Error function:

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

CDF

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

$$\Rightarrow \boxed{\Phi(x) \equiv \tfrac{1}{2} + \tfrac{1}{2} \operatorname{erf}(x/\sqrt{2})} \quad (1)$$

$$\operatorname{erf}(x/\sqrt{2}) \equiv 2\Phi(x) - 1 \quad (2)$$

$$\operatorname{erf}(x/\sqrt{2}) \equiv \Phi(x) - \Phi(-x) \quad (3)$$

Truncated Gaussian Density

$$p(y) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{2\Phi(c/\sigma) - 1} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right) \chi_{[-c,c]}(y) \quad (4)$$

this comes from a "parent" Normal distribution with mean μ and var $= \sigma^2$, i.e. $N(\mu, \sigma^2)$.

From Wikipedia (from numerical integration of $\mathbb{E}[Y]$, $\mathbb{E}[Y^2]$, $\mathbb{E}[Y]^2$), we know that $y \sim p(y)$ from (4) has:

$$\mathbb{E}[Y \mid Y \in [a,b]] = \mu + \sigma \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} \quad (5)$$

$$\operatorname{Var}[Y \mid Y \in [a,b]] = \sigma^2 \left[1 + \frac{\alpha\phi(\alpha) - \beta\phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} - \left(\frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)}\right)^2\right] \quad (6)$$

where $\alpha := a - \mu/\sigma$, $\beta := b - \mu/\sigma$, $\phi(x) = 1/\sqrt{2\pi} \exp(-\frac{1}{2}x^2)$. From now on, let $\mu = 0$.

And you realize that if $\sigma = 1$, $a = -1$, $b = 1$, it's not true that, in particular, $\operatorname{Var}(Y \mid Y \in [-1,1]) = 1$, in fact, it is ≈ 0.29.

Hence, although we start from the "parent" distribution $N(0,1)$, once we truncate the distribution to $[-1,1]$ (i.e. $c = 1$ in (4)), we don't have that $\operatorname{Var}(y) = 1$ in (4) !!

In general, the variance of the truncated is less than the variance of the "parent"!

Now, what we do in the code (in Feng's function) we assume the "parent" Normal distribution $\mathcal{N}(0,1)$, i.e. $\sigma = 1$ (← this is also the default value in the query).

The "problem" occour in exp. 1 of GOAFEM paper because we considered

$$a(x,y) = \mathbb{E}[a](x) + \tilde{\sigma} \sum_m \sqrt{\lambda_m}\, \varphi_m(x)\, y_m \qquad (7)$$

with y_m having the density (4). Ok. So, since we have $c=1$ in Feng's function, i.e., we are already letting $\Gamma_m = [-1,1]$, (and $\sigma=1$ as said above) we compute:

$$\rho_m(y_m) = \frac{1}{\sqrt{2\pi}} \frac{1}{2\Phi(1)-1} \exp\left(-\tfrac{1}{2}y^2\right) \qquad y_m \in [-1,1] \qquad (8)$$

Right. The problem is that $\mathrm{Var}(y_m) \neq 1$, but $\approx 0{,}29$, and KL expansion (7) wants the variance to be 1.

Also, very important, we may put $\sigma \neq 1$ in the query, <u>but there's no value of σ for which (with $c=1$) the variance of $y \sim \rho(y)$ as in (4) can be 1 !!</u>

Then, the idea is to consider (insert a constant $\bar{c}$)

$$a(x,y) = \mathbb{E}[a](x) + \bar{c}\,\tilde{\sigma} \sum_m \sqrt{\lambda_m}\, \varphi_m(x)\, y_m \,, \qquad \bar{c} \in \mathbb{R}_+ \qquad (9)$$

so that $z_m := \bar{c}\, y_m$ are rescaled parameter $\in [-\bar{c}, \bar{c}]$ s.t. $\mathrm{Var}(z_m) = 1 \ \forall m$. We find the density associated with r.v. z_m. Let's put just z (get rid of m):

$$F(z) := \mathbb{P}\left(\frac{Z}{\bar{c}} \leq z\right) = \mathbb{P}(\bar{c}Y \leq z) = \mathbb{P}(Y \leq z/\bar{c}) := \int_{-\infty}^{z/\bar{c}} \rho(y)\,dy \equiv$$

$$\equiv \int_{-1}^{z/\bar{c}} \frac{1}{\sqrt{2\pi}} \frac{1}{2\Phi(1)-1} \exp(-y^2/2)\,dy$$

(hence y_m has density ρ_m in (8), this is our current case here)

$$\equiv \frac{1}{\sqrt{2\pi}} \frac{1}{2\Phi(1)-1} \left[\sqrt{\frac{\pi}{2}}\,\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right]_{-1}^{z/\bar{c}} \equiv \frac{1}{2} \frac{1}{2\Phi(1)-1} \left(\mathrm{erf}\left(\frac{z}{\bar{c}\sqrt{2}}\right) - \mathrm{erf}\left(\frac{-1}{\sqrt{2}}\right)\right) \qquad (10)$$

Hence (10) is the distribution of r.v. $z_m = \bar{c}\, Y_m$.
Now the density is:

$$F'(z) = \frac{d}{dz} F(z) = \frac{1}{2} \frac{1}{2\Phi(1)-1} \left[\frac{d}{dz} \operatorname{erf}\left(\frac{z}{\bar{c}\sqrt{2}}\right) \right] \overset{\text{(from def of erf)}}{=}$$

$$\equiv \frac{1}{2} \frac{1}{2\Phi(1)-1} \left[\frac{1}{\bar{c}\sqrt{2}} \frac{2}{\sqrt{\pi}} e^{-z^2/2\bar{c}^2} \right] \equiv \frac{1}{\bar{c}} \frac{1}{\sqrt{2\pi}} \frac{1}{2\Phi(1)-1} \exp\left(-\frac{z^2}{2\bar{c}^2}\right)$$

$$=: \gamma(z) \qquad (11)$$

That is, (11) is the density of the r.v. $z_m = \bar{c}\, Y_m$ where Y_m has truncated Normal density (8) on $[-1,1]$.

Now one can choose $\bar{c} \in \mathbb{R}$ so that z_m have unit var. $\forall m$. Thus:

$$\operatorname{Var}(z_m) = \mathbb{E}[z_m^2] - \underbrace{\mathbb{E}[z_m]^2}_{=0} \quad (\text{as } Y_m \text{ have } \mu = 0)$$

$$\equiv \int_{-c}^{c} (z_m^2)\, \gamma_m(z_m)\, dz_m \quad (\text{as } z_m \in [-\bar{c}, \bar{c}])$$

$$\equiv \int_{-c}^{c} z_m^2 \frac{1}{\bar{c}} \frac{1}{\sqrt{2\pi}} \frac{1}{2\Phi(1)-1} \exp\left(-z_m^2/2\bar{c}^2\right) dz_m \qquad (12)$$

ooo and now with Matlab we see that $\boxed{\bar{c} \simeq 1.85336168}$ (13)

Thus, using $\bar{c}$ in (13) in kL expression (9) we ensure that r.v. $\bar{c}\, y_m$ have $\simeq$ variance equal to 1 $\frac{11}{00}$