CQE Center for Quantitative Economics

# Higher-order statistics for DSGE models

Willi Mutschler[†]

43/2015

[†] Department of Economics, University of Münster, Germany

wissen.leben
WWU Münster

# Higher-order statistics for DSGE models

Willi Mutschler*

*Center for Quantitative Economics, Westfälische Wilhelms-Universität Münster*

**Abstract**

This note derives closed-form expressions for unconditional moments, cumulants and polyspectra of order higher than two for linear and nonlinear (pruned) DSGE models. The procedures are demonstrated by means of the Smets and Wouters (2007) model (first-order approximation), the An and Schorfheide (2007) model (second-order approximation) and the canonical neoclassical growth model (third-order approximation). Both the Gaussian as well as Student's t-distribution are considered as the underlying stochastic process. Useful Matrix tools and computational aspects are also discussed.

*Keywords:* higher-order moments, cumulants, polyspectra, nonlinear DSGE, pruning
*JEL:* C10, C51, E1

☆This paper was presented at the 3rd Rimini Time Series Econometrics Workshop and the 11th annual DYNARE conference. The author thanks seminar participants and Mark Trede for helpful comments and encouragement.

*Corresponding author. Center for Quantitative Economics, Westfälische Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster, Tel.: +49-251-83-22914, Fax: +49-251-83-22012, Email: willi@mutschler.eu.

## 1. Introduction

Since a Gaussian process is completely characterized by its first two moments, most linear DSGE models focus on Gaussian innovations for simplicity.[1] If, however, we relax linearity or use non-Gaussian innovations, it is natural to analyze whether we are able to exploit information from higher-order moments for the calibration, estimation and identification of parameters. Researchers in mathematics, statistics and signal processing have developed tools, called higher-order statistics (HOS), to solve detection, estimation and identification problems when the noise source is non-Gaussian or we are faced with nonlinearities; however, applications in the macroeconometric literature are rather sparse.[2] The basic tools of HOS are cumulants, which are defined as the coefficients in the Taylor expansion of the log moment generating function in the time-domain; and polyspectra, which are defined as Fourier transformations of the cumulants in the frequency-domain.

In this note, we derive closed-form expressions for unconditional third- and fourth-order moments, cumulants and corresponding polyspectra for non-Gaussian or nonlinear DSGE models. We limit ourselves to fourth-order statistics, since third-order cumulants and the bispectrum capture nonlinearities (or non-Gaussianity) for a skewed process, whereas the fourth-order cumulants and the trispectrum can be used in the case of a non-Gaussian symmetric probability distribution. Regarding the approximation of nonlinear DSGE models we focus on the pruning scheme proposed by Kim et al. (2008) and operationalized by Andreasen et al. (2014), since the pruned state-space (PSS from now on) is a linear, stationary and ergodic state-space system.[3] In the PSS, however, Gaussian innovations do not imply Gaussian likelihood, leaving scope for higher-order statistics to capture information from nonlinearities and non-Gaussianity. In the following exposition we limit ourselves to Taylor approximations and pruning up to third-order, since an extension beyond third-order is - apart from tedious notation - straightforward.

[1] Two notable exceptions are Curdia et al. (2014) and Chib & Ramamurthy (2014).

[2] For introductory literature and tutorials on HOS, see the textbooks of Brillinger (2001), Nikias & Petropulu (1993), Priestley (1983) and the references therein.

[3] Pruning may seem an ad-hoc procedure, however, it can also be theoretically founded as a Taylor expansion in the perturbation parameter (Lombardo & Uhlig, 2014) or on an infinite moving average representation (Lan & Meyer-Gohde, 2013). Schmitt-Grohé & Uribe (2004) also implicitly use pruning in their code to compute moments for a second-order approximation.

Accordingly, we demonstrate our procedures by means of the Smets & Wouters (2007) model for a first-order approximation, the An & Schorfheide (2007) model for a second-order approximation and the canonical neoclassical growth model for a third-order approximation. For all models we consider both the Gaussian as well as Student's t distribution as the underlying shock process and compare our theoretical results with simulated higher-order moments. We focus particularly on skewness and excess kurtosis in our simulations, since these are typical measures an applied researcher would like to match in a calibration exercise. On the other hand auto- and cross-skewness as well as kurtosis may contain valuable information in an estimation exercise. Our DYNARE code is model-independent and can be found on the homepage of the author.

## 2. Higher-order statistics for linear time-invariant state-space systems

Consider the linear time-invariant state-space system

$$\widetilde{z}_{t+1} = A\widetilde{z}_t + B\xi_{t+1} \tag{1}$$

$$\widetilde{y}_{t+1} = C\widetilde{z}_t + D\xi_{t+1} \tag{2}$$

with states $z_t$, controls $y_t$ and stochastic innovations $\xi_t$. A tilde denotes deviations from the unconditional mean, e.g. $\widetilde{y}_t := y_t - E(y_t)$. For the sake of notation, we assume that all control variables are observable. The vector of innovations $\xi_t$ has $E(\xi_t) = 0$ and finite covariance matrix $E(\xi_t\xi_t') =: \Sigma_\xi$. Furthermore, $\xi_t$ is kth-order white noise with finite higher-order moments, which implies $y_t$ is a $k$th-order stationary process.[4]

Formally, the $k$th-order (k=2,3,4) cumulants of the $k$th-order stationary, mean zero vector process $\widetilde{z}_t$ $(t_1, t_2, t_3 \geq 0)$ are given by the $n_z^k$ vectors $\mathcal{C}_{k,z}$ as

$$\begin{aligned}
\mathcal{C}_{2,z}(t_1) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1}], \\
\mathcal{C}_{3,z}(t_1, t_2) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1} \otimes \widetilde{z}_{t_2}], \\
\mathcal{C}_{4,z}(t_1, t_2, t_3) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1} \otimes \widetilde{z}_{t_2} \otimes \widetilde{z}_{t_3}] - \mathcal{C}_{2,z}(t_1) \otimes \mathcal{C}_{2,z}(t_2 - t_3) \\
&\quad - P'_{n_z}\left(\mathcal{C}_{2,z}(t_2) \otimes \mathcal{C}_{2,z}(t_3 - t_1)\right) - P_{n_z}\left(\mathcal{C}_{2,z}(t_3) \otimes \mathcal{C}_{2,z}(t_1 - t_2)\right),
\end{aligned}$$

[4] This is basically an extension of the usual covariance stationarity assumption. See also Priestley (1983, p. 105) for a formal definition of stationary up to order k.

where $P_{n_z} = I_{n_z} \otimes U_{n_z^2 \times n_z}$ and $U_{n_z^2 \times n_z}$ is a $(n_z^3 \times n_z^3)$ permutation matrix with unity entries in elements $[(i-1)n_z + j, (j-1)n_z^2], i = 1, \ldots, n_z^2$ and $j = 1, \ldots, n_z$, and zeros else. Here, we adopt the compact notation of Swami & Mendel (1990) and store all product-moments of a mean zero vector-valued process in a vector using Kronecker products. For example, the second moments of $\widetilde{z}_t$ can either be stored in a $n_z \times n_z$ matrix $E(\widetilde{z}_t \cdot \widetilde{z}_t') =: \Sigma_z$ or in the $n_z^2 \times 1$ vector $E(\widetilde{z}_t \otimes \widetilde{z}_t) = vec(\Sigma_z)$; this notion naturally carries over to higher orders. There is an intimate relationship between moments and cumulants: If two probability distributions have identical moments, they will have identical cumulants as well. In particular, the second cumulant is equal to the autocovariance matrix and the third cumulant to the autocoskewness matrix. The fourth-order cumulant, however, is the fourth-order product-moment (autocokurtosis matrix) less permutations of second-order moments. In general, for cumulants higher than three, we need to know the lower-order moments or cumulants. Nevertheless, using cumulants is preferable for several reasons. For instance, all cumulants of a Gaussian process of order three and above are zero, whereas the same applies only to odd product-moments. Furthermore, the cumulant of two statistically independent random processes equals the sum of the cumulants of the individual processes (which is not true for higher-order moments). And lastly, cumulants of a white noise sequence, such as $\xi_t$, are Kronecker delta functions, so that their polyspectra are flat (Mendel, 1991).[5]

Assuming that $\mathcal{C}_{k,z}(t_1, \ldots, t_{k-1})$ is absolutely summable, the $k$th-order polyspectrum $\mathcal{S}_{k,z}$ is defined as the ($k$-1)-dimensional Fourier transform of the $k$th-order cumulant

$$\mathcal{S}_{k,z}(\omega_1, \ldots, \omega_{k-1}) := \frac{1}{(2\pi)^{k-1}} \sum_{t_1=-\infty}^{\infty} \cdots \sum_{t_{k-1}=-\infty}^{\infty} \mathcal{C}_{k,z}(t_1, \ldots, t_{k-1}) \cdot exp\{-i \sum_{j=1}^{k-1} \omega_j t_j\},$$

with $\omega_j \in [-\pi; \pi]$ and imaginary $i$ (see Swami et al. (1994) for further details). The second-, third- and fourth-order spectra are called the power spectrum, bispectrum and trispectrum, respectively. The power spectrum corresponds to the well-studied spectral density, which is a decomposition of the autocorrelation structure of the underlying process (Wiener-Khinchin theorem). The bispectrum can be viewed as a decomposition of

[5]For a mathematical discussion of using cumulants instead of moments in terms of ergodicity and proper functions, see Brillinger (1965).

the third moments (auto- and cross-skewness) over frequency and is useful for considering systems with asymmetric nonlinearities. In studying symmetric nonlinearities, the trispectrum is a more powerful tool, as it represents a decomposition of (auto- and cross-) kurtosis over frequency. Furthermore, both the bi- and trispectrum will be equal to zero for a Gaussian process, such that departures from Gaussianity will be reflected in these higher-order spectra.

Standard results from VAR(1) systems and insights from HOS can be used, regarding the computation of unconditional cumulants and polyspectra. The $k$th-order cumulants of $\xi_t$ are

$$\mathcal{C}_{k,\xi}(t_1,\ldots,t_{k-1}) = \begin{cases} \Gamma_{k,\xi} & \text{if } t_1 = \cdots = t_{k-1} = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and corresponding polyspectra $\mathcal{S}_{k,\xi}(\omega_1,\ldots,\omega_{k-1}) = (2\pi)^{1-k}\Gamma_{k,\xi}$ are flat. Letting $[\otimes_{j=1}^{k} X(j)] = X(1) \otimes X(2) \otimes \cdots \otimes X(k)$ for objects $X(j)$, Swami & Mendel (1990) show that the cumulants of the state vector $\widetilde{z}_t$

$$\mathcal{C}_{k,z}(t_1,\ldots,t_{k-1}) = [\otimes_{j=0}^{k-1} A^{t_j}] \cdot \mathcal{C}_{k,z}(0,\ldots,0)$$

are given in terms of their zero-lag cumulants

$$\mathcal{C}_{k,z}(0,\ldots,0) = (I_{n_z^k} - [\otimes_{j=1}^{k} A])^{-1} \cdot [\otimes_{j=1}^{k} B] \cdot \Gamma_{k,\xi}$$

which can be computed efficiently using iterative algorithms for generalized Sylvester equations (see Appendix A). Furthermore, there is considerable symmetry (by using appropriate permutation matrices); in particular, all second-order cumulants can be computed from $t_1 > 0$, all third-order cumulants from $t_1 \geq t_2 > 0$ and all fourth-order cumulants from $t_1 \geq t_2 \geq t_3 > 0$. Since there is a linear relationship between $\widetilde{y}_t$ and $\widetilde{z}_{t-1}$ in (2), we obtain closed-form expressions for the $k$th-order cumulants of observables. That is, for $t_j > 0$

$$\mathcal{C}_{k,y}(0,\ldots,0) = [\otimes_{j=1}^{k} C]\mathcal{C}_{k,z}(0,\ldots,0) + [\otimes_{j=1}^{k} D]\Gamma_{k,\xi}, \tag{3}$$

$$\mathcal{C}_{k,y}(t_1,\ldots,t_{k-1}) = [\otimes_{j=1}^{k} C]\mathcal{C}_{k,z}(t_1,\ldots,t_{k-1}). \tag{4}$$

Regarding the computation of polyspectra, consider the vector moving average representation (VMA) of $\widetilde{z}_t = \sum_{j=0}^{\infty} A^j B \xi_{t-j}$. Using equation (2) and lag operator $L$, we obtain

the VMA for our controls

$$\widetilde{y}_t = \sum_{j=0}^{\infty} CA^j B\xi_{t-j-1} + D\xi_t = H_\xi(L^{-1})\xi_t$$

with transfer function $H_\xi(\mathfrak{z}) = D + C\left(\mathfrak{z}I_{n_z} - A\right)^{-1} B$ for $\mathfrak{z} \in \mathbb{C}$. Setting $\mathfrak{z}_j = e^{-i\omega_j}$, with imaginary $i$ and $\omega_j \in [-\pi; \pi]$, we obtain the Fourier transformations of the cumulants of $\widetilde{y}_t$, i.e. the power spectrum $\mathcal{S}_{2,y}$, bispectrum $\mathcal{S}_{3,y}$ and trispectrum $\mathcal{S}_{4,y}$:

$$\mathcal{S}_{2,y}(\omega_1) = (2\pi)^{-1}\left[H(\mathfrak{z}_1^{-1}) \otimes H(\mathfrak{z}_1)\right]\Gamma_{2,\xi}, \tag{5}$$

$$\mathcal{S}_{3,y}(\omega_1, \omega_2) = (2\pi)^{-2}\left[H(\mathfrak{z}_1^{-1} \cdot \mathfrak{z}_2^{-1}) \otimes H(\mathfrak{z}_1) \otimes H(\mathfrak{z}_2)\right]\Gamma_{3,\xi}, \tag{6}$$

$$\mathcal{S}_{4,y}(\omega_1, \omega_2, \omega_3) = (2\pi)^{-3}\left[H(\mathfrak{z}_1^{-1} \cdot \mathfrak{z}_2^{-1} \cdot \mathfrak{z}_3^{-1}) \otimes H(\mathfrak{z}_1) \otimes H(\mathfrak{z}_2) \otimes H(\mathfrak{z}_3)\right]\Gamma_{4,\xi}. \tag{7}$$

## 3. State-space representation of linear and nonlinear DSGE models

Let $E_t$ be the expectation operator conditional on information available at time $t$, then $E_t f\left(u_{t+1}, x_{t+1}, y_{t+1}, u_t, x_t, y_t, \sigma\right) = 0$ is called the general DSGE model with states $x_t$, controls $y_t$, stochastic innovations $u_t$, and perturbation parameter $\sigma$, which can be cast into a nonlinear first-order system of expectational difference equations $f$. The solution of such rational expectation models is characterized by so-called *policy-functions*, $g$ and $h$, that solve (at least approximately) the system of equations $f$.

$$x_{t+1} = h(x_t, u_{t+1}, \sigma) \tag{8}$$

$$y_{t+1} = g(x_t, u_{t+1}, \sigma) \tag{9}$$

For the vector of innovations $u_t$ we assume $E(u_t) = 0$ and finite covariance matrix $E(u_t u_t') =: \Sigma_u$. Furthermore, $u_t$ is nth-order white noise with finite higher-order moments, where n depends on the order of approximation.[6] Apart from the existence of moments and white noise property, we do not need to impose any distributional assumptions on $u_t$.[7]

[6] Because we focus on first four cumulants of observables in the PSS, we require at least finite eighth moments for a second-order approximation and finite twelfth moments for a third-order approximation. In other words, $u_t$ is at least an eighth- or twelfth-order white noise process, which implies $y_t$ being stationary of order four.

[7] Our DYNARE toolbox can handle both the Gaussian as well as Student's t-distribution as the underlying shock process, provided the moments exist.

The approximations of the policy functions are a straightforward application of Taylor series expansions in the state variables. We use perturbation techniques to solve the model around the nonstochastic steady state given by $\bar{x} = h(\bar{x}, 0, 0)$ and $\bar{y} = g(\bar{x}, 0, 0)$. Various simulation studies show, however, that Taylor approximations of an order higher than one may generate explosive time paths, even though the first-order approximation is stable. This is due to artificial fixed points of the approximation, see Kim et al. (2008, p. 3408) for a univariate example. Thus, the model may be neither stationary nor imply an ergodic probability distribution, both of which assumptions are essential for calibration, estimation and identification. Thus, we use the pruning scheme, in which one omits terms from the policy functions that have higher-order effects than the approximation order. That is, we decompose the state vector into first-order ($\hat{x}_t^f$), second-order ($\hat{x}_t^s$) and third-order ($\hat{x}_t^{rd}$) effects, ($\hat{x}_t = \hat{x}_t^f + \hat{x}_t^s + \hat{x}_t^{rd}$), and set up the law of motions for these variables, preserving only effects up to first-, second, and third-order respectively (see the technical appendix of Andreasen et al. (2014) for details.):

$$\hat{x}_{t+1}^f = h_x \hat{x}_t^f + h_u u_{t+1} \tag{10}$$

$$\hat{x}_{t+1}^s = h_x \hat{x}_t^s + \frac{1}{2}\left[H_{xx}\left(\hat{x}_t^f \otimes \hat{x}_t^f\right) + 2H_{xu}\left(\hat{x}_t^f \otimes u_{t+1}\right) + H_{uu}\left(u_{t+1} \otimes u_{t+1}\right) + h_{\sigma\sigma}\sigma^2\right] \tag{11}$$

$$\begin{aligned}\hat{x}_{t+1}^{rd} = {} & h_x \hat{x}_t^{rd} + H_{xx}\left(\hat{x}_t^f \otimes \hat{x}_t^s\right) + H_{xu}\left(\hat{x}_t^s \otimes u_{t+1}\right) + \frac{3}{6}H_{x\sigma\sigma}\hat{x}_t^f + \frac{3}{6}H_{u\sigma\sigma}u_{t+1} \\ & + \frac{1}{6}H_{xxx}\left(\hat{x}_t^f \otimes \hat{x}_t^f \otimes \hat{x}_t^f\right) + \frac{1}{6}H_{uuu}\left(u_{t+1} \otimes u_{t+1} \otimes u_{t+1}\right) \\ & + \frac{3}{6}H_{xxu}\left(\hat{x}_t^f \otimes \hat{x}_t^f \otimes u_{t+1}\right) + \frac{3}{6}H_{xuu}\left(\hat{x}_t^f \otimes u_{t+1} \otimes u_{t+1}\right)\end{aligned} \tag{12}$$

A hat denotes deviations from steady-state, e.g. $\hat{y}_t = y_t - \bar{y}$. Note, that there are no higher-order effects in $u_{t+1}$. The law of motions for the controls can be derived analogously.

**Proposition 1** (Pruned state-space)**.** *Given an extended state vector $z_t$ and an extended vector of innovations $\xi_t$, the pruned solution of a DSGE model can be rewritten as a linear time-invariant state-space system:*

$$z_{t+1} = c + Az_t + B\xi_{t+1} \tag{13}$$

$$y_{t+1} = \bar{y} + d + Cz_t + D\xi_{t+1} \tag{14}$$

PROOF. See Andreasen et al. (2014) and the online appendix for the exact expressions depending on the order of approximation.

It can be shown that $\xi_t$ is zero mean white noise with finite moments. However, for higher-order approximations $\xi_t$ is non-Gaussian, even if the underlying process $u_t$ is normally distributed, therefore leaving scope for higher-order moments to contain additional information.[8] Moreover, if the first-order approximation is stable, i.e. all Eigenvalues of $h_x$ have modulus less than one, the pruned state-space is then also stable. In other words, all higher-order terms are unique and all Eigenvalues of A have modulus less than one. The mean of the extended state vector is equal to $E(z_t) = (I_{n_z} - A)^{-1}c$. Since there is a linear relationship between $y_t$ and $z_{t-1}$ in (14), we obtain $E(y_t) = \bar{y} + CE(z_t) + d$. After subtracting the mean, we are therefore able to use the expressions of section 2 to compute higher-order cumulants and polyspectra.

## 4. Monte-Carlo analysis

In this section we demonstrate our formulas by a Monte-Carlo analysis using three well-known DSGE models: Smets & Wouters (2007) for a first-order approximation (see Table 1), An & Schorfheide (2007) for a second-order approximation (see Table 2) and the neoclassical growth model as in Schmitt-Grohé & Uribe (2004) for a third-order approximation (see Table 3). It is well known that simulating higher-order moments one requires a large sample size as well as many simulation runs, since one deals with outliers taken to the powers of three and above.[9] Therefore, for each model, we simulate 1000 trajectories of the PSS with 10000 data points each (after discarding 1000

[8] Regarding the computation of $\Gamma_{k,\xi}$, see the online appendix. The idea is, that $\Gamma_{k,\xi}$ can be partitioned into several submatrices which can be computed symbolically element-by-element, and contain many duplicate entries. For instance, note that $E[\xi_t \otimes \xi_t \otimes \xi_t]$ is of dimension $n_\xi^3$, but the number of distinct elements is $n_\xi(n_\xi + 1)(n_\xi + 2)/6$, because $\xi_{i,t}\xi_{j,t}\xi_{k,t} = \xi_{j,t}\xi_{i,t}\xi_{k,t} = \xi_{i,t}\xi_{k,t}\xi_{j,t}$ and so forth. We can use special matrix algebra analogous to the duplication matrix, called triplication and quadruplication matrix (Meijer, 2005), to ease the computations for higher-order product-moments of $\xi_t$.

[9] Bai & Ng (2005) derive sampling distributions for the coefficients of skewness and kurtosis for serially correlated data. They also assume stationarity up to eighth order and show in a simulation exercise of an AR(1) process that test statistics for skewness have acceptable finite sample size and power, whereas for kurtosis the size distortions are tremendous. See also Bao (2013) on finite sample biases.

points) and using antithetic shocks.[10] We use the original parametrization of the models, however, we impose both the Gaussian as well as Student's t distribution as the underlying shock process. We then compute the sample variance, skewness and excess kurtosis of the stochastic innovations and observables of each trajectory and average over all Monte-Carlo runs. Lastly, we compare these to their theoretical counterparts using the formulas derived in this paper.[11] Tables 1 to 3 summarize the results. For a first-order approximation the empirical variance, skewness and excess kurtosis are very close to their theoretical values no matter which distribution is imposed on the shocks. In higher-order approximations the discrepancies in the skewness and in particular excess kurtosis are evident: matching higher-order moments in simulation studies is hard. This is already evident in the statistics of the underlying stochastic innovations which are directly drawn from a random number generator. Increasing the number of Monte Carlo runs as well as sample size would on the one hand increase the precision but on the other hand also the computational time. For an applied researcher who uses a try-and-error approach to match third-order or fourth-order characteristics of an observable variable in a calibration exercise this is unfeasible. Hence, we conclude that our expressions are a convenient and fast way to compute higher-order statistics for linear and nonlinear (pruned) DSGE models.

## 5. Conclusion

Whenever we are confronted with nonlinearities or non-Gaussian stochastic innovations, it is natural to focus on higher-order moments for the calibration and estimation of parameters. The contribution of this note is twofold. First, we derive expressions for unconditional moments, cumulants and polyspectra for linear and nonlinear (pruned) DSGE models. Second, we provide a DYNARE toolbox which computes higher-order

[10] In the code one can change all settings regarding the Monte-Carlo framework in a graphical user-interface, i.e. number of trajectories, sample size, burn-in phase and use of antithetic shocks or not.

[11] Note that the second-order zero-lag cumulant of $y_t$ is equal to the covariance matrix. Skewness can either be computed via standardized product moments or via the ratio of the third zero-lag cumulant and the 1.5th power of the second zero-lag cumulant. Furthermore, excess kurtosis is the fourth zero-lag cumulant normalized by the square of the second-order cumulant.

statistics for linear and nonlinear (pruned) DSGE models. Since higher-order cumulants and polyspectra measure the departure from Gaussianity, these expressions and the code can be used to gain additional information from higher-order statistics. For instance, Mutschler (2015) shows that this approach imposes additional restrictions, which can be used to identify parameters that are unidentified in a Gaussian first-order approximation.

## Appendix A. Using generalized Sylvester equations for cumulants

The zero-lag cumulants (k=2,3,4)

$$\mathcal{C}_{k,z} = (I_{n_z^k} - [\otimes_{j=1}^{k} A])^{-1} \cdot [\otimes_{j=1}^{k} B] \cdot \Gamma_{k,\xi}$$

require the inversion of the big matrix $\left(I_{n_z^k} - [\otimes_{j=1}^{k} A]\right)$. Since $\mathcal{C}_{k,z}$ and $\Gamma_{k,\xi}$ are vectors, we can use properties of the Kronecker-product and rewrite the equations to

$$\begin{aligned}
[\underset{n_z \times n_z}{\ \mathcal{C}_{2,z}\ }] &= A[\underset{n_z \times n_z}{\ \mathcal{C}_{2,z}\ }]A' + B[\underset{n_\xi \times n_\xi}{\ \Gamma_{2,\xi}\ }]B', \\
[\underset{n_z^2 \times n_z}{\ \mathcal{C}_{3,z}\ }] &= (A \otimes A)[\underset{n_z^2 \times n_z}{\ \mathcal{C}_{3,z}\ }]A' + (B \otimes B)[\underset{n_\xi^2 \times n_\xi}{\ \Gamma_{3,\xi}\ }]B', \\
[\underset{n_z^2 \times n_z^2}{\ \mathcal{C}_{4,z}\ }] &= (A \otimes A)[\underset{n_z^2 \times n_z^2}{\ \mathcal{C}_{4,z}\ }](A \otimes A)' + (B \otimes B)[\underset{n_\xi^2 \times n_\xi^2}{\ \Gamma_{4,\xi}\ }](B \otimes B)',
\end{aligned}$$

where $[\underset{n \times m}{\ }]$ reshapes a $n \cdot m$ vector into a $n \times m$ matrix. In other words, we reduce the inversion problem to a generalized Sylvester equation, which can be efficiently solved.

## References


An, S., & Schorfheide, F. (2007). Bayesian analysis of DSGE models. *Econometric Reviews*, *26*, 113–172. doi:10.1080/07474930701220071.

Andreasen, M. M., Fernández-Villaverde, J., & Rubio-Ramírez, J. F. (2014). *The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications*. Working Paper Aarhus University.

Bai, J., & Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *Journal of Business & Economic Statistics*, *23*, 49–60. doi:10.1198/073500104000000271.

Bao, Y. (2013). On sample skewness and kurtosis. *Econometric Reviews*, *32*, 415–448. doi:10.1080/07474938.2012.690665.

Brillinger, D. R. (1965). An introduction to polyspectra. *The Annals of Mathematical Statistics*, *36*, 1351–1374. doi:10.1214/aoms/1177699896.

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory*. Classics in Applied Mathematics (Book 36). San Francisco: Society for Industrial and Applied Mathematics.

Chib, S., & Ramamurthy, S. (2014). DSGE Models with Student-t Errors. *Econometric Reviews*, *33*, 152–171. doi:10.1080/07474938.2013.807152.

Curdia, V., Del Negro, M., & Greenwald, D. L. (2014). Rare shocks, great recessions. *Journal of Applied Econometrics*, *29*, 1031–1052. doi:10.1002/jae.2395.

Kim, J., Kim, S., Schaumburg, E., & Sims, C. A. (2008). Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. *Journal of Economic Dynamics and Control*, *32*, 3397–3414. doi:10.1016/j.jedc.2008.02.003.

Lan, H., & Meyer-Gohde, A. (2013). *Pruning in Perturbation DSGE Models - Guidance from Nonlinear Moving Average Approximations*. Discussion Papers SFB 649 Humbold University Berlin.

Lombardo, G., & Uhlig, H. (2014). *A theory of pruning*. ECB Working Paper Series Number 1696 European Central Bank.

Meijer, E. (2005). Matrix algebra for higher order moments. *Linear Algebra and its Applications*, *410*, 112 – 134. doi:10.1016/j.laa.2005.02.040.

Mendel, J. M. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, *79*, 278–305. doi:10.1109/5.75086.

Mutschler, W. (2015). Identification of DSGE models the effect of higher-order approximation and pruning. *Journal of Economic Dynamics and Control*, *56*, 34 – 54. doi:http://dx.doi.org/10.1016/j.jedc.2015.04.007.

Nikias, C., & Petropulu, A. (1993). *Higher-order Spectra Analysis: A Nonlinear Signal Processing Framework*. Prentice Hall signal processing series. PTR Prentice Hall.

Priestley, M. (1983). *Spectral Analysis and Time Series, Two-Volume Set: Volumes I and II*. Elsevier Science.

Schmitt-Grohé, S., & Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control*, *28*, 755–775. doi:10.1016/S0165-1889(03)00043-5.

Smets, F., & Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, *97*, 586–606. doi:10.1257/aer.97.3.586.

Swami, A., Giannakis, G., & Shamsunder, S. (1994). Multichannel ARMA processes. *IEEE Transactions on Signal Processing*, *42*, 898–913. doi:10.1109/78.285653.

Swami, A., & Mendel, J. (1990). Time and lag recursive computation of cumulants from a state-space model. *IEEE Transactions on Automatic Control*, *35*, 4–17. doi:10.1109/9.45138.

Table A.1: Smets and Wouters (2007): First-Order State Space System

| | VARIANCE | | | | SKEWNESS | | | | EXCESS KURTOSIS | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Gaussian | | Student-t | | Gaussian | | Student-t | | Gaussian | | Student-t | |
| shocks | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) |
| ea | 0.21 | 0.21 | 0.27 | 0.27 | -0.00 | 0 | 0.00 | 0 | -0.00 | 0 | 0.99 | 1 |
| eb | 3.43 | 3.43 | 4.28 | 4.28 | -0.00 | 0 | -0.00 | 0 | -0.00 | 0 | 1.02 | 1 |
| eg | 0.37 | 0.37 | 0.46 | 0.46 | -0.00 | 0 | -0.00 | 0 | -0.00 | 0 | 0.98 | 1 |
| eqs | 0.36 | 0.36 | 0.45 | 0.45 | -0.00 | 0 | 0.00 | 0 | -0.00 | 0 | 0.99 | 1 |
| em | 0.06 | 0.06 | 0.07 | 0.07 | -0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.98 | 1 |
| epinf | 0.02 | 0.02 | 0.03 | 0.03 | -0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.99 | 1 |
| ew | 0.04 | 0.04 | 0.05 | 0.05 | -0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 1.00 | 1 |
| observables | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) |
| labobs | 159.47 | 159.36 | 199.67 | 199.20 | 0.00 | 0 | 0.00 | 0 | -0.01 | 0 | 0.11 | 0.13 |
| robs | 17.42 | 17.41 | 21.77 | 21.76 | 0.00 | 0 | -0.00 | 0 | -0.01 | 0 | 0.14 | 0.15 |
| pinfobs | 3.03 | 3.03 | 3.79 | 3.79 | 0.00 | 0 | -0.00 | 0 | 0.00 | 0 | 0.09 | 0.11 |
| dy | 47.90 | 47.88 | 59.85 | 59.85 | 0.00 | 0 | -0.00 | 0 | 0.00 | 0 | 0.68 | 0.67 |
| dc | 55.95 | 55.93 | 69.91 | 69.91 | 0.00 | 0 | -0.00 | 0 | 0.00 | 0 | 0.69 | 0.68 |
| dinve | 50.95 | 50.93 | 63.66 | 63.66 | 0.00 | 0 | -0.00 | 0 | 0.00 | 0 | 0.61 | 0.60 |
| dw | 0.59 | 0.59 | 0.73 | 0.73 | 0.00 | 0 | -0.00 | 0 | 0.00 | 0 | 0.51 | 0.52 |

Simulated (S) and theoretical (T) statistics for stochastic innovations and observables.

Table A.2: An and Schorfheide (2007): Second-Order Pruned State Space System

| | VARIANCE | | | | SKEWNESS | | | | EXCESS KURTOSIS | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Gaussian | | Student-t | | Gaussian | | Student-t | | Gaussian | | Student-t | |
| shocks | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) |
| eR | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 0.00 | 0 | -0.00 | 0 | 0.54 | 0.55 |
| eg | 0.00 | 0.00 | 0.00 | 0.00 | -0.00 | 0 | 0.00 | 0 | -0.00 | 0 | 0.55 | 0.55 |
| ez | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.54 | 0.55 |
| observables | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) |
| YGR | 1.24 | 1.24 | 1.44 | 1.44 | 0.29 | 0.15 | 0.36 | 0.22 | 0.18 | 0.14 | 0.66 | 0.59 |
| INFL | 8.00 | 8.01 | 9.24 | 9.24 | 0.10 | 0.03 | 0.12 | 0.04 | 0.01 | 0.01 | 0.17 | 0.16 |
| INT | 10.88 | 10.89 | 12.56 | 12.57 | 0.08 | 0.01 | 0.09 | 0.01 | -0.00 | 0.00 | 0.04 | 0.04 |

Simulated (S) and theoretical (T) statistics for observable variables.

Table A.3: Neoclassical Growth Model: Third-Order Pruned State Space System

| | VARIANCE | | | | SKEWNESS | | | | EXCESS KURTOSIS | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Gaussian | | Student-t | | Gaussian | | Student-t | | Gaussian | | Student-t | |
| | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) | (S) | (T) |
| ea | 1.00 | 1 | 1.15 | 1.15 | -0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.54 | 0.55 |
| c | 0.70 | 0.71 | 0.81 | 0.82 | -0.12 | -0.17 | -0.17 | -0.23 | 0.02 | 0.06 | 0.43 | 0.55 |

Simulated (S) and theoretical (T) statistics for stochastic innovation ea and observable variable c.