Publication Date: May 2, 2018
Effective Date: May 2, 2018
Addendum 11 to the CRI Technical Report (Version: 2017, Update 1)

This document details four changes in the parameter estimation method used in the CRI system. The CRI estimation method prior to these changes is documented in Technical Report (Version: 2017, Update 1). The four changes are:

- Lowering the grid search criterion for the tempering sequence, ξ_p.
- Adopting a new design of the resampling and move steps, and a mixture proposal distribution coupled with random block replacements is implemented.
- Changing the stopping criterion for the final tempering step.
- Introducing re-initialization in the sequential updating and periodic updating steps.

I. Grid search for the tempering sequence ξ_p

At the beginning of each tempering step, p, a reweighting procedure is run (Equation 21 of Technical Report Version: 2017, Update 1):

$$\bar{\omega}^{(k,n+1,p)} = \bar{\omega}^{(k,n+1,p-1)} \times \mathcal{L}_{n,\min(N-n,\ell)}^{\xi_p - \xi_{p-1}}(\bar{\theta}^{(k,n+1,p)}),$$

where ξ_p is chosen to ensure that a minimum effective sample size (ESS) is maintained, where ESS is defined as

$$ESS = \frac{\left(\sum_{k=1}^{N} \bar{\omega}^{(k,n+1,p)}\right)^2}{\sum_{k=1}^{N}\left(\bar{\omega}^{(k,n+1,p)}\right)^2}.$$

The newly adopted minimum ESS is 25% of the sample size, which equals 250 with the CRI's use of the SMC sample of 1,000 parameter particles. This is done by a grid search, where the ESS is evaluated at a grid of candidate values for ξ_p. The one that produces the ESS that is larger than and closest to 250 is chosen. By changing the criterion from 500 to 250, bigger steps for ξ_p are taken to speed the algorithm without adversely affecting the quality of the estimation result.

II. Resampling and move steps

To avoid particle impoverishment caused by sequential importance sampling, where by nature most of the weight will likely be concentrated in a small number of particles, a resample-move step has to be conducted whenever $\xi_p < 1$. Here, resampling directs the particle cloud towards more likely areas of the sampling space, while the move step enriches particle diversity.

- *Resampling Step:* The particles are resampled proportional to their weights. If $I^{(k,n+1,p)} \in (1, \ldots, K)$ are particle indices sampled proportional to $\bar{\omega}^{(k,n+1,p)}$, the equally weighted particles are obtained as

$$\bar{\theta}^{(k,n+1,p)} = \bar{\theta}^{(I^{(k,n+1,p)},n+1,p)},$$
$$\bar{\omega}^{(k,n+1,p)} = \frac{1}{K}.$$

- *Move Step:* Each particle is passed through a Markov kernel $\mathcal{K}_{n+1,p}(\bar{\theta}^{k,n+1,p},\cdot)$ that leaves $\bar{\gamma}_{n+1,p}(\theta)$ invariant, typically a Metropolis-Hastings kernel:

  1. Propose $\theta^{*(k)} \sim \mathcal{Q}_{n+1,p}\left(\cdot \,\middle|\, \bar{\theta}^{(k,n+1,p)}\right)$.
  2. Compute the acceptance rate α, where:

$$\alpha = \min\left(1, \frac{\bar{\gamma}_{n+1,p}(\theta^{*(k)})\mathcal{Q}_{n+1,p}\left(\bar{\theta}^{(k,n+1,p)} \middle| \theta^{*(k)}\right)}{\bar{\gamma}_{n+1,p}(\bar{\theta}^{(k,n+1,p)})\mathcal{Q}_{n+1,p}\left(\theta^{*(k)} \middle| \bar{\theta}^{(k,n+1,p)}\right)}\right).$$

  3. With probability α, set $\bar{\theta}^{(k,n+1,p)} = \theta^{*(k)}$, otherwise keep the old particle.

This step will enrich the support of the particle cloud while conserving its distribution. If the particle set is a poor representation of the target distribution, the move step can also help to adjust the location of the support. Crucially, given the importance of the sampling setup, the proposal distribution $\mathcal{Q}_{n+1,p}\left(\cdot \,\middle|\, \bar{\theta}^{(k,n+1,p)}\right)$ can be adapted to the existing particle set.

In the CRI implementation, we define three (or four) NS parameters corresponding to each covariate as one block. A mixture distribution is designed to combine with equal probabilities (1) a block independent normal distribution using the means and the standard deviations derived from the existing particle set and (2) a random walk proposal based on a scaled-down covariance matrix used in the block independent proposal; that is,

$$\theta^{*(k)} \sim \frac{1}{2}\mathcal{N}(\boldsymbol{\mu}, \Sigma) + \frac{1}{2}\mathcal{N}(\bar{\theta}^{(k,n+1,p)}, \Sigma^*),$$

where $\boldsymbol{\mu}$ is the sample mean vector of $\bar{\theta}^{(k,n+1,p)}$ and Σ is the covariance matrix with a block diagonal structure, i.e., the covariances across blocks are all zero. $\sigma^{*2}_{i,j}$, which is the (i,j)-th element of Σ^*, is set to be $(0.2\sigma_{i,j})^2$ (the i,jth element of Σ), to propose around the original values. Mixing the independent and random walk proposals can effectively boost the support (i.e., a higher ESS) by offering local alternatives to those parameters with already high likelihood, especially when there exist discrepancies between the true distribution and its approximating normal distribution.

Moreover, we do not propose to replace an entire parameter particle, and implement a random block proposal. For each particle, say, comprising sixteen blocks, we randomly select a random number of blocks (from five to ten) and only propose new values for the selected blocks, while keeping the remaining blocks at their original values. This design can increase the acceptance rate and still offer rich enough replacements. To ensure a good replacement for every block, we perform multiple such Metropolis-Hastings steps each time until the accumulated acceptance rate exceeds 100% and the ESS reaches at least 75% of sample size.

Finally, proposed particles must satisfy some pre-defined constraints laid out in the CRI Technical Report:

- d must be positive
- particles must produce an increasing or decreasing structure of the NS function for the first five months in order to ensure smooth term structure of the forward intensity parameters

- the coefficients for some covariates, such as liquidity level and trend, are required to be non-positive over all forward starting times

Using the mixture proposal creates a minor complication. The sampler for the truncated values does not carry the same norming constant due to the inclusion of the random walk proposal so that it cannot be ignored in the importance weight. To address the issue, we treat those sampled parameters violating the above mentioned constraints as if there were legitimate particles, but assign the likelihood $\bar{\gamma}_{n+1,p}(\theta^{*(k)})$ of any such proposed particle a value of 0. In short, such particles will never be accepted.

III. At the final tempering step

After reaching the final tempering step (i.e., ξ_p =1), additional moves are performed until the accumulated acceptance rate exceeds 200% instead of 100% at the prior steps. This is to improve the final quality of the SMC sample of parameter particles in representing the target distribution.

IV. Re-initialization

For computing self-normalized statistics for the CRI-PD model's parameters, we rely on a sequential batch-resampling routine of Chopin (2002) in the full-sample run. This SMC approach is the expanding-data SMC technique according to the classification in Duan and Fulop (2015). Although expanding data approach is more computationally efficient, we noticed that approximation errors may sometimes get accumulated after repeatedly updating the SMC parameter particle set by adding data one month at a time. We thus introduce a parameter re-initialization every 10 sequential updating steps to remove the potentially accumulated approximation errors. Apart from the full-sample run, we always re-initialize the SMC parameter particle set in our monthly real-time updates of parameters after incorporating newly arrived data. The monthly real-time updating procedure is described in 1.3.5 of Technical Report (Version: 2017, Update 1).

Re-initialization is the same as the initialization at the beginning of the SMC, except that the relevant means and variances-covariances are computed with the updated SMC parameter particle set so that re-initialization can take advantage of updated information on the sampling distribution.

References

Credit Research Initiative, 2017, "NUS-RMI Credit Research Initiative Technical Report Version: 2017 update 1", National University of Singapore.

Duan, J.C., and Fulop, A., 2015, "Density-Tempered Marginalized Sequential Monte Carlo Samplers", Journal of Business & Economic Statistics, 33, 2, 192-202.

Chopin, N., 2002, "A Sequential Particle Filter Method for Static Models", Biometrika, 89:539–551.