Solution, Filtering and Estimation of Models with the ZLB

This draft merely serves as a summary of the methodology in Boehl (2021), Boehl and Strobel (2020), Boehl et al. (2020) and Boehl and Lieberknecht (2021) and is provided for convenience. Please do not cite!
Latest version: `http://gregorboehl.com/live/obc_boehl.pdf`

Gregor Boehl

University of Bonn

January 3, 2021

1 Introduction

This draft sketches a solution method for linear models with occasionally binding constraints, a nonlinear filter/smoother, and a tempered version of the Ensemble Monte Carlo Markov chain method. Combined, these methods allow the Bayesian estimation of medium and large-scale DSGE models while fully accounting for an occasionally binding zero lower bound (ZLB) on interest rates. The content on the method to solve models with occasionally binding constraints is developed in Boehl (2021), whereas the filtering bits can be found in slightly less detail in Boehl and Strobel (2020). Both are the formal academic reference for these methods.

OBC method

Section 2 summarises the method of Boehl (2021) used to solve for an endogenous zero lower bound on the nominal interest rate as an occasionally binding constraint. While this method shares many features with the algorithm introduced in Guerrieri and Iacoviello (2015), and will return an identical solution (if the solution is unique), it comprises a considerable advantage in computation speed. I propose a transformation of the piecewise-linear dynamic system which allow to provide a closed-form state-space representation for the complete expected trajectories of the endogenous variables as a function of the states and expected duration at the ZLB (the "ZLB spell duration"). Further, I provide two necessary equilibrium conditions. Provided with this representation, finding the expected ZLB spell duration given the state of the economy can be done by simply iterating over the sets of spell durations. Using the closed-form solution together with the

*I am grateful to Alexander Meyer-Gohde, Kenneth Judd, Alexander Richter, Felix Strobel, Carlos Zarazaga and participants of several conferences and seminars for discussions and helpful comments on the contents of this paper. I am indebted to Edward Herbst and Pablo Winant for making their code available.

Email address: `gboehl@uni-bonn.de`, `https://gregorboehl.com`

equilibrium conditions allows to check for a model equilibrium instantaneously instead of simulating a complete anticipated equilibrium path for a given ZLB spell. This enhances the computational speed of the algorithm substantially compared to the methods by Guerrieri and Iacoviello (2015) and Holden (2016). As the outcomes of the presented method and those cited above are identical, I do not compare accuracy but refer to the other papers instead.

Nonlinear filtering

Secondly, I present a Bayesian filter in Section 3 that can handle high-dimension and potentially highly nonlinear laws-of-motion, allowing the filtering of medium and large-scale nonlinear economic models. This is central because accurate parameter estimation requires precise estimates of the likelihood, while still maintaining that the inference of the likelihood must also be very fast. In practice, the latter implies to minimize the number of particle evaluations while respecting the nonlinearity of the transition function. The filter is introduced in Boehl and Strobel (2020) to the economic literature, where I also provide a short discussion and comparison of asymptotic properties.

A generic Bayesian filter allows for uncertainty about initial states and potential measurement errors. For this purpose, we use the Ensemble Kalman Filter (EnKF) introduced in Evensen (1994), which is a hybrid of the particle filter and Kalman filter technology. Similar to the particle filter, a set of points (the *ensemble*) is sent through the transition function during the prediction step. However, instead of re-sampling (as with the particle filter), the EnKF approximates a state-dependent system matrix which can be used within a Kalman-like updating step. The EnKF allows to efficiently approximate the state distribution of large-scale nonlinear systems with only a few hundred particles.

Counterfactual simulations, which are important for economic and econometric analysis, require that the smoothened series of shocks can exactly reproduce the filtered data. I propose a procedure of nonlinear path-adjustment to calculate the smoothed/historic shock innovations, building two steps on top of the EnKF: the first step is an ensemble version of the Rauch-Tung-Striebel smoother (Rauch et al., 1965; Raanes, 2016). In a second step, iterative global optimization methods are used to maintain that the shock innovations fully respect the nonlinear transition function while taking the approximated distribution of smoothed states into account.

Why another nonlinear filter?

There is a growing literature on applying particle filters (also called *Sequential Monte Carlo methods*) to economic models and data (see e.g. An and Schorfheide, 2007; Fernández-Villaverde and Rubio-Ramírez, 2007; Herbst and Schorfheide, 2019). These methods are relatively simple to implement but require an extremely high number of particles (i.e. transition function evaluations). For the benchmark model of Smets and Wouters (2007) estimates of the number of necessary particles range from at least 40.000 particles in Herbst and Schorfheide (2019) to – more realistically due to the *curse of dimensionality* – about 1.500.000 particles as in Gust et al. (2017). There is not yet a consensus on the necessary number of particles depending on the dimensionality of the problem. Additionally to the problem of high computational costs, Binning and Maih, 2015 argue that particle filter methods can further be subject to numeric instability.

A potential resolution to the problem is the Unscented Kalman Filter (UKF, Julier et al., 2000). The UKF relies on a deterministic sampling technique, the so-called Sigma points, that aims to minimize the number of points necessary at each iteration. These points are then propagated through the true nonlinear transition function, then mean and covariance are calculated analogously to the linear Kalman filter (this is called the *unscented transform*).[1] Andreasen (2013) document that Sigma point filters as the one discussed further below are not only much faster but can also be more accurate than particle filters. The quality of the estimates however crucially depends on the choice of appropriate Sigma points that represent the nonlinearity of the dynamic system sufficiently well. Inherent to the concept, the approximation for each direction in a certain dimension relies on only one sample point. If this point is chosen badly while the transition function is strongly convex in at least one dimension – as for instance it is the case with the ZLB – the gradient will be overestimated and the respective step in the state will be understated. The opposite occurs for concave transition functions.

Hence, in general the UKF tends to overestimate the tails and in the worst case, the filter will simply diverge. This means that the quality of the filtering results crucially depends on the parameterization of the Sigma points. Unfortunately, a calibration that yields reasonable precise estimates for one draw from the parameter distribution must not necessarily do so for a different draw. Further problematic is the filters' dependence on the matrix square root for find the optimal sigma points. This problem deteriorates when the covariance matrix is close to singular, which is quite likely the case for models with occasionally binding constraints.[2] Compared to the UKF, the EnKF does not rely on parameterized deterministic sampling techniques and is hence parameter-free.

Cuba-Borda et al. (2019), drawing on Fair and Taylor (1980), propose an *inversion filter* for estimation and filtering of shock innovations, which can be understood as solving the nonlinear one-to-one mapping from shocks and observables. This filter does not allow for uncertainty about initial states or potential measurement errors and leaves no leeway in interpreting the data given such uncertainty specifications. This has the drawback that bad initial values or moderate jumps in the observables can result in large approximation errors. A *learning period*, as suggested by some authors, will not change this property as, in the absence of potential measurement errors, the course of the dynamics is deterministic. Especially ignoring uncertainty about the initial distribution of states can have severe effects on the economic interpretation of the results. Correspondingly, I find that the estimation of the likelihood only based on the inferred innovations of exogenous variables does not suffice as an approximation of the data-likelihood given the full distribution of states. Note that the filter can potentially be numerically unstable because it is based on local optimization methods.

While it can be shown that the particle filter is asymptotically unbiased when the number of samples goes to infinity, this is not the case for the EnKF. The approximation of the state distribution – and thereby the inference of the likelihood – is based on a linear approximation around the mean of the distribution. This also implies that for linear systems the EnKF is in fact asymptotically unbiased and identical to the standard Kalman Filter and can handle any assumptions regarding the initial state distribution and measurement errors. The inversion filter is only unbiased given very specific assumption on the initial states and measurement errors: it is unbiased if all initial states are known

[1] Technical details can be found in Julier et al. (2000), Wan and Van Der Merwe (2000) and Julier (2002).

[2] If the constraint binds, the derivative with respect to the constrained variable is zero.

with certainty, and in the absence of measurement errors. Especially the first assumption – initial states are known with certainty – is arguably quite strong, and might render the use of this filter for likelihood inference and smoothing of standard medium scale models at least problematic.

Sampling

In Section 4 I also briefly comment on sampling methods, in particular the Differential evolution Ensemble-MCMC method (Ter Braak, 2006; ter Braak and Vrugt, 2008, DE-MCMC). I find that the adaptation of such rather advanced methods is necessary for three reasons. First, DSGE models generally have a very rugged likelihood landscape, with large regions of indeterminacy and other regions where the likelihood function is almost flat. Secondly, with nonlinear models an additional problem constitutes trough the process of nonlinear filtering, which is generally also based on sampling. For this reason, likelihood estimates will, to some degree, be noisy. Neither "classic" samplers, nor local optimization methods are well suited to deal with such problems reliably, in particular given that economic models can have quite high dimensional parameter spaces. Global optimizers do a better job, but are prone to errors. In my experience, many estimations based on local optimization and MH sampling are not robust to either initial values of the optimizer or the choice of the optimizer. Lastly, medium scale economic models at the ZLB can be subject to reversal effects (see e.g. Carlstrom et al. (2015)) or do not allow for ZLB equilibria given certain shock sizes for other reasons. The solution method proposed here correctly specifies that reversals (in the spirit of Carlstrom et al., 2015) cannot be a rational expectations equilibrium, but this goes at the cost of a further increase of the parameter region for which the model cannot be solved, which in turn is challenging when sampling from the parameter space distribution. An additional advantage of Ensemble-MCMC samplers is that the parallelization of ensemble-based methods is very straightforward.

Reference implementation

My reference implementation of the method and the filter enables to effectively estimation large-scale DSGE models in about two to three hours for the benchmark of Smets and Wouters (2007).[3] Note that computational advantage much depends on efficient implementation. The reference implementation is written in the powerful and freely available language `Python`.[4] The method and filter – together with a parser and additional econometric tools – are implemented in the `pydsge` package which is available at `https://github.com/gboehl/pydsge`.

[3] Benchmark taken on a machine with 40 cores of 3.10GHz each.

[4] `Python` can provide speed benchmarks that are en-par with compiled languages such as Fortran while comprising the advantages of a high-level programming language. I would like to promote free and open software and advocate the avoidance of proprietary languages. Open source alternatives already provide by far more efficient and more flexible environments while avoiding barriers for scientific advancing such as licensing and closed-source code.

2 Solution Method

This section presents the solution method. The model can be cast in the form

$$E_t x_{t+1} = N x_t + h \max\left\{p E_t x_{t+1} + m x_t, \bar{r}\right\}, \tag{1}$$

with $x_t = \begin{vmatrix} v_t \\ w_{t-1} \end{vmatrix}$, where v_t is the vector of forward looking variables and w_{t-1} are the states updated by the time-t shocks.[5] N is the system matrix and h contains the time-t coefficients of the constraint. The vectors p and m represent the constraint equation, which in our case is the equation defining the notational rate. This reads $r_t = \max\{p x_{t+1} + m x_t, \bar{r}\}$. Villemot et al. (2011) provides the means to cast any dynamic system in the form $A x_{t+1} + B x_t = 0$. Matrix A can be inverted e.g. by applying the singular value decomposition and substituting out static variables.

Denote the system in which the constraint is slack (the *unconstrained system*) as

$$E_t x_{t+1} = \hat{N} x_t, \tag{2}$$

with

$$\hat{N} = (I - h \otimes p)^{-1} (N + h \otimes m), \tag{3}$$

and note that it is always possible to find an invertible $(I - h \otimes p)$ by multiplying m, p and $\bar{r}$ by an appropiate scalar while at the same time dividing h by the same scalar.

I will first outline the solution method, taking the durations for which the constraint holds as given. Then I present a simple iteration scheme to endogenize the expected durations for discrete expectations. Let us first assume that the constraint binds in the current period t. System (1) can be rewritten as

$$E_t x_{t+1} = \begin{cases} \hat{N} x_t & \forall\, p E_t x_{t+1} + m x_t - \bar{r} \geq 0 \\ \\ N x_t + g \bar{r} & \forall\, p E_t x_{t+1} + m x_t - \bar{r} < 0. \end{cases} \tag{4}$$

Let k be the expected ZLB spell in period t. Denote the desired rational expectations solution to (1) given k and the state variables w_t as the function S such that

$$v_t = S(k, w_{t-1}). \tag{5}$$

I will use $S(k)$ as a shorthand notation where w_{t-1} are understood. Also, denote as $x_t|k$ the solution conditional on expecting the constraint to hold for k periods. For the unconstrained system $\hat{N}$, $S(0, w_{t-1}) = v_t$ can be found using familiar methods like the QZ-decomposition as suggested by Klein (2000). Denote this (linear) solution by $S(0) = \Omega$:

$$v_t = \Omega w_{t-1} \quad \forall\, p E_t x_{t+1} + m x_t > \bar{r} \tag{6}$$

[5] See e.g. Rendahl (2017) on how manipulate the system such that the exogenous shocks are directly included in the state vector.

For $\Psi = \begin{vmatrix} I & -\Omega \end{vmatrix}$, Equation (6) implies that

$$E_t \left\{ \Psi \begin{vmatrix} v_{t+k+1} \\ w_{t+k} \end{vmatrix} \right\} = 0 \quad \forall\, pE_t x_{t+k+1} + m x_{t+k} \geq \bar{r}, \tag{7}$$

i.e. for every future period $t + k$ in which the system is expected to be unconstrained.

Now assume that the constraint binds at time t and will continue to do so until period $t + k$. Iterating System (4) forward yields

$$E_t \left\{ \begin{vmatrix} v_{t+k} \\ w_{t+k-1} \end{vmatrix} \right\} = N^k x_t + (I - N)^{-1}(I - N^k)h\bar{r}, \tag{8}$$

where $(I - N)^{-1}(I - N^k) = \sum_{i=0}^{k-1} N^i$ is the transformation for a geometric series of matrices. Finally, I can combine Equations (7) and (8) to find a solution of type (5) of the endogenous variables v_t in terms of the state variables w_{t-1} given k:

$$S(k, w_{t-1}) = \left\{ v_t : \Psi N^k \begin{vmatrix} v_t \\ w_{t-1} \end{vmatrix} := -\Psi(I - N)^{-1}(I - N^k)h\bar{r} \right\}. \tag{9}$$

Since h is a vector of constants, the whole RHS of (9) is given.

Let us now relax the assumption that the constraint holds immediately in time t. This case is in particular relevant for models with persistent endogenous state variables. It is straightforward to take Equation (9) as a starting point, and to allow for a number of periods l in the unconstrained system $\hat{N}$ *until* the system is at the constraint:

$$S(k, l, w_{t-1}) = \left\{ v_t : \Psi N^k \hat{N}^l \begin{vmatrix} v_t \\ w_{t-1} \end{vmatrix} = -\Psi(I - N)^{-1}(I - N^k)h\bar{r} \right\}. \tag{10}$$

Using Equations (8) and (10) I can express the expectations on the state conditional on (l, k) of the economy in period s, $E_t x_s|(l, k)$, as the function L with

$$\begin{aligned} E_t x_s|(l, k) = L_s(l, k, w_{t-1}) = & N^{\max\{s-l,0\}} \hat{N}^{\min\{l,s\}} \begin{vmatrix} S(l, k, w_{t-1}) \\ w_{t-1} \end{vmatrix} \\ & + (I - N)^{-1}(I - N^{\max\{s-l,0\}})h\bar{r}. \end{aligned} \tag{11}$$

Note that $L_1(0, 0, w_{t-1}) = \begin{vmatrix} \Omega \\ I \end{vmatrix} w_{t-1}$ is the generic solution to the unconstrained system.

Solving for (l, k)

Let us again first treat the simpler case in which I assume that any shock causes the constraint to bind immediately in time t (the *non-transitory* case). The following proposition summarizes the conditions for (x_t, w_{t-1}, k) to be a rational expectations equilibrium:

Proposition 1 (non-transitory equilibrium)**.** *Assuming non-transition, a number of expected periods k at the constraint is a rational expectations equilibrium iff*

$$pE_t[x_{t+1}|k^*] + m x_t|k^* \geq \bar{r} > pE_t[x_{t+1}|k] + m x_t|k \tag{12}$$

for all $k^* > k \geq 0$, hence if in expectations the system is constrained for exactly k^* periods.

Let us proceed to the case where agents expect the unconstrained system to prevail for some transition time before the constraint binds for k periods. Using specification (11), Definition 2 summarizes the respective equilibrium conditions.

Proposition 2 (transitory equilibrium)**.** *A pair (l^*, k^*) is a rational expectations equilibrium iff*

$$pE_t[x_s|k^*] + mx_s|k^* \geq \bar{r} \quad \forall s < l^* \wedge s \geq k^* + l^* \tag{13}$$

and

$$pE_t[x_s|k^*] + mx_s|k^* < \bar{r} \quad \forall l^* \leq s < k^* + l^*. \tag{14}$$

In other words, (l, k) are part of an equilibrium, if in expectations, the constraint starts binding *exactly* in period $t + l$ and ends to bind *exactly* in period $t + l + k$.

Unfortunately there is no closed form solution for (l, k) given w_{t-1}. A set of (l, k) that satisfies Theorem 2 must be found using an iterative scheme. As this constitutes an iterative scheme on an integer domain, a theoretical assessment is difficult because most theoretical work on similar algorithms deals with real valued functions. While there are limits to the assessment of whether equations (13) and (14), given w_{t-1} have any solution, some insights regarding the existence and uniqueness of such solutions on such solutions are provided by Holden (2017).

The specification of the equilibrium conditions, although arguably slightly more formal, is similar to the one in Guerrieri and Iacoviello (2015). The crucial advantage of the formulation here is the closed form expression of $E_t[x_{t+s}|(l, k)]$.

An optimal iterative scheme must be hand-tailored to the problem. For the purpose of the estimation of large-scale DSGE models, in which the constraint is the zero lower bound on nominal interest rates, I use the following iterative scheme:

```
l, k = 0, 0
for l in range(l_max):
    if b L(l, 0, l, v) - r_bar < 0:
        # break loop since constraint binds
        break
    if l is l_max - 1:
        # return that l=k=0 is an equilibrium
        return 0, 0
...
```

If this is the case, exit. Otherwise assume $k > 0$ and iterate over l and k until the equilibrium conditions in (12), (13) and (14) are satisfied.

```
...
for l in range(l_max):
    for k in range(1, k_max):
        if l:
            if b L(l, k, 0, v) - r_bar < 0:
                continue   # continue skips the inner loop
            if b L(l, k, l-1, v) - r_bar < 0:
```

```
            continue
        if b L(l, k, k+l, v) - r_bar < 0:
            continue
        if b L(l, k, l, v) - r_bar > 0:
            continue
        if b L(l, k, k+l-1, v) - r_bar > 0:
            continue
        # if we made it here, this must be an equilibrium
        return l, k
# if the loop went though without finding an equilibrium, throw a warning
warn('No equilibrium exists!11')
```

This scheme is very efficient for the specific problem because in more than 50% of the cases, the method will already exit in the first loop because the ZLB is not binding and not expecting to bind in the near future. If it does not exit, than for post-2008 data points it is predominantly the case that the ZLB already is binding. In this case $l = 0$ and only k is to be determined. As, according to the Primary Dealer Survey, most market participants expected the ZLB to be binding for about eight quarters, the procedure will on average need 8 guesses (plus three for the first loop) until an equilibrium is found. l will normally only be positive in 2008, when the economy is not yet at the ZLB, but the respective shocks, that trigger a binding ZLB in later periods, have already materialized.

While the above procedure is taylored to work most efficiently in the context of estimating DSGE models with the ZLB, it is generic and applicable to any sort of constraint. The resulting transition function is linear for the region where the ZLB does not bind and (increasingly) nonlinear when it binds. For the model presented here, the implementation in the `pydsge` package will find the state-space representation for about 200.000 particles draws per second and CPU.

3 Nonlinear Filtering

In this section I briefly summarize the nonlinear filtering methodology, which is an adaptation of the Ensemble Kalman Filter (Evensen, 1994, EnKF) for the general type of nonlinear problems faced in macroeconomics. Denote a (potentially nonlinear) hidden Markov-Model (HMM) by

$$x_t = g(x_{t-1}, \varepsilon_t) \tag{15}$$

$$z_t = h(x_t) + \nu_t \tag{16}$$

with $\varepsilon_t \sim \mathcal{N}(0, Q)$ and $\nu_t \sim N(0, R)$. Let $\mathbf{X}_t = [\mathbf{x}_t^1, \cdots, \mathbf{x}_t^N] \in \mathbb{R}^{n \times N}$ be the ensemble at time t, which consists of N vectors of the state. Further denote by $(\bar{x}_t, P_t)$ the mean and the covariance matrix of the unconditional distribution of states for period t. Initialize the ensemble by sampling N times from the prior distribution

$$\mathbf{X}_0 \overset{N}{\sim} \mathcal{N}(\bar{x}_0, P_0). \tag{17}$$

It is advisable to use a low-discrepancy sequence (or *quasirandom sequence*) such as the Sobol (1967) sequence for sampling the initial distribution. In our papers we use latin hypercube sampling as proposed by McKay et al. (2000). I find that, across quasirandom methods, it provides the smallest approximation error of the likelihood approximation across random seeds.

Step 1: Predict

Predict the prior-ensemble $\mathbf{X}_{t|t-1}$ at time t by applying the transition function to the posterior ensemble from last period. Use the observation function to obtain a prior-ensemble of observables:

$$\mathbf{X}_{t|t-1} = g(\mathbf{X}_{t-1|t-1}, \boldsymbol{\varepsilon}_t), \tag{18}$$

$$\mathbf{Z}_{t|t-1} = h(\mathbf{X}_{t|t-1}) + \boldsymbol{\nu}_t, \tag{19}$$

where $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\nu}_t$ are each N realizations drawn from the respective distributions.

Step 2: Update

Denote by $\bar{\mathbf{X}}_t = \mathbf{X}_t(\mathbf{I}_N - \mathbf{1}\mathbf{1}^\intercal/N)$ the *anomalies* of the ensemble, i.e. the deviations from the ensemble mean. Recall that the covariance matrix of the prior distribution at t is $\frac{\bar{\mathbf{X}}_t \bar{\mathbf{X}}_t^\intercal}{N-1}$. The Kalman mechanism then yields an *update*-step of

$$\mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \bar{\mathbf{X}}_{t|t-1}\bar{\mathbf{Z}}_{t|t-1}^\intercal \left(\bar{\mathbf{Z}}_{t|t-1}\bar{\mathbf{Z}}_{t|t-1}^\intercal\right)^{-1} \left(z_t \mathbf{1}^\intercal - \mathbf{Z}_{t|t-1}\right). \tag{20}$$

The mechanism is similar to the unscented Kalman filter (UKF) UKF but with particles instead of deterministic Sigma points, and statistical linearization instead of the unscented transform. The advantage of the EnKF over the UKF is that its output does not depend on the parametrization of the filter. Conceptionally this procedure can hence be seen as a *transposition* of the EnKF.[6]

The likelihood at each iteration can be then determined by

$$ll_t = \varphi\left(z_t | \bar{z}_t, \frac{\bar{\mathbf{Y}}_t \bar{\mathbf{Y}}_t^\intercal}{N-1} + R\right) \tag{21}$$

3.1 *Smoothing and iterative path-adjusting*

For economic analysis we are also interested in the series of shocks, $\{\varepsilon_t\}_{t=0}^{T-1}$, that fully recovers the mode of the smoothened states. The econometric process of using *all* available information on *all* estimates is called smoothing. For this purpose, we employ the Rauch-Tung-Striebel smoother (Rauch et al., 1965) in its Ensemble formulation similar to Raanes (2016).

Denote by T the period of the last observation available and update each ensemble according to the backwards recursion[7]

$$\mathbf{X}_{t|T} = \mathbf{X}_{t|t} + \bar{\mathbf{X}}_{t|t}\bar{\mathbf{X}}_{t+1|t}^{+}\left[\mathbf{X}_{t+1|T} - \mathbf{X}_{t+1|t}\right]. \tag{23}$$

[6] Notationally both are equivalent. The regular EnKF assumes the size of the state spaces to be larger than N, and accordingly the term $\left(\bar{\mathbf{Z}}_{t|t-1}\bar{\mathbf{Z}}_{t|t-1}^\intercal\right)$ to be rank deficient. The mechanism then builds on the properties of the pseudoinverse (the latter provides a least squares solution to a system of linear equations), which is used instead of the regular matrix inverse.

[7] Although it is formally correct that

$$\bar{\mathbf{X}}_{t|t}\bar{\mathbf{X}}_{t+1|t}^\intercal \left(\bar{\mathbf{X}}_{t+1|t}\bar{\mathbf{X}}_{t+1|t}^\intercal\right)^{+} = \bar{\mathbf{X}}_{t|t}\bar{\mathbf{X}}_{t+1|t}^{+}, \tag{22}$$

This creates a series $\left\{\mathbf{X}_{t|T}\right\}_{t=0}^{T}$ of representatives of the distributions of states at each point in time, reflecting all the available information. We now want to ensure that the mode of the distribution fully reflects the nonlinearity of the transition function while retaining a reasonably good approximation of the full distribution. We call this process *nonlinear path-adjustment*. It is important that the smoothened distributions are targeted instead of, e.g., just the distributions of observables and shocks. Only when the full smoothened distributions are targeted it can be maintained that *all* available information from the observables is taken into account. This procedure implicitly assumes that the smoothened distributions approximate the actual transition function sufficiently well and only minor adjustments remain necessary. Since in general there are (many) more states than exogenous shocks, the fitting problem is underdefined and matching precision will depend on the size of the relative (co)variance of each variable. Small observation errors lead to small variances around observable states and tight fitting during path-adjustment while loosely identified states grant more leeway.

Initiate the algorithm with $\hat{x}_0 = \mathrm{E}\mathbf{X}_{0|T}$ (the mean vector over the ensemble members), define $P_{t|T} = \mathrm{Cov}\{\mathbf{X}_{t|T}\}$ and for each period t recursively find

$$\hat{\varepsilon}_t = \arg\max_{\varepsilon} \left\{ \log f\left(g(\hat{x}_{t-1}, \varepsilon) | \bar{x}_{t|T}, P_{t|T} \right) \right\}, \tag{24}$$

$$\hat{x}_t = g(\hat{x}_{t-1}, \hat{\varepsilon}_t), \tag{25}$$

which can be done using standard iterative methods. For our papers I use the Covariance matrix adaptation evolution strategy (Hansen, 2006, CMA-ES). This global heuristic optimizer has a fairly good track record of converging to a global optimum rather quickly.

The resulting series of $\hat{x}_t$ corresponds to the estimated mode given the initial mean and approximated covariances and is completely recoverable by $\hat{\varepsilon}_t$. Naturally, it represents the nonlinearity of the transition function while taking all available information into account. Since the deviation between mode $\hat{x}_t$ and mean $\bar{x}_t$ is in general marginal, I refer to

$$\{\hat{x}_t, P_t\}_{t=0}^{T} \tag{26}$$

as the *path-adjusted smoothed distributions*.[8]

4 Sampling

For posterior sampling methods that rely on mode maximization must be avoided, as these prone to get stuck in local maxima. Even if the global mode could be found using respective optimization techniques, some odd-shaped likelihood functions might not allow the MCMC algorithm to fully explore the posterior distribution e.g. because the posterior is bimodal, because it is flat with many local spikes, or because of intersections where no likelihood can be evaluated. I propose a tempering extension to the differential

the implementation using the LHS of this equation is numerically more stable when using standard implementations of the *pseudo-inverse* based on the SVD.

[8] Unfortunately the adjustment step can not be done during the filtering stage already. Iterative adjustment before the prediction step, would bias the transition of the covariance. Likewise, adjusting after the prediction step will require the repeating the prediction and updating step leading to a potentially infinite loop. See e.g. Ungarala (2012) for details.

evolution Monte carlo Marcov chain (DE-MCMC) suggested by Ter Braak (2006) and ter Braak and Vrugt (2008). Tempering can be done in many ways, e.g. one could folow the lines of Herbst and Schorfheide (2014). The tempering algorithm used in our papers builds on top of the implementation of Goodman and Weare (2010).

The DE-MCMC method is a class of ensemble MCMC methods. Instead of using a single or small number of chains that are state dependent (as e.g. in the Metropolis algorithm), ensemble samplers use a large number of chains (the "ensemble"). For each iteration, proposals are generated based on the current state of the full ensemble. These methods are hence self-tuning and do – if at all – only require hyper parameters. Another advantage of ensemble samplers is that massive parallelization is straightforward. Ensemble methods have been extensively applied, in particular in the field of astrophysics.

Ensemble initialization can be done in several ways. Goodman and Weare (2010) suggest to initialize it as a small ball around some initial value. This however bears the risk that the ensemble can not fully unfold due to odd or irregularly shaped posteriors. I suggest to initialize the ensemble with the prior distribution. This will put equal initial weight to each region of the parameter space. In order for particles not to not leave low-density regions too early, a tempering scheme can be used. The temperature λ here is the weight of the likelihood for the posterior. A $\lambda = 0$ posterior hence is identical to the prior. A gradual increase of λ towards one allows to also sample odd-shaped or near-flat distributions as some particles will remain in low-likelihood-high-prior regions for an extended period of time. Initializing the ensemble with the prior distribution ensures that the full relevant parameter space is considered, independently of multi-modality or possible discontinuities.

References

An, S., Schorfheide, F., 2007. Bayesian analysis of dsge models. Econometric reviews 26, 113–172.

Andreasen, M.M., 2013. Non-linear dsge models and the central difference kalman filter. Journal of Applied Econometrics 28, 929–955.

Binning, A., Maih, J., 2015. Sigma point filters for dynamic nonlinear regime switching models. Technical Report.

Boehl, G., 2021. Efficient Solution and Computation of Models with Occasionally Binding Constraints. Technical Report. URL: `https://gregorboehl.com/live/obc_boehl.pdf`.

Boehl, G., Goy, G., Strobel, F., 2020. A structural investigation of quantitative easing. IMFS Working Paper Series 142. Goethe University Frankfurt, Institute for Monetary and Financial Stability (IMFS).

Boehl, G., Lieberknecht, P., 2021. The Hockey Stick Phillips Curve and the Zero Lower Bound. Technical Report. URL: `https://gregorboehl.com`.

Boehl, G., Strobel, F., 2020. US business cycle dynamics at the zero lower bound. IMFS Working Paper Series 143. Goethe University Frankfurt, Institute for Monetary and Financial Stability (IMFS).

ter Braak, C.J., Vrugt, J.A., 2008. Differential evolution markov chain with snooker updater and fewer chains. Statistics and Computing 18, 435–446.

Carlstrom, C.T., Fuerst, T.S., Paustian, M., 2015. Inflation and output in new keynesian models with a transient interest rate peg. Journal of Monetary Economics 76, 230–243.

Cuba-Borda, P., Guerrieri, L., Iacoviello, M., Zhong, M., 2019. Likelihood evaluation of models with occasionally binding constraints. Journal of Applied Econometrics 34, 1073–1085.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. Journal of Geophysical Research: Oceans 99, 10143–10162.

Fair, R.C., Taylor, J.B., 1980. Solution and Maximum Likelihood Estimation of Dynamic Nonlinear RationalExpectations Models. Technical Report. National Bureau of Economic Research.

Fernández-Villaverde, J., Rubio-Ramírez, J.F., 2007. Estimating macroeconomic models: A likelihood approach. The Review of Economic Studies 74, 1059–1087.

Goodman, J., Weare, J., 2010. Ensemble samplers with affine invariance. Communications in applied mathematics and computational science 5, 65–80.

Guerrieri, L., Iacoviello, M., 2015. Occbin: A toolkit for solving dynamic models with occasionally binding constraints easily. Journal of Monetary Economics 70, 22–38.

Gust, C., Herbst, E., López-Salido, D., Smith, M.E., 2017. The empirical implications of the interest-rate lower bound. American Economic Review 107, 1971–2006.

Hansen, N., 2006. The cma evolution strategy: a comparing review, in: Towards a new evolutionary computation. Springer, pp. 75–102.

Herbst, E., Schorfheide, F., 2014. Sequential monte carlo sampling for dsge models. Journal of Applied Econometrics 29, 1073–1098.

Herbst, E., Schorfheide, F., 2019. Tempered particle filtering. Journal of Econometrics 210, 26–44.

Holden, T.D., 2016. Computation of solutions to dynamic models with occasionally binding constraints. Technical Report.

Holden, T.D., 2017. Existence and uniqueness of solutions to dynamic models with occasionally binding constraints. Technical Report.

Julier, S., Uhlmann, J., Durrant-Whyte, H.F., 2000. A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Transactions on automatic control 45, 477–482.

Julier, S.J., 2002. The scaled unscented transformation, in: American Control Conference, 2002. Proceedings of the 2002, IEEE. pp. 4555–4559.

Klein, P., 2000. Using the generalized schur form to solve a multivariate linear rational expectations model. Journal of economic dynamics and control 24, 1405–1423.

McKay, M.D., Beckman, R.J., Conover, W.J., 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 42, 55–61.

Raanes, P.N., 2016. On the ensemble rauch-tung-striebel smoother and its equivalence to the ensemble kalman smoother. Quarterly Journal of the Royal Meteorological Society 142, 1259–1264.

Rauch, H.E., Striebel, C., Tung, F., 1965. Maximum likelihood estimates of linear dynamic systems. AIAA journal 3, 1445–1450.

Rendahl, P., 2017. Linear time iteration. Technical Report. IHS Economics Series.

Smets, F., Wouters, R., 2007. Shocks and frictions in us business cycles: A bayesian dsge approach. American Economic Review 97, 586–606.

Sobol, I.M., 1967. On the distribution of points in a cube and the approximate evaluation of integrals. Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki 7, 784–802.

Ter Braak, C.J., 2006. A markov chain monte carlo version of the genetic algorithm differential evolution: easy bayesian computing for real parameter spaces. Statistics and Computing 16, 239–249.

Ungarala, S., 2012. On the iterated forms of kalman filters using statistical linearization. Journal of Process Control 22, 935–943.

Villemot, S., et al., 2011. Solving rational expectations models at first order: what Dynare does. Technical Report. Citeseer.

Wan, E.A., Van Der Merwe, R., 2000. The unscented kalman filter for nonlinear estimation, in: Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000, Ieee. pp. 153–158.