Notes

1) The Importance sampling algorithm (4.1) in the paper is about estimating the portfolio lose probability $P(L > x)$.

2) The algorithm is guided by the (delta-gamma) quadratic approximation of the portfolio loss distribution where the loss $L \approx a_0 + Q$ where:
   a) $Q = a'\Delta S + \Delta S\, A\, \Delta S' =_d (a'C)X + X'(C'AC)X$, Where:
      i) $\Delta S \; is\ the\ multivariate\ t\ \left(f_{v,\Sigma}\right) vector\ of\ the\ change\ in\ risk\ factors$
      ii) $a' = -\delta$, δ is the first order sensitivity vector (Question: how will we estimate that? May be 1st order approximation of the 1st derivatives?)
      iii) $A = -\frac{1}{2}\,\Gamma$, Where Γ is the second order sensitivity matrix (Question: how will we estimate that? May be 2nd order approximation of the 2nd derivatives?
      iv) $C\,C' = \Sigma$, among all C there is one that makes C'AC diagonal matrix (Question: Is C is the lower triangle Cholesky decomposition matrix that would make C'AC diagonal?)
      v) X is the multivariate t vector ($f_{v,\mathrm{I}}$) with uncorrelated but not necessarily independent components
      vi) $Q =_d (a'C)X + X'(C'AC)X = \sum_{j=1}^{m}(b_j X_j + \lambda_j X_j^2)$, $b = a'C, \lambda_j$ is the diagonal elements of $C'AC$
      vii) Since X_j, X_j^2 have heavy tailed distributions ($X_j^2 \sim$ F), they don't have moment generating functions.
      viii) Indirect approach of using $Q_x = \left(\frac{Y}{v}\right)(Q - x)$, where Y is the Chi Square random variable in $X = \frac{Z}{\sqrt{\frac{Y}{v}}} \sim t_v$
      ix) $P(Q < x) = P(Q - x < 0) = P\left(\left(\frac{Y}{v}\right)(Q - x) < 0\right) = P(Q_x < 0) = F_x(0)$
      x) Q_x is not a heavy tailed distribution : (Question: Why?)
      xi) $Q_x\ MGF = \phi_x(\theta) = \left(1 + \frac{2\theta x}{v} - \sum_{j=1}^{m}\left(\frac{(\theta b_j)^2}{v}\right) * \left(\frac{1}{1-2\theta\lambda_j}\right)\right)^{-\frac{v}{2}} \prod_{j=1}^{m} \frac{1}{\sqrt{1-2\theta\lambda_j}}$
      xii) Using exponential change of measure $dP_\theta = e^{\theta Q - \psi(\theta)}\, dP$,
      $P(L > x) = E_\theta\left[e^{-\theta Q + \psi(\theta)}\, \mathbb{I}\{L > x\}\right]$, where $\psi(\theta) = \log(\phi(\theta))$ = log(MGF(θ))
      xiii) Theorem: 4.1 If $\theta\,\lambda_1 < \frac{1}{2}\ and\ \alpha(\theta) < \bar{\theta}_Y$ then $dP_\theta = e^{\theta Q_x - \psi_x(\theta)}\, dP$
      $P(L > y) = E_\theta\left[e^{-\theta Q_x + \psi_x(\theta)}\, \mathbb{I}\{L > y\}\right] \approx P(Q > x) \rightarrow y = x + a0$

      $\psi_x(\theta) = \log(\phi_x(\theta))$

      xiv) In the case that the distribution of X under P measure (Y is chi-square), the distribution of Y under P_θ is gamma with shape $\frac{v}{2}$ and scale $= \frac{2}{1-2\alpha(\theta)}$

$$\alpha(\theta) = -\frac{\theta x}{v} + \frac{1}{2}\sum_{j=1}^{m}\left(\frac{(\theta b_j)^2}{v}\right) * \left(\frac{1}{1-2\theta\lambda_j}\right)$$
$$P(L > y) = E_\theta\left[e^{-\theta Q_x + \psi_x(\theta)}\,\mathbb{I}\{L > y\}\right]$$

xv) We would like to choose θ that would minimize the estimators second moment (hence the variance)
$E_\theta 2\left[e^{-2\theta Q_x + 2\psi_x(\theta)}\,\mathbb{I}\{L > y\}\right] = E_\theta\left[e^{-\theta Q_x + \psi_x(\theta)}\,\mathbb{I}\{L > y\}\right]$ using the quadratic approx. that $L \approx a_0 + Q$, we can bound the second moment

$$E_\theta\left[e^{-\theta Q_x + \psi_x(\theta)}\,\mathbb{I}\{Q > x\}\right] \le e^{\psi_x(\theta)}$$

$\psi_x(\theta)$ is convex, so that the upper bound is minimized by $\frac{d}{d\theta}\,\psi_x(\theta) = 0$

$$\frac{d}{d\theta}\,\psi_x(\theta) = \frac{d}{d\theta}\left(\log\left(\phi_x(\theta)\right)\right) =$$
$$\frac{d}{d\theta}\left(\log\left(\left(1 + \frac{2\theta x}{v} - \sum_{j=1}^{m}\left(\frac{(\theta b_j)^2}{v}\right) * \left(\frac{1}{1-2\theta\lambda_j}\right)\right)^{-\frac{v}{2}} \prod_{j=1}^{m}\frac{1}{\sqrt{1-2\theta\lambda_j}}\right)\right) = 0$$

That can be found numerically (using optimization)

3) The algorithm (4.1)

a) Prepare A, C matrices? (*Open for discussion*)
b) Choose Value of θ using optimization of $\frac{d}{d\theta}\,\psi_x(\theta) = 0$
c) Calculate $\psi_x(\theta)$
d) Generate Y from gamma(shape, scale) , the scale parameter is function if function of θ
e) Generate iid $Z_j \sim N\left(\mu_j(\theta),\ \sigma_j^2(\theta)\right)$
f) Set X = $X = \frac{Z}{\sqrt{\frac{Y}{v}}}$
g) Set ΔS = C X
h) Calculate Q = $a'\Delta S + \Delta S\ A\ \Delta S'$ *and* $Q_x = \left(\frac{Y}{v}\right)(Q - x)$,
i) Multiply the loss indicator by the LR to get

$$M_x(\theta) = e^{-\theta Q_x + \psi_x(\theta)}\,\mathbb{I}\{L > y\}$$

j) Calculate $P(L > y) = E_\theta\left[e^{-\theta Q_x + \psi_x(\theta)}\,\mathbb{I}\{L > y\}\right]$ = mean($M_x(\theta)$) over n indpendent replications