S5  Estimating the variance for the parallel LF-NS scheme

In the following we discuss in detail the approximation error for the Bayesian evidence

\[ Z = \int_0^1 L(x)dx. \]

LF-NS approximates the above integral by approximating it on the final prior volume \( x_{m,r} \) and the remaining volume separately

\[ Z = \int_0^{x_{m,r}} L(x)dx + \int_{x_{m,r}}^1 L(x)dx. \]

The integral \( Z^m_r \) is approximated through a finite sum

\[ \int_{x_{m,r}}^1 L(x)dx \approx \sum_{i=1}^{m} \sum_{j=1}^{r} L(x_{i,j})(x_{i,j-1} - x_{i,j}) = \sum_{i=1}^{m} \sum_{j=1}^{r} \epsilon_{i,j}(x_{i,j-1} - x_{i,j}) =: Z^m_r. \]  

(5.1)

Since the prior volumes \( x_{i,j} \) are in general not know, the numerical approximation \( Z^m_r \) is approximated itself through the estimator

\[ \hat{Z}^m_r = \sum_{i=1}^{m} \sum_{j=1}^{r} \epsilon_{i,j}(\hat{x}_{i,j-1} - \hat{x}_{i,j}) \approx \hat{Z}^m_r, \]

where the prior volumes \( x_{i,j} \) are treated as random variables and approximated by their means \( \hat{x}_{i,j} \). The error in estimating \( Z_D \) through \( \hat{Z}^m_r \) is negligible compare to the error in estimating \( \hat{Z}^m_r \) through \( \hat{Z}^m_r \) (see the discussion in \[2\] or \[1\]).

To emphasize its dependence on the final volume \( x_{m,r} \), we rewrite the integral \( Z^m_r \) as an integral over the parameter space rather than the volume space

\[ Z^m_r = \int_0^{x_{m,r}} L(x)dx = x_{m,r} \int_0^{L(\theta)} d\Pi(\theta|\hat{L}(\theta) \geq \epsilon_{m,r}) =: L_{m,r} \]

The quantity \( L_{m,r} \) is the average of the likelihoods over the joint prior \( \Pi(\theta|\hat{L}(\theta)) \) constraint to the likelihood regions above \( \epsilon_{m,r} \). We approximate \( L_{m,r} \) with a Monte Carlo estimator \( \hat{L}_{m,r} \) and and the prior volume \( x_{m,r} \) with its mean \( \hat{x}_{m,r} \)

\[ Z^m_r \approx \hat{x}_{m,r} \hat{L}_{m,r} =: \hat{Z}^m_r. \]

Thus, the final estimation error can be written as

\[ \| Z - (\hat{Z}^m_r + \hat{Z}^m_r)\| \]
The error \( S5.1 \) Variance of \( \eta \)

with its Monte Carlo estimate prior volume. Its variance is 

\[
\text{The super script } (i) \text{ in } t_j^{(i)} \text{ emphasizes that it is the } i^{\text{th}} \text{ sample of the random variable } t_j. \text{ The random variable } t_j \text{ is distributed as the } j^{\text{th}} \text{ highest number among } N \text{ uniform numbers on the interval } [0, 1], \text{ which is the Beta distribution}
\]

\[
t_j \sim B(N - j + 1, j)
\]

with

\[
\hat{t}_j := \mathbb{E}(t_j) = \frac{N - j + 1}{N + 1}, \quad \mathbb{E}(t_j^2) = \frac{(N - j + 1)(N - j + 2)}{(N + 1)(N + 2)}
\]

and variances of

\[
\text{Var}(t_j) = \frac{(N - j + 1)j}{(N + 2)(N + 1)^2}.
\]

The corresponding values for the variables \( x_{i,j} \) are

\[
\hat{x}_{i,j} := \mathbb{E}(x_{i,j}) = \hat{t}_r^{-1}\hat{t}_j = \frac{(N - r + 1)^{i-1}(N - j + 1)}{(N + 1)^i},
\]

\[
\mathbb{E}(x_{i,j}^2) = \mathbb{E}\left( t_j^{(i)^2} \prod_{k=1}^{i-1} t_k^{(i)} \right) = \mathbb{E}(t_j^2)\mathbb{E}(t_j^2)^{i-1} = \frac{(N - j + 1)(N - j + 2)(N - r + 1)^{i-1}(N - r + 2)^{i-1}}{(N + 1)^i(N + 2)^i}.
\]

With this notation we compute the variances of the errors \( \eta_{L}^{m,r} \) and \( \eta_{D}^{m,r} \).

S5.1 Variance of \( \eta_{L} \)

The error \( \eta_{L} \) is the error of approximating the final prior volume \( x_{m,r} \) with its mean \( \hat{x}_{m,r} \) and replacing the integral \( L_m \) with its Monte Carlo estimate \( L_m \) when approximating the Bayesian evidence over the final prior volume. Its variance is

\[
\text{Var}(\eta_{L}^{m,r}) = \text{Var}(x_{m,r}L_{m,r} - \hat{x}_{m,r}L_m) = \text{Var}(x_{m,r})L_{m,r}^2 + \text{Var}(L_m)\hat{x}_{m,r}^2.
\]
The integral $L_{m,r}$ for the variance estimation can be estimated through the Monte Carlo estimate $\overline{L}_{m,r}$ and the variance of $\overline{L}_{m,r}$ is just

$$\text{Var} \left( \overline{L}_{m,r} \right) = \frac{\sigma_L^2}{N}, \quad \sigma_L^2 = \frac{1}{N-1} \sum_{(\theta,\tilde{l}) \in \mathcal{L}_m} (\tilde{l} - L_{m,r})^2.$$  

The last average is taken over the points in the last live point set $\mathcal{L}_{m,r}$, which are all distributed according to $\Pi(\theta, \tilde{l} | \hat{l} | \theta) > \epsilon_{m,r}$.

### S5.2 Variance of $\eta_D$

Next, we estimate the variance of the estimation through the dead points. We have $\text{Var}(\eta_D) = \text{Var}(\tilde{Z}^{m,r}_D - \tilde{Z}^{m,r}_D) = \text{Var}(\tilde{Z}^{m,r}_D)$. We first rewrite

$$\tilde{Z}^{m,r}_D = \sum_{i=1}^{m} \sum_{j=1}^{r} \epsilon_{i,j} (x_{i,j-1} - x_{i,j}) = \sum_{i=1}^{m} x_{i-1,r} \sum_{j=1}^{r} \epsilon_{i,j} \left( t^{(i)}_{j-1} - t^{(i)}_j \right).$$  

(5.2)

We point out that for each $i$, the samples $x_{i-1,r}$ and $E_i$ are statistically independent since $E_i$ only contain the $i^{th}$ samples of $t$ and $x_{i-1,r} = \prod_{k=1}^{i-1} t^{(k)}_r$ contains only the samples up to $i - 1$. For $m > 1$, we can write

$$\text{Var}(\tilde{Z}^{m,r}_D) = \text{Var}(\tilde{Z}^{m-1,r}_D) + 2 \text{Cov}(x_{m-1,r} E_m, \tilde{Z}^{m-1,r}_D) + \text{Var}(x_{m-1,r} E_m).$$  

(5.3)

We fist compute the general formula for $\text{Cov} \left( x_{m,r}, \tilde{Z}^{m,r}_D \right)$. The product of $x_{m,r}$ and $\tilde{Z}^{m,r}_D$ is

$$x_{m,r} \tilde{Z}^{m,r}_D = \sum_{i=1}^{m} x_{m,r,x_{i-1,r} E_i}$$

$$= \sum_{i=1}^{m} \prod_{l=1}^{i-1} t^{(l)}_r \prod_{k=i+1}^{m} t^{(k)}_r \sum_{j=1}^{r} \epsilon_{i,j} \left( t^{(i)}_{r,t_{j-1}} - t^{(r)}_{r,t_j} \right).$$

In the above formulation we used the convention that \( \prod_{k=m}^{m-1} = 1 \). With this and using that for random variables $X$, $Y$ and $Z$ we have $\mathbb{E} \left( X^2 \right) \mathbb{E} \left( Y Z \right) = \mathbb{E} \left( X \right) \mathbb{E} \left( Y \right) \mathbb{E} \left( Z \right) = \text{Var} \left( X \right) \mathbb{E} \left( Y Z \right) + \mathbb{E} \left( X \right)^2 \text{Cov} \left( Y, Z \right)$ we compute the covariance

$$\text{Cov} \left( x_{m,r}, \tilde{Z}^{m}_D \right) = \sum_{i=1}^{m} \sum_{j=1}^{r} \epsilon_{i,j} \left( \text{Var}(x_{i-1,r} | (\mathbb{E}(t_r,t_{j-1}) - \mathbb{E}(t_r,t_j)) + \tilde{x}^{2}_{m-i,r} (\text{Cov}(t_r,t_{j-1}) - \text{Cov}(t_r,t_j))) \right)$$

(5.4)

\(^{1}\)for $m = 1$ we obviously have $\text{Var} \left( \tilde{Z}^{1}_D \right) = \text{Var}(E_i)$.  

3
\[
\sum_{i=1}^{m} x_{m-i,r} \sum_{j=1}^{r-1} \epsilon_{i,j} \text{Var} (x_{i-1,r}) (\hat{t}_j \hat{t}_{j-1} - \hat{t}_j \hat{t}_j)
\]

\[+
\sum_{i=1}^{m} \hat{x}_{m-i,r} \epsilon_{i,r} \text{Var} (x_{i-1,r}) (\hat{t}_r \hat{t}_{r-1} - \text{Var} (x_{i,r})).\]

Next we compute the variance

\[
\text{Var}(x_{m-1,r} E_m) = \mathbb{E} \left( x_{m-1}^2 \right) \text{Var}(E_m) + \text{Var}(x_{m-1,r} \hat{E}_m). \tag{5.5}
\]

The square of \( E_m \) is

\[
E_m^2 = \left( \sum_{j=1}^{r} \epsilon_{m,j} (t_j^{(m)} - t^{(m)}_j) \right)^2
\]

\[= \sum_{j=1}^{r} \sum_{k=1}^{r} \epsilon_{m,j} \epsilon_{m,k} \left( t_{j-1}^{(m)} - t_j^{(m)} \right) \left( t_{k-1}^{(m)} - t_k^{(m)} \right)
\]

and thus the variance of \( E_m \) is

\[
\text{Var}(E_m) = \sum_{j=1}^{r} \sum_{k=1}^{r} \epsilon_{m,j} \epsilon_{m,k} \text{Cov} (t_{j-1} - t_j, t_{k-1} - t_k).
\]

Putting this variance of \( E_m \) into formula (5.5) we obtain

\[
\text{Var}(x_{m-1,r} E_m) =
\]

\[
= \sum_{j=1}^{r} \sum_{k=1}^{r} \epsilon_{m,j} \epsilon_{m,k} \left( \mathbb{E}(x_{m-1,r}^2) \text{Cov} (t_{j-1} - t_j, t_{k-1} - t_k) + \text{Var}(x_{m-1,r}) (\hat{t}_{j-1} - \hat{t}_j) (\hat{t}_{k-1} - \hat{t}_k) \right)
\]

\[- 2 \sum_{j=2}^{r} \sum_{k=1}^{r} \epsilon_{m,j} \epsilon_{m,j-1} \mathbb{E}(x_{m-1,r}^2) \text{Var} (t_{j-1})
\]

Combing the variance and covariance term in (5.3) we get

\[
\text{Var}(\eta_{m,r}) = \text{Var}(\eta_{m-1,r}) + 2 \hat{E}_m \text{Cov}(x_{m-1}, \hat{Z}_{m-1}^{m-1}) + \text{Var}(x_{m-1,r} E_m). \tag{5.6}
\]
S5.3 Total variance

The total variance is

\[ \text{Var}(\eta_m^L + \eta_m^D) = \text{Var}(\eta_m^L) + \text{Var}(\eta_m^D) + 2 \text{Cov}(\eta_m^L, \eta_m^D). \]

Since \( \eta_L \) and \( \eta_D \) are not independent (they both share the random variable \( x_{m,r} \)), we have to account for their dependence.

\[ \text{Cov}(\eta_m^L, \eta_m^D) = \text{Cov}(x_{m,rL_{m,r}}, \tilde{Z}_{m,rD}) = L_{m,r} \text{Cov}(x_{m,r}, \tilde{Z}_{m,rD}) \]

This covariance can be computed using formula 5.4. The total variance is

\[ \text{Var}(\eta_m^L + \eta_m^D) \]

\[ = \text{Var}(\eta_m^L) + \text{Var}(\eta_m^D) + 2L_{m,r} \text{Cov}(x_{m,r}, \tilde{Z}_{m,rD}) \]
References
