Calibrated Adaptive Probabilistic ODE Solvers Supplementary Materials

A Proof of Proposition 1

Proof. The proof is structured as follows. First, we show by induction that an initial covariance $\Sigma_0 = \breve{\Sigma}_0 \otimes \Gamma$ implies covariances $\Sigma_n^P = \breve{\Sigma}_n^p \otimes \Gamma$, $\Sigma_n^F = \breve{\Sigma}_n^F \otimes \Gamma$, and $S_n = \breve{S}_n \cdot \Gamma$, for all n. Then, for measurement covariances S_n of such form, we can compute the (quasi) maximum likelihood estimate $\hat{\Gamma}$.

Assume $\Sigma_{n-1}^F = \breve{\Sigma}_{n-1}^F \otimes \Gamma$. Using the mixed product property and the associativity of the Kronecker product, the covariance of the prediction Σ_n^P can be written as

$$\Sigma_n^P = A_n \Sigma_{n-1}^F A_n^\intercal + Q_n = \left(\breve{A}_n \otimes I_d\right)\left(\breve{\Sigma}_{n-1}^F \otimes \Gamma\right)\left(\breve{A}_n \otimes I_d\right)^\intercal + \left(\breve{Q}_n \otimes \Gamma\right) = \left(\breve{A}_n \breve{\Sigma}_{n-1}^F \breve{A}_n^\intercal + \breve{Q}_n\right) \otimes \Gamma = \breve{\Sigma}_n^P \otimes \Gamma,$$

where $\breve{\Sigma}_n^P := \breve{A}_n \breve{\Sigma}_{n-1}^F \breve{A}_n^\intercal + \breve{Q}_n$. Next, using $H_n = e_1^\intercal \otimes I_d$ (EKF0), the measurement covariance S_n is given by

$$S_n = H_n \Sigma_n^P H_n^\intercal = (e_1^\intercal \otimes I_d)\left(\breve{\Sigma}_n^P \otimes \Gamma\right)(e_1^\intercal \otimes I_d)^\intercal = \left(e_1^\intercal \breve{\Sigma}_n^P e_1\right) \otimes \Gamma = \breve{S}_n \cdot \Gamma.$$

Finally, the filtering covariance can be computed with the update step Eq. (11) as

$$\begin{aligned}
\Sigma_n^F &= \Sigma_n^P - K_n S_n K_n \\
&= \Sigma_n^P - \Sigma_n^P H_n^\intercal S_n^{-1} H_n \Sigma_n^P \\
&= \left(\breve{\Sigma}_n^P \otimes \Gamma\right) - \left(\breve{\Sigma}_n^P \otimes \Gamma\right)(e_1^\intercal \otimes I_d)^\intercal \left(\breve{S}_n \otimes \Gamma\right)^{-1} (e_1^\intercal \otimes I_d)\left(\breve{\Sigma}_n^P \otimes \Gamma\right) \\
&= \left(\breve{\Sigma}_n^P - \breve{\Sigma}_n^P e_1 \breve{S}_n^{-1} e_1^\intercal \breve{\Sigma}_n^P\right) \otimes \Gamma \\
&= \breve{\Sigma}_n^F \otimes \Gamma,
\end{aligned}$$

with $\breve{\Sigma}_n^P := \breve{\Sigma}_n^P - \breve{\Sigma}_n^P e_1 \breve{S}_n^{-1} e_1^\intercal \breve{\Sigma}_n^P$. This concludes the first part of the proof.

It is left to compute the (quasi) MLE by maximizing the log-likelihood $\log p(z_{1:N}) = \log \prod_{n=1}^N \mathcal{N}(z_n; \hat{z}_n, S_n)$. For diagonal Γ and zero measurements $z_n = 0$, we obtain

$$\begin{aligned}
\hat{\Gamma} &= \arg\max_\Gamma \log p(z_{1:N}) \\
&= \arg\max_\Gamma \sum_{n=1}^N \log \mathcal{N}\left(0; \hat{z}_n, \breve{S}_n \cdot \Gamma\right) \\
&= \arg\max_\Gamma \sum_{n=1}^N \left(-\frac{1}{2}\log|\breve{S}_n \cdot \Gamma| - \frac{1}{2}\hat{z}_n^\intercal \left(\breve{S}_n \cdot \Gamma\right)^{-1} \hat{z}_n\right) \\
&= \arg\max_\Gamma \sum_{n=1}^N \left(-\frac{\log \breve{S}_n^d + \sum_{i=1}^d \log \Gamma_{ii}}{2} - \sum_{i=1}^d \frac{(\hat{z}_n)_i^2}{2\breve{S}_n \Gamma_{ii}}\right) \\
&= \arg\max_\Gamma \sum_{i=1}^d \left(-\frac{N \log \Gamma_{ii}}{2} - \sum_{n=1}^N \frac{(\hat{z}_n)_d^2}{2\breve{S}_n \Gamma_{ii}}\right).
\end{aligned}$$

Each diagonal element $\hat{\Gamma}_{ii}$ can be computed independently by taking the derivative setting it to zero:

$$0 = -\frac{N}{2\Gamma_{ii}} + \sum_{n=1}^{N} \frac{(\hat{z}_n)_d^2}{2\breve{S}_n \Gamma_{ii}^2}, \qquad \forall i \in \{1, \dots, d\}.$$

The solution to this equation provides the (quasi) maximum likelihood estimate for $\hat{\Gamma}$

$$\hat{\Gamma}_{ii} = \frac{1}{N} \sum_{n=1}^{N} \frac{(\hat{z}_n)_i^2}{\breve{S}_n}, \qquad \forall i \in \{1, \dots, d\}.$$

□

B Additional Experiments

These experiments include two additional IVPs.

Logistic Equation We consider an initial value problem, given by the logistic equation

$$\dot{y}(t) = r y(t)(1 - y(t)), \tag{35}$$

with parameter $r = 3$, integration interval $[0, 2.5]$, and initial value $y(0) = 0.1$. Its exact solution is given by

$$y^*(t) = \frac{\exp(rt)}{1/y_0 - 1 + \exp(rt)}. \tag{36}$$

Brusselator The Brusselator is a model for multi-molecular chemical reactions, given by the ODEs

$$\begin{aligned} \dot{y_1} &= 1 + y_1^2 y_2 - 4 y_1, \\ \dot{y_2} &= 3 y_1 - y_1^2 y_2. \end{aligned} \tag{37}$$

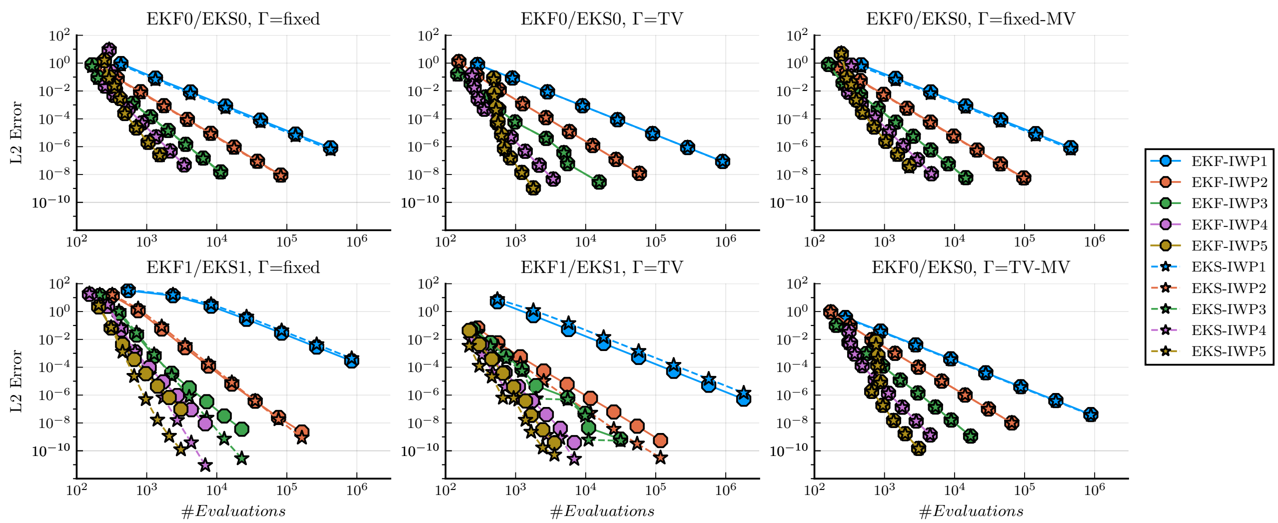
We consider an IVP with initial value $y(0) = [1.5, 3]$ on the time span $[0, 10]$.

B.1 Performance and Calibration on Additional Problems

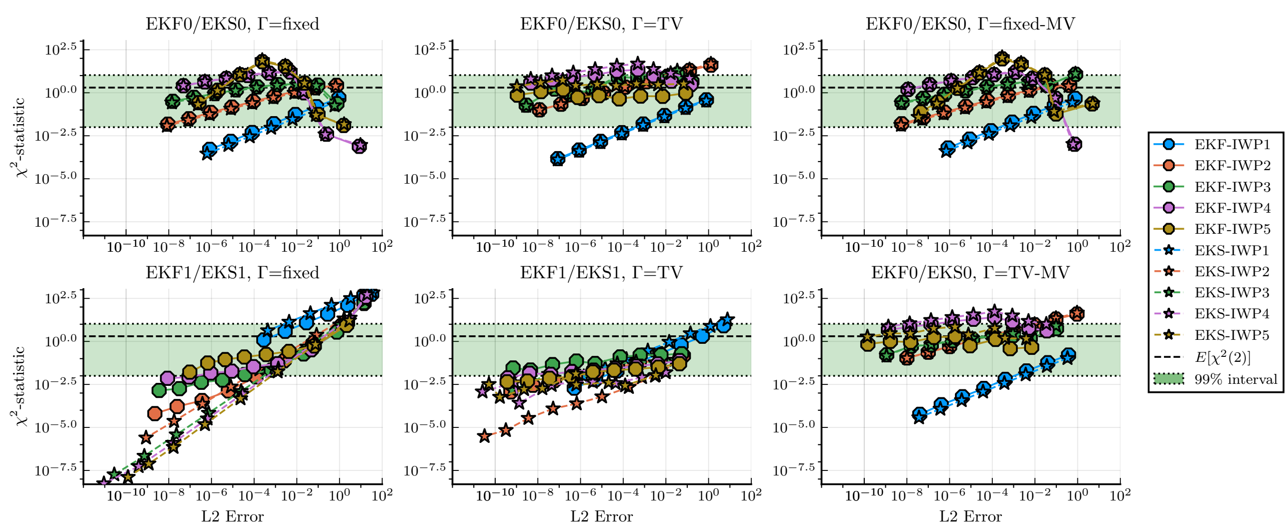
This section extends the results of Section 6.2 with evaluations on additional problems: Lotka-Volterra (Fig. 6), logistic equation (Fig. 7), FitzHugh-Nagumo (Fig. 8), and Brusselator (Fig. 9). In addition, all figures contain classic work-precision diagrams which visualize the relation of achieved error and number of evaluations, where both the evaluations of the vector field and of its Jacobian are counted.

B.2 Comparison to Dormand-Prince 4/5

Comparison of the probabilistic solvers to the classic Dormand-Prince 4/5 method on additional problems. Figure 10 presents the resulting work-precision diagrams. This extends the results of Section 6.3.

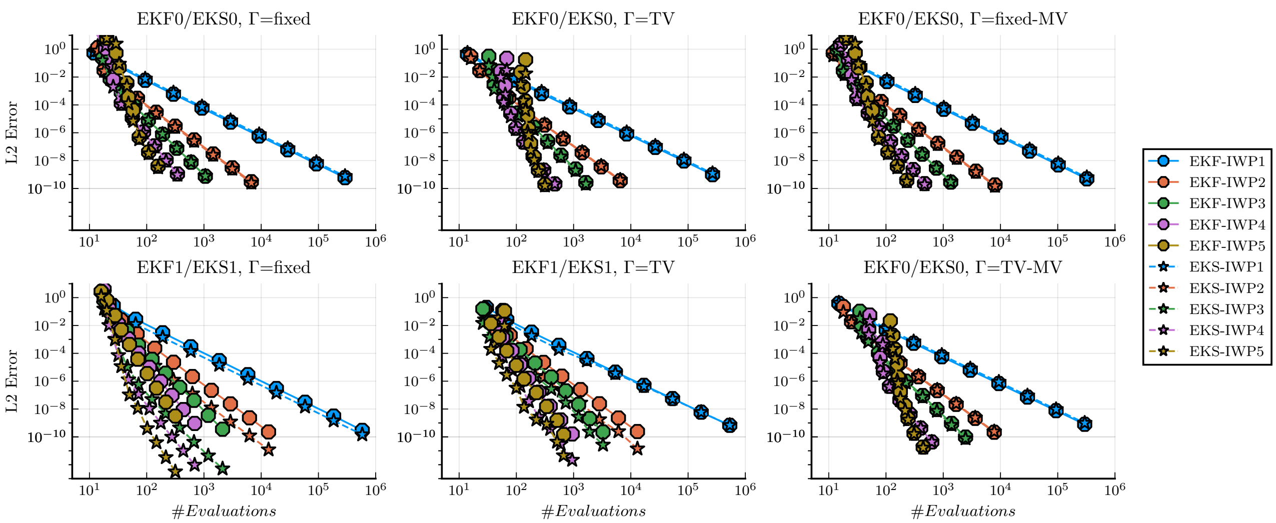


(a) Work-precision diagrams.

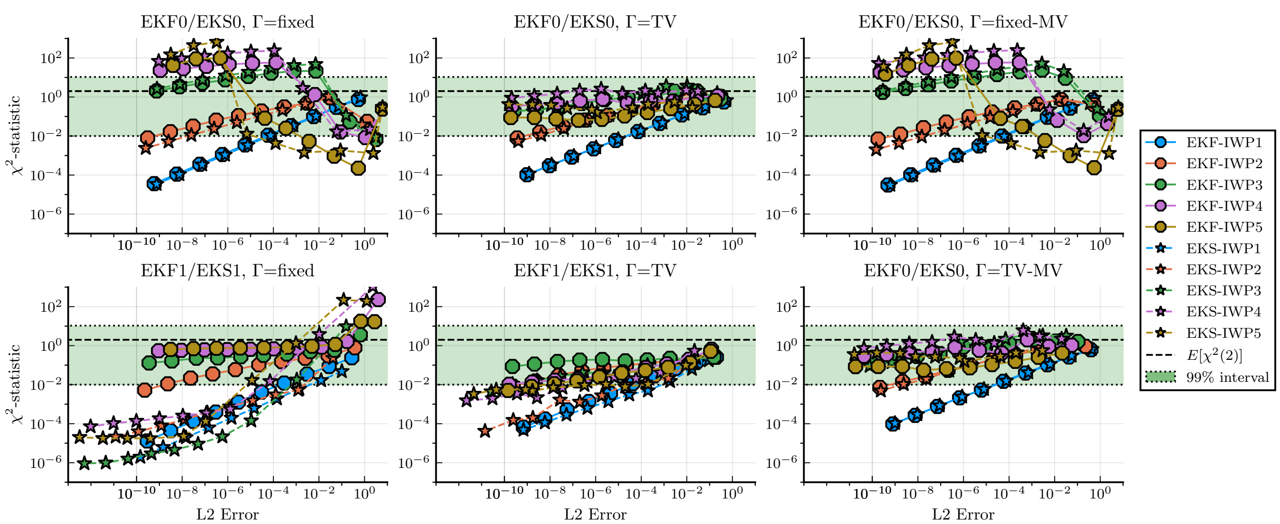


(b) Uncertainty Calibration.

Figure 6: *Accuracy and uncertainty calibration across configurations on the Lotka-Volterra equations.* In each subfigure, a specific combination of filtering algorithm (EKF0/EKS0 or EKS1/EKF1) and calibration method is evaluated, the latter including fixed and time-varying (TV) diffusion models, as well as their multivariate versions (fixed-MV, TV-MV).

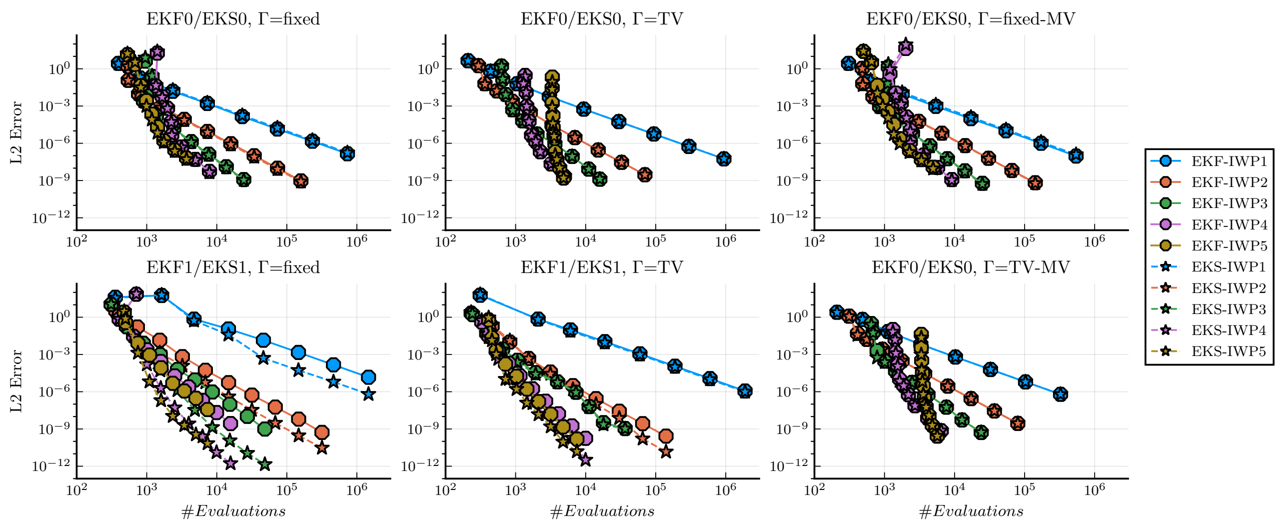


(a) Work-precision diagrams.

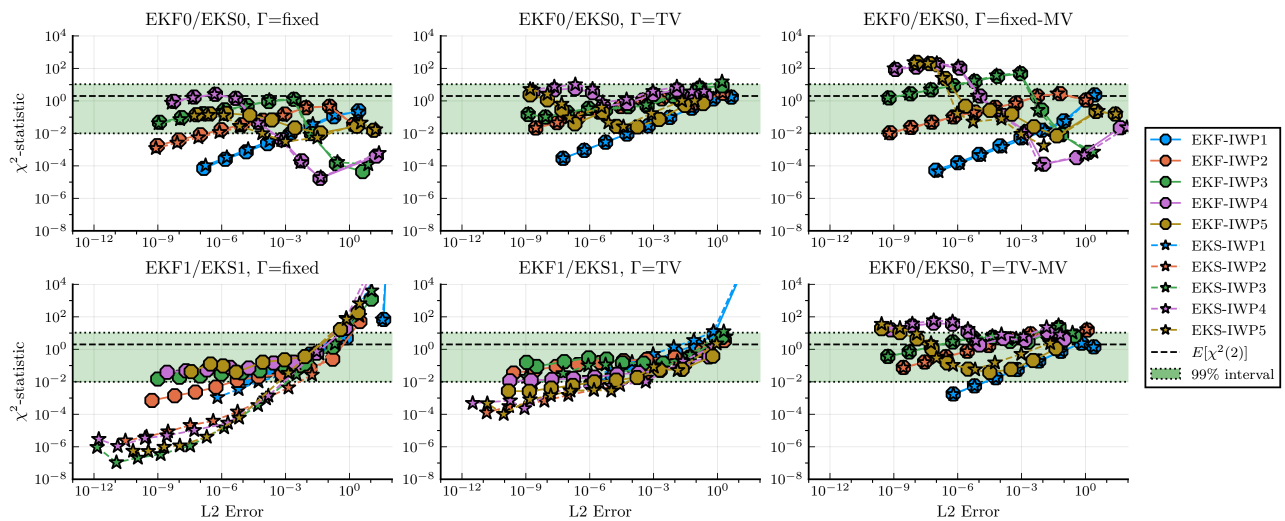


(b) Uncertainty Calibration.

Figure 7: *Accuracy and uncertainty calibration across configurations on the logistic equation.* In each subfigure, a specific combination of filtering algorithm (EKF0/EKS0 or EKS1/EKF1) and calibration method is evaluated, the latter including fixed and time-varying (TV) diffusion models, as well as their multivariate versions (fixed-MV, TV-MV).

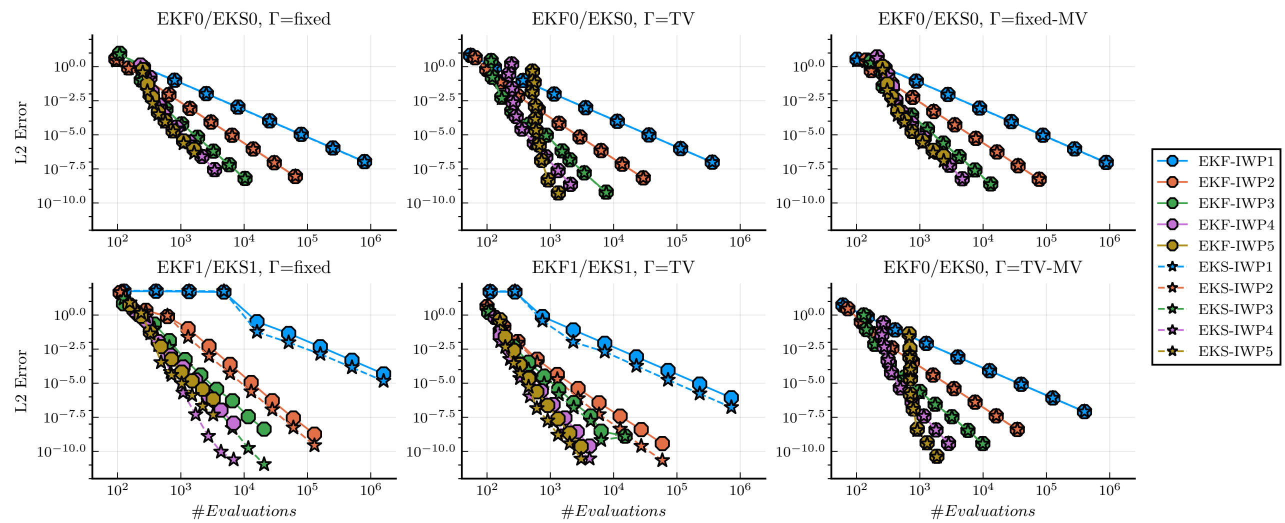


(a) Work-precision diagrams.

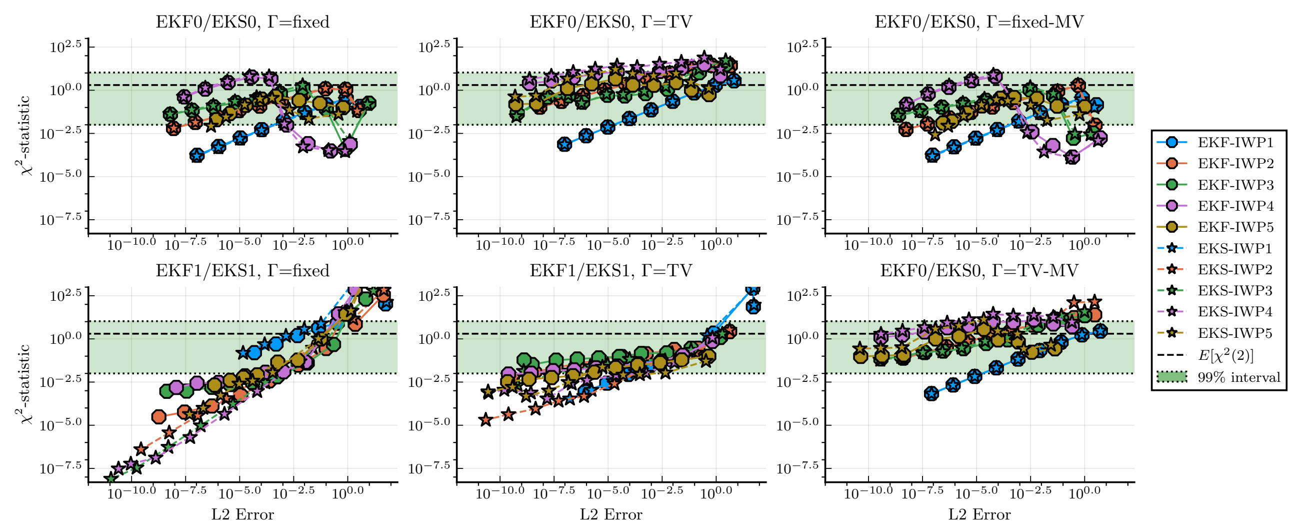


(b) Uncertainty Calibration.

Figure 8: *Accuracy and uncertainty calibration across configurations on the FitzHugh-Nagumo equations.* In each subfigure, a specific combination of filtering algorithm (EKF0/EKS0 or EKS1/EKF1) and calibration method is evaluated, the latter including fixed and time-varying (TV) diffusion models, as well as their multivariate versions (fixed-MV, TV-MV).

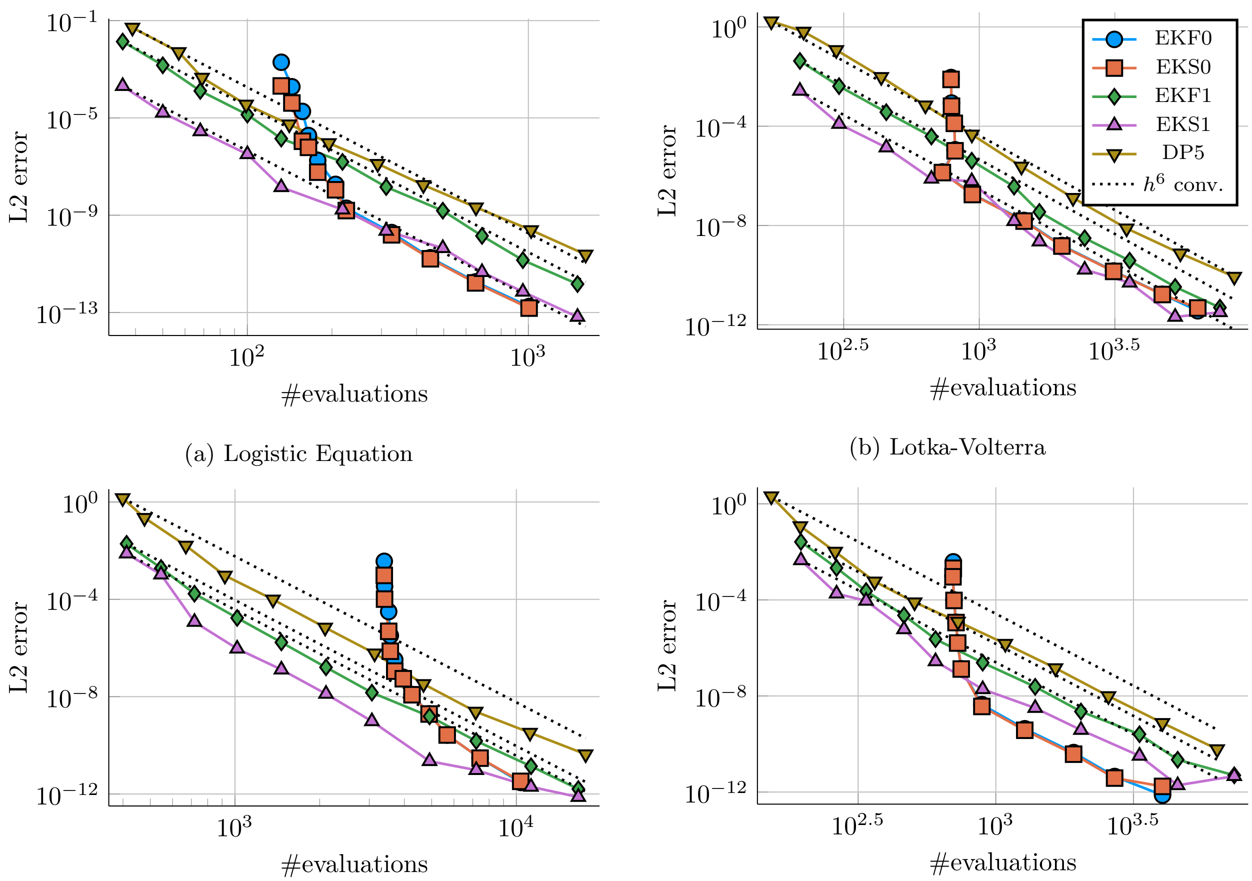


(a) Work-precision diagrams.



(b) Uncertainty Calibration.

Figure 9: *Accuracy and uncertainty calibration across configurations on the Brusselator equations.* In each subfigure, a specific combination of filtering algorithm (EKF0/EKS0 or EKS1/EKF1) and calibration method is evaluated, the latter including fixed and time-varying (TV) diffusion models, as well as their multivariate versions (fixed-MV, TV-MV).



(a) Logistic Equation

(b) Lotka-Volterra

(c) FitzHugh-Nagumo

(d) Brusselator

Figure 10: *Comparison to Dormand-Prince 4/5 (DP5).* All probabilistic ODE solvers use an IWP-5 prior and a scalar, time-varying diffusion model.