Conditional Gauss–Hermite Filtering with Application to Volatility Estimation

Hermann Singer
FernUniversität in Hagen *

October 21, 2008

Abstract

The conditional Gauss–Hermite filter (CGHF) utilizes a decomposition of the filter density $p(y_1, y_2)$ into the product of the conditional density $p(y_1|y_2)$ with $p(y_2)$ where the state vector y is partitioned into (y_1, y_2). In contrast to the usual Gauss–Hermite filter (GHF) it is assumed that the product terms can be approximated by Gaussians. Due to the nonlinear dependence of $\phi(y_1|y_2)$ from y_2, quite complicated densities can be modeled, but the advantages of the normal distribution are preserved. For example, in stochastic volatility models, the joint density $p(y, \sigma)$ strongly deviates from a bivariate Gaussian, whereas $p(y|\sigma)p(\sigma)$ can be well approximated by $\phi(y|\sigma)\phi(\sigma)$. As in the GHF, integrals in the updates can be computed by Gauss–Hermite quadrature. We obtain recursive update fomulas for the conditional moments $E(y_1|y_2)$, $\text{Var}(y_1|y_2)$ and $E(y_2)$, $\text{Var}(y_2)$.

Key Words: Multivariate stochastic differential equations; Nonlinear systems; Discrete time measurements; Continuous-discrete state space model; Conditionally gaussian densities; Stochastic volatility.

1 Introduction

The Gaussian filter (GF) assumes, that the true filter density $p(y)$ can be approximated by a Gaussian distribution $\phi(y)$. Thus, expectation values occuring in the time and measurement update can be computed numerically by Gauss–Hermite integration (GHF, cf. Ito and Xiong; 2000). There are important applications, however, where the joint Gaussian assumption does not lead to satisfactory results. For example, if the volatility parameter of an

*Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung, D-58084 Hagen, Germany, hermann.singer@fernuni-hagen.de

Ornstein-Uhlenbeck process is filtered (Bayesian estimation), the measurements do not lead to any change in the conditional volatility state. This stems from the fact, that the state vector $(y(t), \sigma(t))$ is not bivariate Gaussian anymore, since the process $y(t)$ is driven by the product of the Gaussian volatility and the Wiener process (cf. fig. 5). Similarly, stochastic volatility models (Scott; 1987; Hull and White; 1987) are not satisfactorily filtered by the GHF (and other filters relying on 2 moments, such as the extended Kalman filter EKF, or the unscented Kalman filter UKF). One can solve the problem by using better density approximations such as the Gaussian sum filter (Alspach and Sorenson; 1972; Ito and Xiong; 2000) or by expanding the density into a Fourier series such as the Hermite expansion (Kuznetsov et al.; 1960; Aït-Sahalia; 2002; Singer; 2006b, 2008).

Alternatively, it is proposed to factorize the joint density of all states y by using the partitioned state $y = (y_1, y_2)$ such that $p(y_1|y_2)$ is (approximately) conditionally Gaussian. Then, $p(y_1|y_2)p(y_2) \approx \phi(y_1|y_2)\phi(y_2)$ and the numerical methods for the usual GHF can be adapted. [1]

In section 2, static conditionally Gaussian models are discussed. This is extended to dynamic models in sect. 3 and illustrated by the Ornstein-Uhlenbeck process (Bayesian estimation of the volatility parameter). Section 4 develops the general conditional Gaussian filter whereas in section 5 recursive ML estimation is compared with sequential filtering of σ using several approximate nonlinear filters.

2 Conditionally Gaussian models

As explained, it may be advantageous not to approximate the full joint density, but first to factorize the distribution by conditioning on certain variables in order to get conditional distributions near to a Gaussian.

2.1 Example 1

For example, if $y \sim N(\mu, \sigma)$, the joint density

$$p(y, \mu, \sigma) \quad = \quad \phi(y|\mu, \sigma)p(\mu, \sigma) \propto (2\pi\sigma^2)^{-1/2} \exp[-\tfrac{1}{2}(y-\mu)^2)/\sigma^2] \qquad (1)$$

is of Gaussian shape as a function of μ, but not for σ (cf. figs. 1– 2 where we set $\mu = 0$ and used a prior $p(\sigma) = \phi(\sigma; 2, 1)$). The joint distribution $p(y, \sigma) = p(y|\sigma)p(\sigma)$ displays the variability in the variance of y. It cannot be well approximated by a bivariate Gaussian $\phi(y, \sigma)$, as would be the case for the Gaussian filter. From fig. 2 (right), it can be seen that the posterior mean $E[\sigma|y]$ depends on y and thus we obtain estimates of σ from observations y, although the covariance $\mathrm{Cov}(y, \sigma) = E[y\sigma] - E[y]E[\sigma] = E[E[y|\sigma]\sigma] - E[y]E[\sigma] = 0$, since $E[y|\sigma] = E[y] = \mu = 0$.

[1] We use the notation $\phi(y_1|y_2) = \phi(y_1; E[y_1|y_2], \mathrm{Var}[y_1|y_2])$ for the Gaussian density

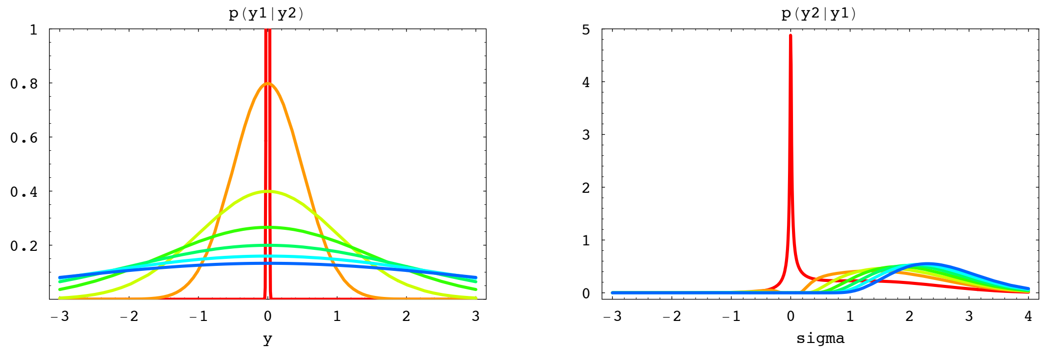

Figure 1: Conditional density $\phi(y|\sigma)$ and posterior density $p(\sigma|y)$ with prior $p(\sigma) = \phi(\sigma; 2, 1)$.

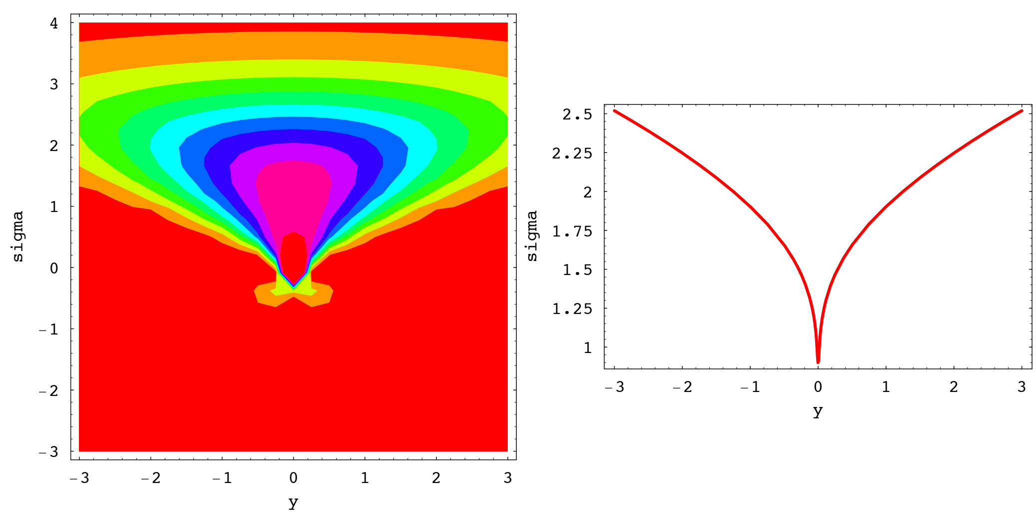

Figure 2: Joint density $\phi(y|\sigma)p(\sigma)$ and posterior mean $E[\sigma|y]$. It depends on y although $\mathrm{Cov}(y, \sigma) = 0$ (see text).

In contrast, figs. 3–4 (setting $\sigma = 1$ and using a prior $p(\mu) = \phi(\mu; 2, 1)$) display the joint distribution of y and μ. Since the role of these variables is symmetric in eqn. 1, the Gaussian shape is preserved and the regression $E[\mu|y]$ is linear. Thus, the estimation of parameters related to the mean (e.g. drift coefficients) is much simpler as compared to volatility parameters, where the regression $E[\sigma|y]$ is nonlinear (cf. fig. 2). In a bivariate Gaussian setting $\phi(y, \sigma)$, only a linear relation is possible (cf. 3) ■

Therefore, the idea is put forward, to represent the joint distribution of states and volatilites $p(y, \sigma)$ not by a joint Gaussian $\phi(y, \sigma)$, as in the Gaussian filter (or EKF, UKF), but by the product $\phi(y|\sigma)\phi(\sigma)$. This allows a fully nonlinear specification of the conditional moments

$$\begin{aligned} E[y|\sigma] &= \mu_1(\sigma) \\ \mathrm{Var}[y|\sigma] &= \Sigma_1(\sigma). \end{aligned} \tag{2}$$

In contrast, the joint Gaussian assumption only allows the normal correlation structure (Liptser and Shiryayev; 2001, ch. 13, theorem 13.1, lemma 14.1)

$$\begin{aligned} E[y|\sigma] &= E[y] + \mathrm{Cov}(y, \sigma)\mathrm{Var}(\sigma)^-(\sigma - E[\sigma]) \\ \mathrm{Var}[y|\sigma] &= \mathrm{Var}(y) - \mathrm{Cov}(y, \sigma)\mathrm{Var}(\sigma)^-\mathrm{Cov}(\sigma, y) \end{aligned} \tag{3}$$

which is linear in the conditional mean and independent of σ for the conditional variance ($^-$ denotes the generalized inverse). Put the other way round, a bivariate Gaussian can be obtained by a linear $\mu_1(\sigma)$ and constant $\Sigma_1(\sigma) = \Sigma_1$.

More generally, the distribution $p(y_1, y_2)$ of the vectors y_1, y_2 is not approximated by $\phi(y_1, y_2)$, but by

$$\begin{aligned} p(y_1, y_2) &= p(y_1|y_2)p(y_2) \approx \phi(y_1|y_2)\phi(y_2) \\ &= \phi(y_1; \mu_1(y_2), \Sigma_1(y_2))\, \phi(y_2; \mu_2, \Sigma_2), \end{aligned} \tag{4}$$

where the conditional moments $\mu_1(y_2) = E[y_1|y_2], \Sigma_1(y_2) = \mathrm{Var}(y_1|y_2)$ are nonlinear functions of the conditioning states y_2. Of course, the choice of y_2 depends on the form of the true distribution $p(y_1, y_2)$. It is chosen such, that $p(y_1|y_2)$ is well approximated by a Gaussian with parameters $\mu_1(y_2), \Sigma_1(y_2)$. In example 1, we must condition on $y_2 = (\mu, \sigma)$ to get exactly $p(y_1|y_2) = \phi(y_1|y_2)$. A jointly Gaussian $\phi(y_1, y_2)$ is included as a special case (linear moments)

$$\begin{aligned} E[y_1|y_2] &= E[y_1] + \mathrm{Cov}(y_1, y_2)\mathrm{Var}(y_2)^-(y_2 - E[y_2]) \\ \mathrm{Var}[y_1|y_2] &= \mathrm{Var}(y_1) - \mathrm{Cov}(y_1, y_2)\mathrm{Var}(y_2)^-\mathrm{Cov}(y_2, y_1). \end{aligned} \tag{5}$$

with conditional moments.

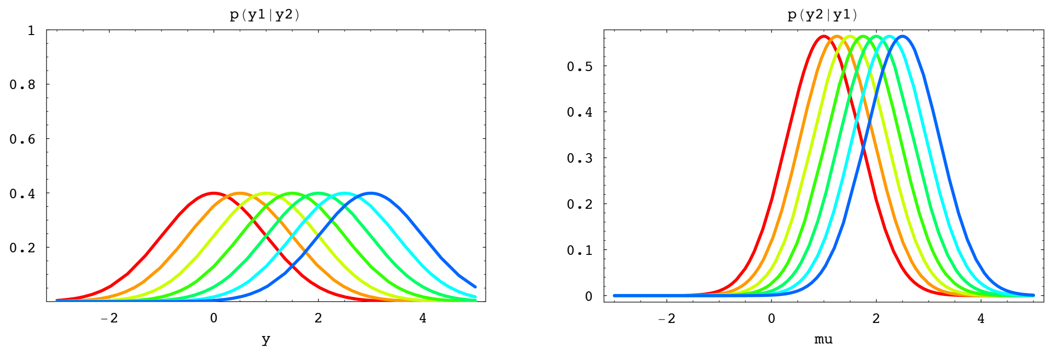

Figure 3: Conditional density $\phi(y|\mu)$ and posterior density $p(\mu|y)$ with prior $p(\mu) = \phi(\mu; 2, 1)$.

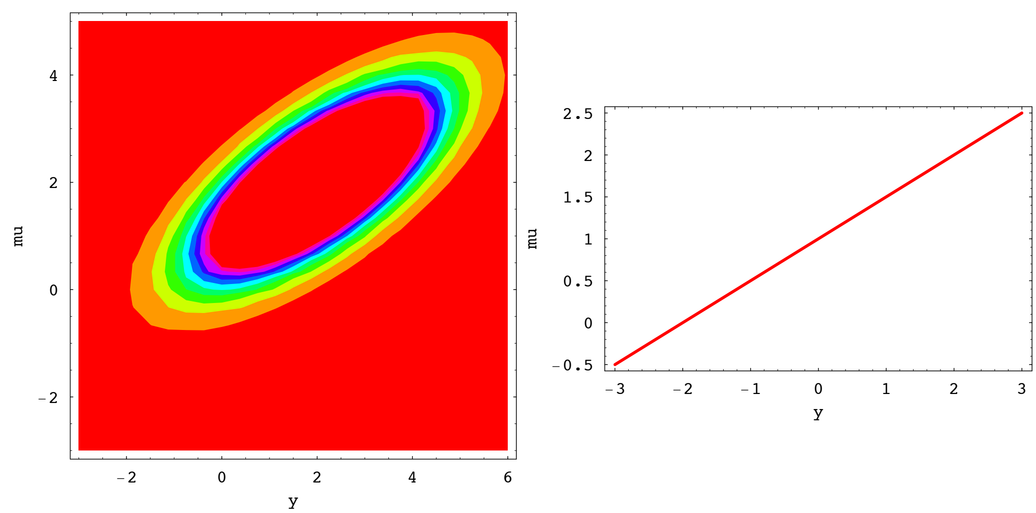

Figure 4: Joint density $\phi(y|\mu)p(\mu)$ and posterior mean $E[\mu|y]$.

3 State space models

We want to filter the *continuous-discrete state space model* (Jazwinski; 1970)

$$dy(t) = f(y(t), t, \psi)dt + g(y(t), t, \psi)dW(t) \tag{6}$$

where discrete time measurements $z_i := z(t_i)$ are taken at times $\{t_0, t_1, \ldots, t_T\}$ and $t_0 \leq t \leq t_T$ according to the measurement equation

$$z_i = h(y(t_i), t_i, \psi) + \epsilon_i. \tag{7}$$

In state equation (6), $W(t)$ denotes an r-dimensional Wiener process and the state is described by the p-dimensional state vector $y(t)$. It fulfils a system of stochastic differential equations in the sense of Itô (Arnold; 1974) with random initial condition $y(t_0) \sim p_0(y, \psi)$. The functions $f : \mathbb{R}^p \times \mathbb{R} \times \mathbb{R}^u \to \mathbb{R}^p$ and $g : \mathbb{R}^p \times \mathbb{R} \times \mathbb{R}^u \to \mathbb{R}^p \times \mathbb{R}^r$ are called drift and diffusion coefficients, respectively. In measurement equation (7), $\epsilon_i \sim N(0, R(t_i, \psi))i.d.$ is a k-dimensional discrete time white noise process (measurement error). Parametric estimation is based on the u-dimensional parameter vector ψ. For notational simplicity, deterministic control variables $x(t)$ are absorbed in the time argument t. Moreover, the functions f and g may also depend on nonanticipative measurements $Z^i = \{z(t_j)|j \leq i\}, t_i \leq t$ and h, R may depend on lagged measurements $Z^{i-1} = \{z(t_j)|j \leq i-1\}$ allowing continuous time ARCH specifications. In the linear case, the system is conditionally Gaussian (cf. Liptser and Shiryayev; 2001, ch. 11). This dependence will be dropped below.

3.1 Example 2: Ornstein-Uhlenbeck process

The linear Gauss-Markov process is given by the SDE

$$dy(t) = \lambda y(t)dt + \sigma dW(t) \tag{8}$$

with measurement equation $(i = 0, \ldots, T)$

$$z_i = y_i + \epsilon_i \tag{9}$$

where $\psi = \{\lambda, \sigma, R = \mathrm{Var}(\epsilon_i)\}$ are unknown (nonrandom) parameters. For simplicity, let λ, R be known. Then, σ can be estimated by exact ML (Singer; 1993, 1995), or as Bayes estimator, using an extended state vector $\eta = \{y, \sigma\}$

$$\begin{aligned} dy &= \lambda y dt + \sigma dW(t) && (10) \\ d\sigma &= 0 && (11) \\ z_i &= y_i + \epsilon_i. && (12) \end{aligned}$$

However, as shown in fig. 5, the moments $\mu = E(y, \sigma)$ and $\Sigma = \mathrm{Var}(y, \sigma)$ cannot filter the volatility state $\sigma(t)$. However, if we note that $dy|\sigma = \lambda y dt +$

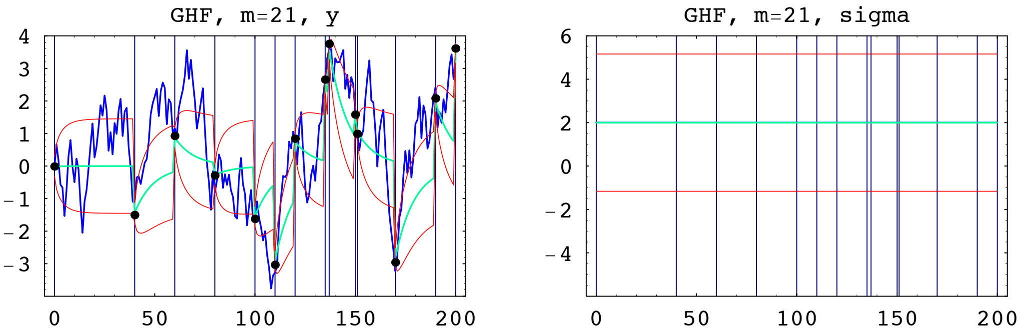

Figure 5: Gauss–Hermite filter for the Ornstein-Uhlenbeck process $y(t)$ with random volatility parameter σ.

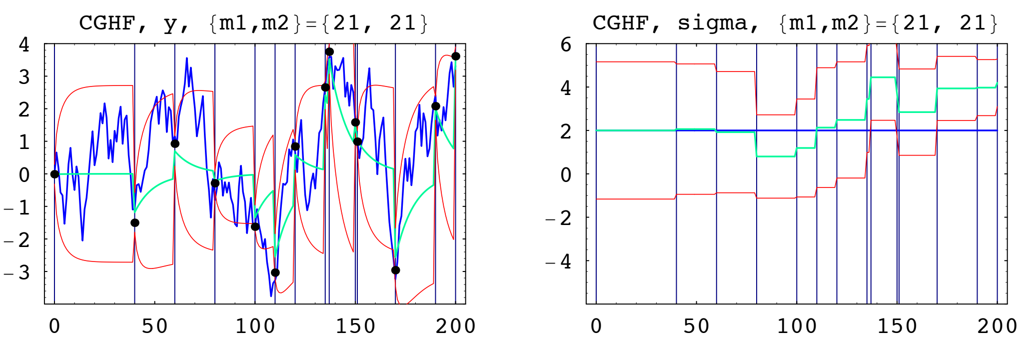

Figure 6: Conditional Gauss–Hermite filter for the Ornstein-Uhlenbeck process $y(t)$ with random volatility parameter σ.

σdW is Gaussian, the idea of the last section turns over to the dynamic context. Using the exact discrete model (EDM) at the measurement times t_i, setting $y_i = y(t_i)$ etc., we obtain

$$\begin{aligned} y_{i+1} &= \lambda_i y_i + \sigma_i u_i && (13) \\ \sigma_{i+1} &= \sigma_i && (14) \\ z_i &= y_i + \epsilon_i && (15) \end{aligned}$$

with the Gaussian error term $u_i = \int_{t_i}^{t_{i+1}} \exp[\lambda(t_{i+1} - s)]dW(s)$ and the AR(1) parameter $\lambda_i = \exp[\lambda(t_{i+1} - t_i)]$.

3.1.1 Time update:

Assume that the posteriori density after measurement z_i is conditionally Gaussian, i.e. $p(y_i|\sigma_i, Z^i) = \phi(y_i; E[y_i|\sigma_i, Z^i], \mathrm{Var}[y_i|\sigma_i, Z^i])$ and $p(\sigma_i|Z^i) = \phi(\sigma_i|Z^i)$. Then, the time update $p(y_{i+1}|\sigma_{i+1}, Z^i)$ is Gaussian with parameters

$$\begin{aligned} E[y_{i+1}|\sigma_{i+1}, Z^i] &= \lambda_i E[y_i|\sigma_i, Z^i] && (16) \\ \mathrm{Var}[y_{i+1}|\sigma_{i+1}, Z^i] &= \lambda_i \mathrm{Var}[y_i|\sigma_i, Z^i]\lambda_i' + \sigma_i \mathrm{Var}(u_i)\sigma_i' && (17) \end{aligned}$$

since $\sigma_{i+1} = \sigma_i$ in this simple example.

3.1.2 Measurement update:

At the time of measurement t_{i+1} the Bayes formula

$$p(y_{i+1}, \sigma_{i+1}|z_{i+1}, Z^i) = \frac{p(z_{i+1}|y_{i+1}, \sigma_{i+1}, Z^i)p(y_{i+1}, \sigma_{i+1}|Z^i)}{p(z_{i+1}|Z^i)} \quad (18)$$

can be evaluated easily due to the Gaussian densities (measurement and a priori density)

$$\begin{aligned} p(z_{i+1}|y_{i+1}, \sigma_{i+1}, Z^i) &= \phi(z_{i+1}; y_{i+1}, R) && (19) \\ p(y_{i+1}, \sigma_{i+1}|Z^i) &= \phi(y_{i+1}|\sigma_{i+1}, Z^i)\phi(\sigma_{i+1}|Z^i). && (20) \end{aligned}$$

Since the measurements are linear, the normal correlation update (3) is exact and one obtains

$$\begin{aligned} p(y_{i+1}, \sigma_{i+1}|Z^{i+1}) &= \phi(y_{i+1}|\sigma_{i+1}, Z^{i+1})p(\sigma_{i+1}|Z^{i+1}) && (21) \\ p(\sigma_{i+1}|Z^{i+1}) &= \phi(z_{i+1}|\sigma_{i+1}, Z^i)\phi(\sigma_{i+1}|Z^i)/p(z_{i+1}|Z^i). && (22) \end{aligned}$$

Thus, the posterior of σ_{i+1} is nongaussian due to the nonlinear dependence of $\mathrm{Var}(z_{i+1}|\sigma_{i+1}, Z^i) = \mathrm{Var}(y_{i+1}|\sigma_{i+1}, Z^i) + R$ from σ_{i+1} (cf. 17). This nonlinear dependence is the reason why the posterior mean

$$\begin{aligned} E[\sigma_{i+1}|Z^{i+1}] &= \int \sigma_{i+1} p(\sigma_{i+1}|Z^{i+1}) d\sigma_{i+1} && (23) \\ &= \int \sigma_{i+1} \phi(z_{i+1}|\sigma_{i+1}, Z^i)\phi(\sigma_{i+1}|Z^i) d\sigma_{i+1}/p(z_{i+1}|Z^i) \end{aligned}$$

is a function of the measurements, in contrast to the usual GHF. The integral can be computed by Gauss–Hermite integration (see appendix A). From the posteriori moments $E[\sigma_{i+1}|Z^{i+1}]$ and $\mathrm{Var}(\sigma_{i+1}|Z^{i+1})$ one can construct a Gaussian distribution and proceed in the recursive filter algorithm with the next time update.
For the posterior mean of the state y_{i+1} we simply obtain the usual normal correlation update

$$\begin{aligned} E[y_{i+1}|\sigma_{i+1}, Z^{i+1}] &= E[y_{i+1}|\sigma_{i+1}, Z^i] \\ &+ \mathrm{Var}(y_{i+1}|\sigma_{i+1}, Z^i)[\mathrm{Var}(y_{i+1}|\sigma_{i+1}, Z^i) + R]^{-1} \\ &\times (z_{i+1} - E[y_{i+1}|\sigma_{i+1}, Z^i]) \end{aligned} \tag{24}$$

etc. The a priori terms are given in (16). Figs. 5–6 display the difference in the performance of the GHF and the CGHF. In this picture, an Ornstein-Uhlenbeck process was simulated according to (8) with parameters $\psi = \{\lambda = -1, \sigma = 2, R = \mathrm{Var}(\epsilon_i) = 0.1\}$ with sampling interval $\delta t = 0.1$. I used a simple Euler-Maruyama scheme (cf., e.g. Kloeden and Platen; 1992). The measurements were taken at times $\tau = \{0, 4, 6, 8, 10, 11, 12, 13.5, 13.7, 15, 15.1, 17, 19, 20\}$. Clearly, the Gauss-Hermite filter (fig. 5) does not filter the volatility process (Bayesian parameter) $d\sigma = 0$, whereas the CGHF, due to the conditional Gaussian filter density, yields estimates of σ from the observations $y(t_i)$ ■

4 Conditional Gauss–Hermite filtering

In this section we derive a sequence of time update and measurement update steps for the filter density $p(y_1, y_2, t|Z^i)$ which is approximated by the product of Gaussians

$$p(y_1, y_2, t|Z^i) \approx \phi(y_1, t|y_2, t, Z^i)\phi(y_2, t|Z^i). \tag{25}$$

The densities are evaluated at the time points $\tau_j = t_0 + j\delta t, j = 0, \ldots, J = (t_T - t_0)/\delta t$, and δt is an arbitrary (but small) discretization interval. The times of measurement are given by $t_i = \tau_{j_i}$. The filter proceeds in a recursive sequence of time update (dynamic moment equations) and measurement updates (Bayes formula; cf. appendix B).
According to the Gaussian assumption (25) one has to consider the conditional moments

$$\begin{aligned} E[y_1(t)|y_2(t), Z^i] &= \mu_1(y_2(t), Z^i) &(26)\\ E[y_2(t)|Z^i] &= \mu_2(t, Z^i) &(27)\\ \mathrm{Var}(y_1(t)|y_2(t), Z^i) &= \Sigma_1(y_2(t), Z^i) &(28)\\ \mathrm{Var}(y_2(t)|Z^i) &= \Sigma_2(t, Z^i) &(29)\end{aligned}$$

and we seek recursive equations for their time evolution.

The state space model 6–7 is written in partitioned form ($y_1 : p_1 \times 1, g_1 : p_1 \times r$ etc.; dropping ψ)

$$\begin{aligned} dy_1(t) &= f_1(y_1, y_2, t)dt + g_1(y_1, y_2, t)dW(t) &(30)\\ dy_2(t) &= f_2(y_1, y_2, t)dt + g_2(y_1, y_2, t)dW(t) &(31)\end{aligned}$$

with measurements at t_i

$$z_i = h(y_1(t_i), y_2(t_i), t_i) + \epsilon_i. \tag{32}$$

4.1 Time update

In a short time step δt, the Euler-Maruyama approximation for the Itô equations (30–31) is

$$\begin{aligned} y_1(t + \delta t) &= y_1(t) + f_1(y_1, y_2, t)\delta t + g_1(y_1, y_2, t)\delta W(t) &(33)\\ y_2(t + \delta t) &= y_2(t) + f_2(y_1, y_2, t)\delta t + g_2(y_1, y_2, t)\delta W(t) &(34)\end{aligned}$$

and we find the moment equations (dropping the dependence on Z^i)

$$\begin{aligned} E[y_1(t + \delta t)|y_2(t)] &= E[y_1(t)|y_2(t)] + E[f_1(y_1, y_2, t)|y_2(t)]\delta t &(35)\\ E[y_2(t + \delta t)] &= E[y_2(t)] + E[f_2(y_1, y_2, t)]\delta t &(36)\end{aligned}$$

The second moments read

$$\begin{aligned} \mathrm{Var}[y_1(t + \delta t)|y_2(t)] &= \mathrm{Var}[y_1(t)|y_2(t)] + \\ &\quad \mathrm{Cov}[y_1(t), f_1(y_1, y_2, t)|y_2(t)]\delta t + \\ &\quad \mathrm{Cov}[f_1(y_1, y_2, t), y_1(t)|y_2(t)]\delta t + \\ &\quad E[g_1 g_1'(y_1, y_2, t)|y_2(t)]\delta t &(37)\\ \mathrm{Var}[y_2(t + \delta t)] &= \mathrm{Var}[y_2(t)] + \\ &\quad \mathrm{Cov}[y_2(t), f_2(y_1, y_2, t)]\delta t + \\ &\quad \mathrm{Cov}[f_2(y_1, y_2, t), y_2(t)]\delta t + \\ &\quad E[g_2 g_2'(y_1, y_2, t)]\delta t. &(38)\end{aligned}$$

The expectation values on the right hand sides are with respect to the distributions $\phi(y_1(t)|y_2(t), Z^i)$ and $\phi(y_2(t)|Z^i)$ and can be evaluated using Gauss–Hermite quadrature (appendix A). For example

$$\begin{aligned} E[f_1(y_1, y_2, t)|y_2(t)] &= \int f_1(y_1, y_2, t)\phi(y_1; \mu_1(y_2), \Sigma_1(y_2))dy_1 &(39)\\ &\approx \sum_{l=1}^{L} f_1(\eta_{1lm}, \eta_{2m}, t)w_{1l} &(40)\end{aligned}$$

where

$$\begin{aligned} \eta_{2m} &= \mu_2 + \Sigma_2^{1/2}\zeta_{2m} : p_2 \times 1 &(41)\\ \eta_{1lm} &= \mu_1(\eta_{2m}) + \Sigma_1^{1/2}(\eta_{2m})\zeta_{1l} : p_1 \times 1 &(42)\end{aligned}$$

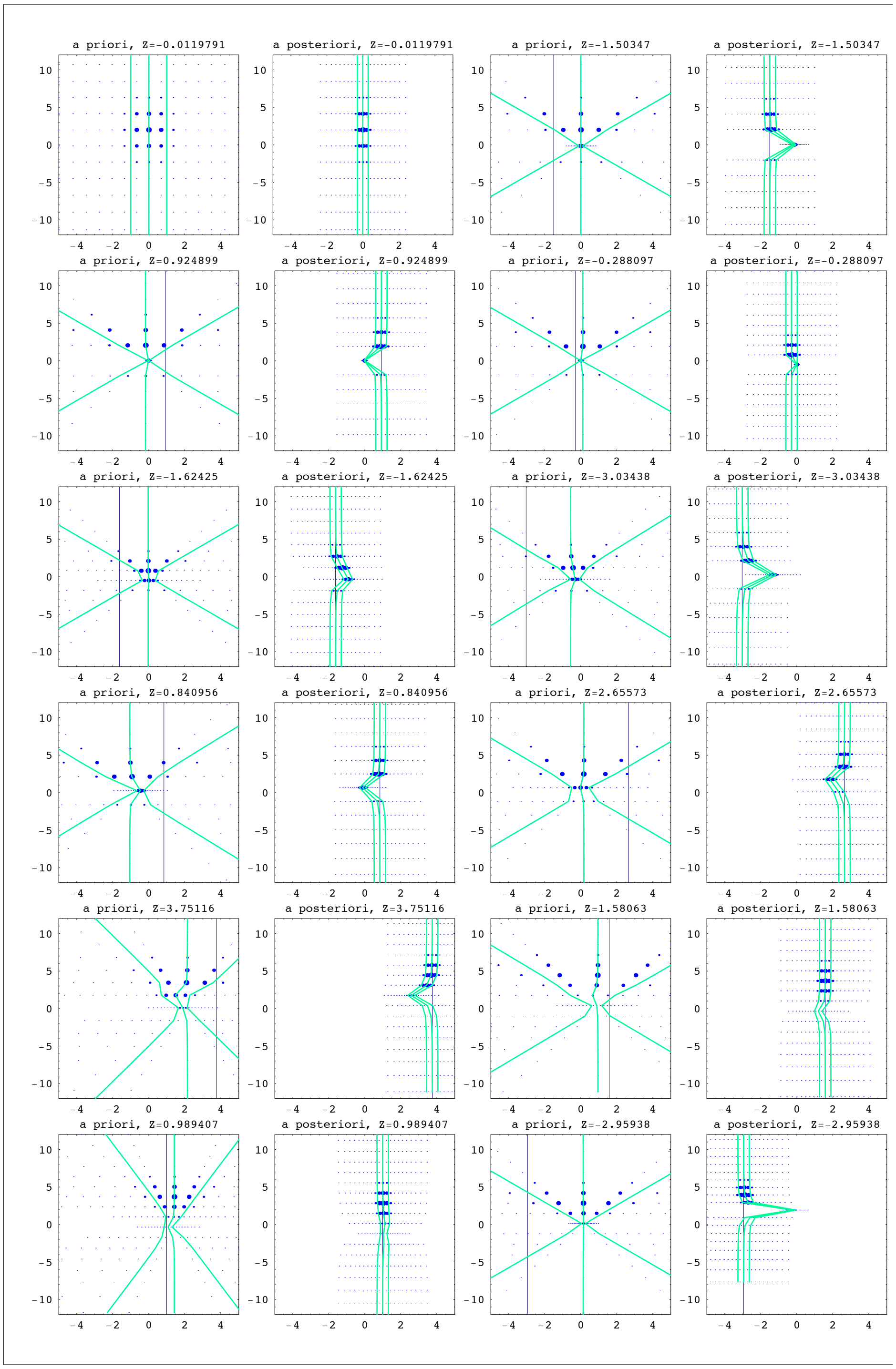

Figure 7: Gauss–Hermite sample points for the Ornstein-Uhlenbeck process ($L = 21, M = 21$). Also displayed is the conditional mean and standard deviation $\mu_1(\eta_{2m}) \pm \Sigma_1^{1/2}(\eta_{2m})$.

are Gauss–Hermite sample points for the integration over y_2 and y_1 (conditional on the values $y_2 = \eta_{2m}$). Thus, one has y_1-sample points η_{1lm} for each y_2-coordinate $\eta_{2m}; l = 1, \ldots, L; m = 1 \ldots M$ (cf. fig. 7).
Similarly,

$$\begin{aligned} E[f_2(y_1, y_2, t)] &= \int\int f_2(y_1, y_2, t)\phi(y_1; \mu_1(y_2), \Sigma_1(y_2)) \\ &\times \phi(y_2; \mu_2, \Sigma_2) dy_1 dy_2 \\ &\approx \sum_{l,m=1}^{L,M} f_2(\eta_{1lm}, \eta_{2m}, t) w_{1l} w_{2m}. \end{aligned} \tag{43}$$

Now it is assumed that $E[y_1(t + \delta t)|y_2(t)] \approx E[y_1(t + \delta t)|y_2(t + \delta t)]$ etc. and using this approximation the time update is continued over the complete time interval $[t_i, t_{i+1}]$.

4.2 Measurement update

At time t_{i+1}, new measurements z_{i+1} come in, which are incorporated by using the Bayes formula (setting $y_{i+1} := y(t_{i+1})$ etc.)

$$p(y_{1,i+1}, y_{2,i+1}|z_{i+1}, Z^i) = \frac{p(z_{i+1}|y_{1,i+1}, y_{2,i+1})p(y_{1,i+1}, y_{2,i+1}|Z^i)}{p(z_{i+1}|Z^i)}. \tag{44}$$

The product of the measurement density

$$p(z_{i+1}|y_{1,i+1}, y_{2,i+1}) = \phi(z_{i+1}; h(y_{1,i+1}, y_{2,i+1}, t_{i+1}), R_{i+1}) \tag{45}$$

with the a priori distribution

$$p(y_{1,i+1}, y_{2,i+1}|Z^i) = \phi(y_{1,i+1}|y_{2,i+1}, Z^i) * \phi(y_{2,i+1}|Z^i) \tag{46}$$

can be evaluated approximately by the normal correlation update as

$$\phi(y_{1,i+1}|y_{2,i+1}, Z^{i+1}) * \phi(z_{i+1}|y_{2,i+1}, Z^i) * \phi(y_{2,i+1}|Z^i) \tag{47}$$

where $\phi(z_{i+1}|y_{2,i+1}, Z^i) = \phi(z_{i+1}; E[h|y_{2,i+1}, Z^i], \mathrm{Var}[h|y_{2,i+1}, Z^i] + R_{i+1})$ is the conditional likelihood of z_{i+1} given $y_{2,i+1}$. The moments of the posterior of $y_1|y_2$ are given by

$$\begin{aligned} E[y_{1,i+1}|y_{2,i+1}, Z^{i+1}] &= E[y_{1,i+1}|F_{i+1}] + \mathrm{Cov}[y_{1,i+1}, z_{i+1}|F_{i+1}] \\ &\times \mathrm{Var}[z_{i+1}|F_{i+1}]^-(z_{i+1} - E[z_{i+1}|F_{i+1}]) \\ \mathrm{Var}[y_{1,i+1}|y_{2,i+1}, Z^{i+1}] &= \mathrm{Var}[y_{1,i+1}|F_{i+1}] - \mathrm{Cov}[y_{1,i+1}, z_{i+1}|F_{i+1}] \\ &\times \mathrm{Var}[z_{i+1}|F_{i+1}]^-\mathrm{Cov}[z_{i+1}, y_{1,i+1}|F_{i+1}] \end{aligned}$$

where $F_{i+1} = \{y_{2,i+1}, Z^i\}$ is shorthand for the conditioning variables. Now the moments of the a priori distribution ($E[y_{1,i+1}|y_{2,i+1}, Z^i], \mathrm{Var}[y_{1,i+1}|y_{2,i+1}, Z^i]$,

$E[y_{2,i+1}|Z^i], \mathrm{Var}[y_{2,i+1}|Z^i]$) are known from the time update and the expectations can be evaluated by Gauss–Hermite integration again. For example

$$
\begin{aligned}
E[z_{i+1}|y_{2,i+1}, Z^i] &= E[h(y_{1,i+1}, y_{2,i+1}, t_{i+1})|y_{2,i+1}, Z^i] && (48)\\
&\approx \sum_{l=1}^{L} h(\eta_{1lm}, \eta_{2m}, t_{i+1}) w_{1l} && (49)
\end{aligned}
$$

where again

$$
\begin{aligned}
\eta_{2m} &= \mu_2 + \Sigma_2^{1/2}\zeta_{2m} && (50)\\
\eta_{1lm} &= \mu_1(\eta_{2m}) + \Sigma_1^{1/2}(\eta_{2m})\zeta_{1l} && (51)
\end{aligned}
$$

are the Gauss–Hermite sample points evaluated at the a priori moments ($\mu_2 = E[y_{2,i+1}|Z^i]$, $\mu_1(y_2) = E[y_{1,i+1}|y_{2,i+1}, Z^i]$) etc.
The posteriori distribution for y_2 is given by (cf. 47)

$$
p(y_{2,i+1}|Z^{i+1}) = \phi(z_{i+1}|y_{2,i+1}, Z^i)\phi(y_{2,i+1}|Z^i)/p(z_{i+1}|Z^i). \qquad (52)
$$

Now, since

$$
\begin{aligned}
E[z_{i+1}|y_{2,i+1}, Z^i] &= E[h(y_{1,i+1}, y_{2,i+1}, t_{i+1})|y_{2,i+1}, Z^i] && (53)\\
\mathrm{Var}[z_{i+1}|y_{2,i+1}, Z^i] &= \mathrm{Var}[h(y_{1,i+1}, y_{2,i+1}, t_{i+1})|y_{2,i+1}, Z^i] + R_{i+1} && (54)
\end{aligned}
$$

are in general nonlinear functions of $y_{2,i+1}$ (cf. example 2; 17), the measurement z_{i+1} is informative for the 'volatility state' $y_{2,i+1}$ and one obtains the likelihood of observation z_{i+1} and the posterior mean of y_2

$$
\begin{aligned}
p(z_{i+1}|Z^i) &= \int \phi(z_{i+1}|y_{2,i+1}, Z^i)\phi(y_{2,i+1}|Z^i) dy_{2,i+1} \\
&\approx \sum_{m=1}^{M} \phi(z_{i+1}|\eta_{2m}, Z^i) w_{2m} && (55)\\
E[y_{2,i+1}|Z^{i+1}] &= p(z_{i+1}|Z^i)^{-1} \int y_{2,i+1}\phi(z_{i+1}|y_{2,i+1}, Z^i) \\
&\times \phi(y_{2,i+1}|Z^i) dy_{2,i+1} \\
&\approx p(z_{i+1}|Z^i)^{-1} \sum_{m=1}^{M} \eta_{2m}\phi(z_{i+1}|\eta_{2m}, Z^i) w_{2m} && (56)
\end{aligned}
$$

(analogously for $\mathrm{Var}[y_{2,i+1}|Z^{i+1}]$).
Starting from the a priori moments $\mu_1(y_2(t_0)) = E[y_1(t_0)|y_2(t_0)], \mu_2 = E[y_2(t_0)]$ etc. one obtains a recursive sequence of measurement and time updates for the moments and the Gauss–Hermite sample points (cf. fig. 7).

4.3 Complete filter

Putting together the measurement update and the time update, one obtains a recursive sequence of moments (26-29) at the measurement times t_i and for the time points $\tau_j = t_i + j\delta t, j = 1, \ldots, (t_{i+1} - t_i)/\delta t$ in between. The unconditional moments (w.r.t. y_2, dropping Z^i) can be computed from the filter terms as

$$\begin{aligned} E[y_1(t)] &= E[E[y_1(t)|y_2(t)]] &\quad (57)\\ \mathrm{Var}[y_1(t)] &= E[(y_1(t) - \mu_1(t))(y_1(t) - \mu_1(t))'] \\ &= E[\mathrm{Var}(y_1(t)|y_2(t))] + \mathrm{Var}(E[y_1(t)|y_2(t)]) &\quad (58) \end{aligned}$$

(residual variance + explained variance).
For the starting values I used $\mu_1(y_2(t_0)) = \mu_1, \Sigma_1(y_2(t_0)) = \Sigma_1$ (independent of y_2) and $\mu_2 = E[y_2(t_0)], \Sigma_2 = \mathrm{Var}[y_2(t_0)]$. Thus, the prior $p_0 = p(y_1(t_0)|y_2(t_0))p(y_2(t_0))$ is a Gaussian distribution with uncorrelated states $y_1(t_0), y_2(t_0)$. After the first measurement update one obtains η_{2m} and $\mu_1(\eta_{2m})$, i.e. the unknown function $\mu_1(y_2)$ is determined on the sample points η_{2m} (same for $\Sigma_1(y_2)$). Iterating, one obtains the regression functions $\mu_1(y_2)$ etc. in a nonparametric way.

5 Example 3: ML vs. recursive Bayesian estimation

In example 2, the Ornstein-Uhlenbeck process was discussed. It is interesting to compare ML estimation of σ with recursive Bayesian filtering. As noted, the likelihood can be computed exactly by using the Kalman filter, i.e.

$$\begin{aligned} \mu_{i+1|i} &= \lambda_i \mu_{i|i} \\ \Sigma_{i+1|i} &= \lambda_i \Sigma_{i|i} \lambda_i' + \Omega_i \\ \mu_{i+1|i+1} &= \mu_{i+1|i} + K_i(z_{i+1} - \mu_{i+1|i}) \\ K_i &= \Sigma_{i+1|i}(\Sigma_{i+1|i} + R)^- \\ \Sigma_{i+1|i+1} &= (I - K_i)\Sigma_{i+1|i} \\ L_{i+1}(z_{i+1}) &= \phi(z_{i+1}; \mu_{i+1|i}, \Sigma_{i+1|i} + R). \end{aligned}$$

In the formulae above, K_i is the Kalman gain and $\Omega_i = \mathrm{Var}(\sigma u_i) = \sigma^2(1 - \exp(2\lambda \Delta t_i))/(2\lambda_i)$ is the variance of the system error u_i (cf. 13). As usual, $\mu_{i+1|i} = E[y_{i+1}|Z^i]$ etc. denotes the conditional expectations. Starting from a flat prior $\Sigma_{0|-1} = 10$ one obtains the ML estimator by maximizing $l(\sigma) = \sum_{i=0}^{t} \log(L_i)$. The estimator $\hat{\sigma}(t)$ was computed recursively for the data set $Z^t, t = 1, \ldots, T$.
Fig. 8 shows a comparision of the sequential estimates of σ using maximum likelihood (ML), the newly developed CGHF, the generalized Gauss-Hermite

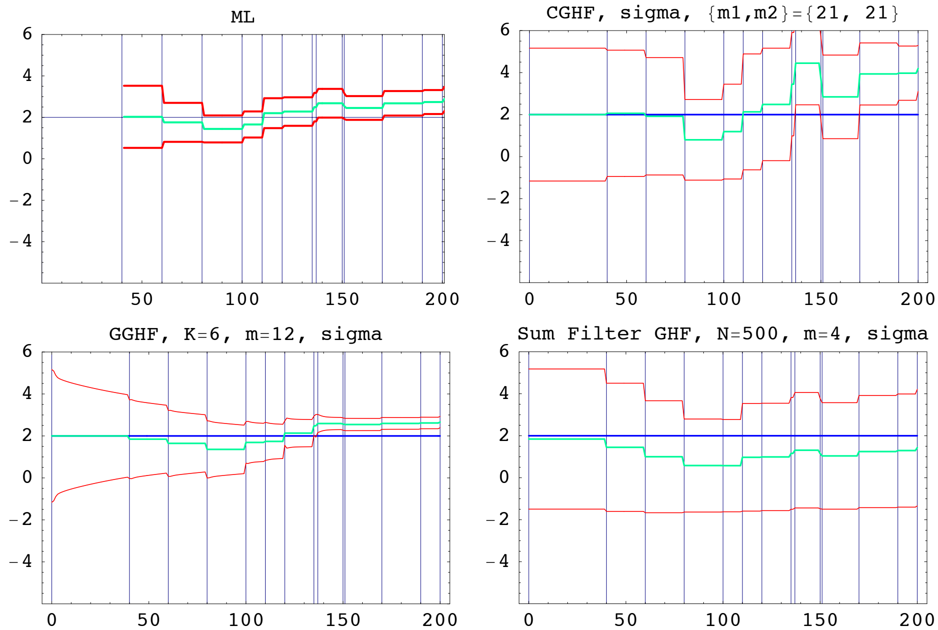

Figure 8: Sequential estimation of the parameter σ (discrete sampling): Maximum likelihood (top, left), CGHF (top, right), GGHF (bottom, left), Gaussian sum filter (bottom, right). Estimates $\hat{\sigma}(t) \pm \mathrm{std}(\hat{\sigma}(t))$ (see text).

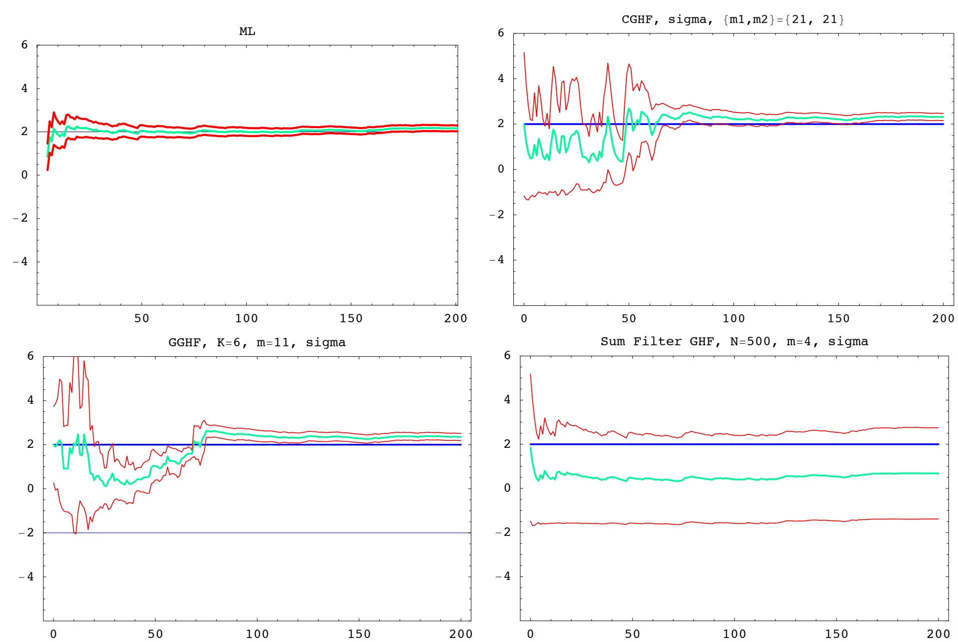

Figure 9: Sequential estimation of the parameter σ (quasi-continuous sampling): ML (top, left), CGHF (top, right), GGHF (bottom, left), Gaussian sum filter (bottom, right). Estimates $\hat{\sigma}(t) \pm \mathrm{std}(\hat{\sigma}(t))$ (see text).

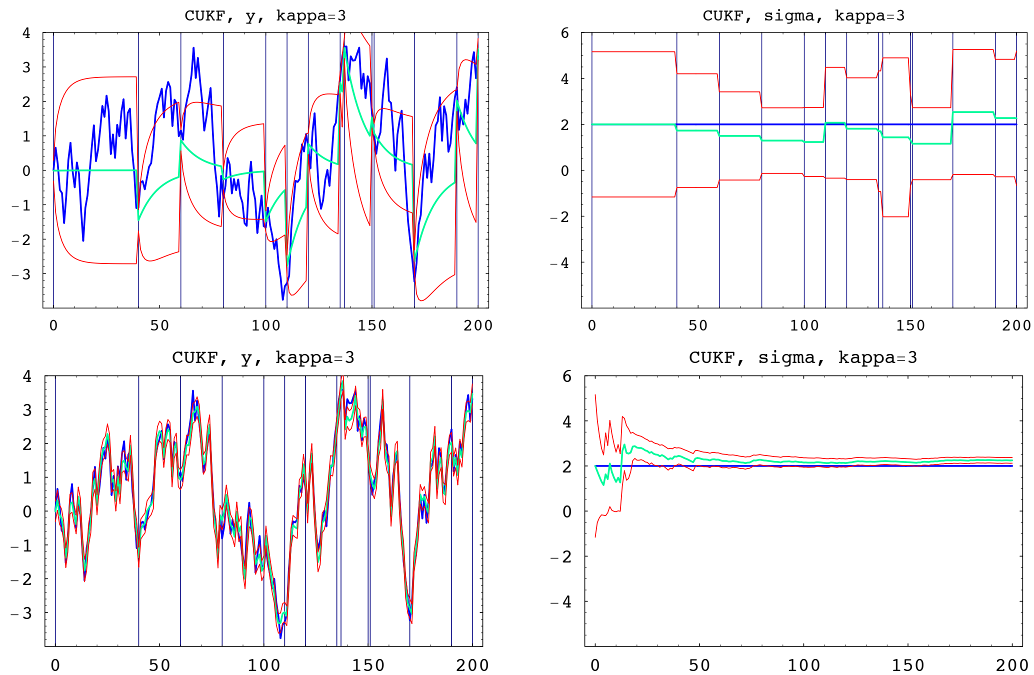

Figure 10: CUKF ($\kappa = 3$): Sequential estimation of the parameter σ: discrete sampling (top), quasi-continuous sampling (bottom).

filter (GGHF, cf. Singer; 2006b) and the Gaussian sum filter, implemented with GHF updates (cf. Ito and Xiong; 2000). If more measurements are used, the convergence is more quickly (cf. fig. 9, using all data). Clearly, ML works best, but for each estimate $\hat{\sigma}(t)$, a nonlinear optimization algorithm has to be solved (I used a quasi Newton algorithm with BFGS secant update and numerical score). The estimate $\hat{\sigma}(t-1)$ was used as starting value for the next maximization at time t. In contrast, the nonlinear filters work sequentially without iterative optimization. In comparision, the CGHF works best and is the fastest. Alternatively, the integrations can be done using the unscented transform leading to a conditional unscented Kalman filter (CUKF). The CPU times for the several algorithms are shown in table 1.

6 Conclusion

We have shown how the filtering of volatility parameters can be achieved by a simple assumption. Instead of taking the joint Gaussian $\phi(y_1, y_2)$, the Gaussian product $\phi(y_1|y_2)\phi(y_2)$ was used leading to a nonlinear dependence of $E[y_1|y_2]$ and $\mathrm{Var}(y_1|y_2)$ on y_2. In contrast, a joint Gaussian assumption can only accomodate a linear regression $E[y_1|y_2] = a + by_2$ and a constant variance $\mathrm{Var}[y_1|y_2] = \mathrm{Var}[y_1] - \mathrm{Cov}(y_1, y_2)\mathrm{Var}(y_2)^{-}\mathrm{Cov}(y_2, y_1)$. However, in stochastic volatility models, the variance of y_1 is dependent on y_2. The Gaus-

method	CPU time (sec)
discrete sampling	
CGHF $(M, L) = (3, 3)$	0.45
CGHF $(M, L) = (11, 11)$	1.89
CGHF $(M, L) = (21, 21)$	5.27
GGHF $(K, M) = (6, 12)$	27.27
Sum filter GHF $(N, M) = (500, 4)$	86.05
CUKF $(\kappa = 3)$	0.45
continuous sampling	
CGHF $(M, L) = (3, 3)$	6.32
CGHF $(M, L) = (11, 11)$	22.69
CGHF $(M, L) = (21, 21)$	55.96
GGHF $(K, M) = (6, 11)$	78.58
Sum filter GHF $(N, M) = (500, 4)$	145.56
CUKF $(\kappa = 3)$	6.37

Table 1: CPU times of several algorithms.

sian product is the simplest assumption for this type of nonlinear problems and leads to an efficient algorithm using Gauss-Hermite quadrature. Alternatively, the integrations can be done using the unscented transform leading to a conditional unscented Kalman filter (CUKF).

Appendix A: Gauss–Hermite integration

The moment equations of the (C)GHF require the computation of expectations of the type $E[f(Y)]$, where Y is a random variable with density $p(y)$. For the Gaussian filter, one may assume that the true $p(y)$ is approximated by a Gaussian distribution $\phi(y; \mu, \sigma^2)$ with the same mean μ and variance σ^2. Then, the Gaussian integral

$$\begin{aligned} E_\phi[f(Y)] &= \int f(y)\phi(y; \mu, \sigma^2)dy = \int f(\mu + \sigma z)\phi(z; 0, 1)dz \\ &\approx \sum_{l=1}^{m} f(\mu + \sigma\zeta_l)w_l = \sum_{l=1}^{m} f(\eta_l)w_l \end{aligned}$$

may be approximated by Gauss–Hermite quadrature (Ito and Xiong; 2000) Here, (ζ_l, w_l) are quadrature points and weights, respectively. If such an approximation is used, one obtains the Gauss–Hermite filter (GHF). Filters using Gaussian densities are called Gaussian filters (GF). More generally, the density may be approximated by the product of conditionally Gaussian densities $\phi(y_1|y_2)\phi(y_2)$ (CGHF) which again yields integrals w.r.t. the Gaussian density i.e. $E[f(Y)] = \int f(y_1, y_2)\phi(y_1|y_2)\phi(y_2)dy_1dy_2$.

In the multivariate case, the integration is performed using standardization with some matrix square root (e.g. the Cholesky decomposition)

$$\begin{aligned} E_\phi[f(Y)] &= \int f(y)\phi(y;\mu,\Sigma)dy \\ &= \int f(\mu+\Sigma^{1/2}z)\phi(z;0,I)dz_1...dz_p \\ &\approx \sum_{l_1,...,l_p} f(\mu+\Sigma^{1/2}\{\zeta_{l_1},...,\zeta_{l_p}\})w_{l_1,...,l_p} \\ &= \sum_{l_1,...,l_p} f(\eta_{l_1},...,\eta_{l_p})w_{l_1,...,l_p}, \end{aligned}$$

since $\phi(z;0,I) = \phi(z_1;0,1)...\phi(z_p;0,1)$ allows stepwise application of the univariate quadrature formula and $\{\zeta_{l_1},...,\zeta_{l_p}\}, l_j = 1,...,m$, is the p-tupel of Gauss–Hermite quadrature points with weights $w_{l_1,...,l_p} = w_{l_1}...w_{l_p}$.

Appendix B: Continuous-discrete filtering scheme

The Gauss-Hermite filter (GHF) is a recursive sequence of time and measurement updates for the conditional moments $\mu(t|t_i) = E[y(t)|Z^i]$ and $\Sigma(t|t_i) = \text{Var}[y(t)|Z^i]$, where expectation values are computed according to (59) using Gauss-Hermite quadrature ($^-$ denoting the generalized inverse; $\Omega := gg'$, cf. eq. 6):

Initial condition: $t = t_0$

$$\begin{aligned} \mu(t_0|t_0) &= \mu + \text{Cov}(y_0,h_0) \\ &\times\ (\text{Var}(h_0)+R(t_0))^-(z_0 - E[h_0]) \\ \Sigma(t_0|t_0) &= \Sigma - \text{Cov}(y_0,h_0) \\ &\times\ (\text{Var}(h_0)+R(t_0))^-\text{Cov}(h_0,y_0) \\ L_0 &= \phi(z_0;E[h_0],\text{Var}(h_0)+R(t_0)) \\ \eta_l &= \eta_l(\mu,\Sigma); \mu = E[y_0], \Sigma = \text{Var}(y_0) \text{ (quadrature points).} \end{aligned}$$

$i = 0,\dots,T-1$:
Time update: $t \in [t_i, t_{i+1}]$

$$\begin{aligned} \tau_j &= t_i + j\delta t; j = 0,...,J_i - 1 = (t_{i+1}-t_i)/\delta t - 1 \\ \mu(\tau_{j+1}|t_i) &= \mu(\tau_j|t_i) + E[f(y(\tau_j),\tau_j)|Z^i]\delta t \\ \Sigma(\tau_{j+1}|t_i) &= \Sigma(\tau_j|t_i) + \{\text{Cov}[f(y(\tau_j),\tau_j),y(\tau_j)|Z^i] + \\ &\quad \text{Cov}[y(\tau_j),f(y(\tau_j),\tau_j)|Z^i] + E[\Omega(y(\tau_j),\tau_j)|Z^i]\}\delta t \\ \eta_l &= \eta_l(\mu(\tau_j|t_i),\Sigma(\tau_j|t_i)) \text{ (quadrature points)} \end{aligned}$$

Measurement update: $t = t_{i+1}$

$$
\begin{aligned}
\mu(t_{i+1}|t_{i+1}) &= \mu(t_{i+1}|t_i) + \mathrm{Cov}(y_{i+1}, h_{i+1}|Z^i) \\
&\times \ (\mathrm{Var}(h_{i+1}|Z^i) + R(t_{i+1}))^-(z_{i+1} - E[h_{i+1}|Z^i]) \\
\Sigma(t_{i+1}|t_{i+1}) &= \Sigma(t_{i+1}|t_i) - \mathrm{Cov}(y_{i+1}, h_{i+1}|Z^i) \\
&\times \ (\mathrm{Var}(h_{i+1}|Z^i) + R(t_{i+1}))^-\mathrm{Cov}(h_{i+1}, y_{i+1}|Z^i) \\
L_{i+1} &= \phi(z_{i+1}; E[h_{i+1}|Z^i], \mathrm{Var}(h_{i+1}|Z^i) + R(t_{i+1})) \\
\eta_l &= \eta_l(\mu(t_{i+1}|t_i), \Sigma(t_{i+1}|t_i)) \text{ (quadrature points)}
\end{aligned}
$$

Remarks

1. The discretization interval δt is a small value controlling the accuracy of the Euler scheme implicit in the time update. Since the quadrature points are functions of the mean and variance, the moment equations (35,37) are a coupled system of nonlinear differential equations for the sample points of the Gauss-Hermite scheme. Therefore, other approximation methods such as the Heun scheme or higher order Runge-Kutta schemes could be used.

2. The time update neglects second order terms. Inclusion of $E[f - E(f)][f - E(f)]'\delta t^2$ leads to a positive semidefinite update, which is numerically more stable.

3. The measurement update is the optimal linear update (normal correlation; Liptser and Shiryayev, 2001, ch. 13, theorem 13.1, lemma 14.1)

$$
\begin{aligned}
E[y_{i+1}|z_{i+1}, Z^i] &= E[y_{i+1}|Z^i] + \mathrm{Cov}(y_{i+1}, z_{i+1}|Z^i) \\
&\times \ \mathrm{Var}(z_{i+1}|Z^i)^-(z_{i+1} - E[z_{i+1}|Z^i]) \\
\mathrm{Var}[y_{i+1}|z_{i+1}, Z^i] &= \mathrm{Var}[y_{i+1}|Z^i] - \mathrm{Cov}(y_{i+1}, z_{i+1}|Z^i) \\
&\times \ \mathrm{Var}(z_{i+1}|Z^i)^-\mathrm{Cov}(z_{i+1}, y_{i+1}|Z^i).
\end{aligned}
$$

with measurement equation (7) inserted and covariances computed by Gauss-Hermite integration. It is linear in z but includes the nonlinear measurements $z = h(y) + \epsilon$ in the expectation values and covariance terms. It does not require any Taylor expansions and can be used for discontinuous measurement functions as in threshold models (ordinal data). A direct implementation of the Bayes formula (44) would lead to the asymmetric a posteriori density

$$
\begin{aligned}
p(y_{i+1}|Z^{i+1}) &= \sum_{l=1}^{m} w_l^* \delta(y_{i+1} - \eta_l) \\
w_l^* &= w_l p(z_{i+1}|\eta_l) / \sum_{l=1}^{m} w_l p(z_{i+1}|\eta_l)
\end{aligned}
$$

where the a priori density is $p(y_{i+1}|Z^i) = \sum_{l=1}^m w_l \delta(y_{i+1} - \eta_l)$ with Gauss-Hermite sample points $\eta_l = \eta_l(\mu(t_{i+1}|t_i), \Sigma(t_{i+1}|t_i))$. Computing the a posteriori moments

$$\begin{aligned}
\mu(t_{i+1}|t_{i+1}) &= \int y_{i+1} p(y_{i+1}|Z^{i+1}) dy_{i+1} = \sum_{l=1}^m w_l^* \eta_l \\
\Sigma(t_{i+1}|t_{i+1}) &= \int (y_{i+1} - \mu(t_{i+1}|t_{i+1}) \\
&\times \; (y_{i+1} - \mu(t_{i+1}|t_{i+1})' p(y_{i+1}|Z^{i+1}) dy_{i+1} \\
&= \sum_{l=1}^m w_l^* [\eta_l - \mu(t_{i+1}|t_{i+1})][\eta_l - \mu(t_{i+1}|t_{i+1})]'
\end{aligned}$$

one can construct a symmetric a posteriori distribution with the same first and second moments.

4. Taylor expansion of f, Ω and h around μ leads to the usual EKF and SNF. Using sigma points instead of Gaussian quadrature points yields the unscented Kalman filter UKF (cf. Julier and Uhlmann; 1997; Julier et al.; 2000; Julier and Uhlmann; 2004; Singer; 2006a)

References

Aït-Sahalia, Y. (2002). Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach, *Econometrica* **70,1**: 223–262.

Alspach, D. and Sorenson, H. (1972). Nonlinear Bayesian estimation using Gaussian sum approximations, *IEEE Transactions on Automatic Control* **17**: 439–448.

Arnold, L. (1974). *Stochastic Differential Equations*, John Wiley, New York.

Hull, J. and White, A. (1987). The Pricing of Options with Stochastic Volatilities, *Journal of Finance* **XLII,2**: 281–300.

Ito, K. and Xiong, K. (2000). Gaussian filters for nonlinear filtering problems, *IEEE Transactions on Automatic Control* **45, 5**: 910–927.

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory*, Academic Press, New York.

Julier, S. and Uhlmann, J. (1997). A new extension of the Kalman filter to nonlinear systems, *The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Control*, Orlando, Florida.

Julier, S. and Uhlmann, J. (2004). Unscented Kalman Filtering and Nonlinear Estimation, *Proceedings of the IEEE* **92, 3**: 401–422.

Julier, S., Uhlmann, J. and Durrant-White, H. F. (2000). A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators, *IEEE Transactions on Automatic Control* **45, 3**: 477–482.

Kloeden, P. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin.

Kuznetsov, P., Stratonovich, R. and Tikhonov, V. (1960). Quasi-moment functions in the theory of random processes, *Theory of Probability and its Applications* **V, 1**: 80–97.

Liptser, R. and Shiryayev, A. (2001). *Statistics of Random Processes, Volumes I and II*, second edn, Springer, New York, Heidelberg, Berlin.

Scott, L. (1987). Option pricing when the variance changes randomly: Theory, estimation, and an application, *Journal of Financial and Quantitative Analysis* **22**: 419–438.

Singer, H. (1993). Continuous-time dynamical systems with sampled data, errors of measurement and unobserved components, *Journal of Time Series Analysis* **14, 5**: 527–545.

Singer, H. (1995). Analytical score function for irregularly sampled continuous time stochastic processes with control variables and missing values, *Econometric Theory* **11**: 721–735.

Singer, H. (2006a). Continuous-Discrete Unscented Kalman Filtering, *Diskussionsbeiträge Fachbereich Wirtschaftswissenschaft 384*, FernUniversität in Hagen. http://www.fernuni-hagen.de/FBWIWI/forschung/beitraege/pdf/db384.pdf.

Singer, H. (2006b). Generalized Gauss-Hermite Filtering for Multivariate Diffusion Processes, *Diskussionsbeiträge Fachbereich Wirtschaftswissenschaft 402*, FernUniversität in Hagen. http://www.fernuni-hagen.de/FBWIWI/forschung/beitraege/pdf/db402.pdf.

Singer, H. (2008). Generalized Gauss-Hermite Filtering, *Advances in Statistical Analysis* **92, 2**: 179–195.