}-\mu)^T &= E\Big[E[(\mathbf{v}-\mu)(\mathbf{v}-\mu)^T|\mathbf{v}_{-p}]\Big] && (9)
\end{aligned}$$

$$= \begin{bmatrix} E[(\mathbf{v}_{-p}-\mu_{-p})(\mathbf{v}_{-p}-\mu_{-p})^T] & E\Big[(\mathbf{v}_{-p}-\mu_{-p})E[(v_p-\mu_p)|\mathbf{v}_{-p}]\Big] \\ E\Big[(\mathbf{v}_{-p}-\mu_{-p})^T E[(v_p-\mu_p)|\mathbf{v}_{-p}]\Big] & E\Big[E[(v_p-\mu_p)^2|\mathbf{v}_{-p}]\Big] \end{bmatrix}.$$

It can be seen from the above derivation that to calculate $E[(\mathbf{v}-\mu)|\mathbf{v}\in D]$ and $E[(\mathbf{v}-\mu)(\mathbf{v}-\mu)^T|\mathbf{v}\in D]$, we only need to calculate the $E[(\mathbf{v}_{-p}-\mu_{-p})|\mathbf{v}_{-p}\in D_1]$ and $E[(\mathbf{v}_{-p}-\mu_{-p})(\mathbf{v}_{-p}-\mu_{-p})^T|\mathbf{v}_{-p}\in D_1]$, where $D_1=\{v_1>0,\cdots,v_{p-1}>0\}$. It is easy to see that the distribution of $\mathbf{v}_{-p}$ subject to the constraint D_1 is $\mathbf{v}_{-p}\sim N_{p-1}(\mu_{-p},\Sigma_{-p,-p})I_{D_1}(\mathbf{v}_{-p})/Pr(\mathbf{v}_{-p}\in D_1)$. Hence, the moments of rank truncated angle truncated multinormal distribution with $p-1$ dimensions.

Let $\eta = (\eta_1, \cdots, \eta_{p-1})^T \sim N(\mathbf{0}, \mathbf{I}_{p-1})$, where $\mathbf{I}_{p-1}$ is an identity matrix. Let $\phi(x)$ and $G(x)$ be the pdf and cdf of standard normal random variate, respectively. Let Γ be the Cholesky factor of $\Sigma_{-p,-p}$, such that $\Sigma_{-p,-p} = \Gamma\Gamma^T$. Then according to Hajivassiliou, *et al.* (1996), we have

$$E[(\mathbf{v}_{-p} - \mu_{-p})|\mathbf{v}_{-p} \in D_1] = \Gamma \int c_0^{-1} \eta \omega(\eta) \prod_{j=1}^{p-1} \phi(\eta_j|\mathbf{C}_j(\eta_{<j}))d\eta, \tag{10}$$

$$E[(\mathbf{v}_{-p} - \mu_{-p})(\mathbf{v}_{-p} - \mu_{-p})^T|\mathbf{v}_{-p} \in D_1] = \Gamma \left(\int c_0^{-1} \eta\eta^T \omega(\eta) \prod_{j=1}^{p-1} \phi(\eta_j|\mathbf{C}_j(\eta_{<j}))d\eta \right) \Gamma^T, \tag{11}$$

where

$$\mathbf{C}_j(\eta_{<j}) = \{\eta_j|(-\mu_j - \Gamma_{j.<j}\eta_{<j})/\Gamma_{jj} < \eta_j < +\infty\}, \tag{12}$$

$$\phi(\eta_j|\mathbf{C}_j(\eta_{<j})) = \phi(\eta_j) I_{\mathbf{C}_j(\eta_{<j})}(\eta_j)/G(\mathbf{C}_j(\eta_{<j})), \tag{13}$$

$$G(\mathbf{C}_j(\eta_{<j})) = 1 - G((-\mu_j - \Gamma_{j.<j}\eta_{<j})/\Gamma_{jj}), \tag{14}$$

$$\omega(\eta) = \prod_{j=1}^{p-1} G(\mathbf{C}_j(\eta_{<j})), \tag{15}$$

$$c_0 = \int \omega(\eta) \prod_{j=1}^{p-1} \phi(\eta_j|\mathbf{C}_j(\eta_{<j}))d\eta, \tag{16}$$

and $\phi(\eta_j|\mathbf{C}_j(\eta_{<j}))$ is the conditional density of η_j given the event $\mathbf{C}_j(\eta_{<j})$. An observation η with its elements η_j can therefore be drawn recursively from the one-dimensional conditional distribution $\phi(\eta_j|\mathbf{C}_j(\eta_{<j}))$ by taking

$$\eta_j = G^{-1}(u_j G(-\mu_j - \Gamma_{j.<j}\eta_{<j})/\Gamma_{jj} + 1 - u_j),$$

where u_j are draws from the uniform [0,1] density. The sampling method that draws a sample $\eta^1, \cdots, \eta^m$ with each of these observations drawn using the recursive method that

$$\Gamma \frac{\sum_{i=1}^{m} \eta^i \omega(\eta^i)}{\sum_{i=1}^{m} \omega(\eta^i)} \qquad \text{and} \qquad \Gamma \frac{\sum_{i=1}^{m} \eta^i \eta^{iT} \omega(\eta^i)}{\sum_{i=1}^{m} \omega(\eta^i)} \Gamma^T \tag{17}$$

are good approximations of the integrals in (10) and (11). As a result, the conditional expectations in the left-hand side of (2) and (3) can be obtained by GHK simulator.