Limitations of Joint and Dual Nonlinear Kalman Estimators in Low-Cost Bioprocess Monitoring

Cristovão F. Iglesias Jr [1] **Luis Pessoa** [2] **Claudio Miceli** [2] **Miodrag Bolic** [1]

Abstract

The biopharmaceutical industry constantly presses for fast and low-cost bioprocess monitoring strategies. However, a recent study has shown that the Joint Extended Kalman Filter (JEKF) is inefficient in this monitoring type under biomanufacturing conditions. This work investigates the Dual Extended Kalman Filter (DEKF), Joint Unscented Kalman Filter (JUKF), and Joint Cubature Kalman Filter (JCKF) under these challenging conditions. Our theoretical analysis also reveals inefficiencies in DEKF, while our empirical tests using a synthetic dataset indicate that JUKF and JCKF only perform well with specific initial conditions for the state error covariance matrix, although with unconventional Kalman gain behavior. These results suggest nonlinear Kalman estimators in biomanufacturing still merit further investigation.

1. Introduction

Bioprocess monitoring is crucial for the development of biopharmaceutical processes (Rathore et al., 2023; Narayanan et al., 2023) and to achieve the goals of several initiatives, such as Quality by Design, Process Analytical tools, and Pharma 4.0 (Christakis et al., 2023; Mandenius & Gustavsson, 2015; Murugan, 2021; Sokolov et al., 2021). However, there is constant pressure for monitoring strategies that are fast and low cost. The literature indicates that soft sensors based on nonlinear Kalman estimators (NKE) with unstructured mechanistic model (UMM) can enable bioprocess monitoring (Christakis et al., 2023; Mandenius & Gustavsson, 2015; Murugan, 2021; Sokolov et al., 2021). NKE known as Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), and Cubature Kalman Filter (CKF) are composed of the error covariance matrices of the process (**Q**), measurement (**R**), and state (**P**). The integration of NKEs and UMM can be performed with joint and dual strategies to estimate the states and parameters of a UMM. The joint NKE (JNKE) considers a single joint state variable vector, which includes both the states and parameters of a UMM (Ljung, 1979; Kopp & Orford, 1963; Haykin & Haykin, 2001), and the dual NKE (DNKE) employs two consecutive NKE, separating the estimation of the system states and parameters (Ji & Brown, 2009). However, although Joint EKF (JEKF) has been used in many different bioprocess monitoring applications (Yousefi-Darani et al., 2021b; Herwig et al., 2021a;b; Iglesias Jr et al., 2022; Iglesias & Bolic, 2022), it has limitations for fast and low-cost bioprocess monitoring. It was recently demonstrated that joint estimation of states and unshared parameters (parameters not shared with other components of UMM) using an extended Kalman filter "fails" under the biomanufacturing conditions for fast and low-cost bioprocess monitoring (Iglesias Jr & Bolic, 2024). In biomanufacturing, a generic UMM allows modeling the macro-scale of a phenomenon without known mechanisms. This is extremely useful because, using JNKE with UMM, we can enable the real-time monitoring of bioprocesses with unknown mechanisms (Iglesias Jr et al., 2023). The UMM used in biomanufacturing typically comprises ordinary differential equations (ODE) with unshared parameters and weak terms. However, Iglesias et al. (Iglesias Jr & Bolic, 2024) proved that this characteristic of UMM in biomanufacturing, along with the use of **P**(t=0) and **Q** with uncorrelated elements, as well as the presence of only one measured state variable, are the conditions where the JEKF cannot estimate the unshared parameters and the state simultaneously. This issue, characterized by a zero Kalman gain for the unshared parameter during the entire process, can be addressed with the Specific initiAl coNdiTiOn (SANTO) approach where the idea is to add a quantity to the state error covariance between the measured state Variable (MSV) and an Unshared Parameter (UP) in **P**(t=0) (Iglesias Jr & Bolic, 2024). Since Dual EKF (DEKF) is a well-known alternative to Joint EKF, and the UKF and CKF are known for better performance than EKF (in addition to bypassing the need for first-order Taylor series expansion for linearization)(Khodarahmi & Maihami,

*Equal contribution [1]EECS, University of Ottawa, Ottawa, Canada [2]Federal University of Rio de Janeiro, RJ, Brazil. Correspondence to: Cristovão F. Iglesias Jr <cfrei096@uottawa.ca>.

LatinX in AI workshop at the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

2022), three questions remain unanswered.

In this work, we investigate the following research question: **RQ0)** Can DEKF estimate the unshared parameters and the state simultaneously under the biomanufacturing conditions for fast and low-cost bioprocess monitoring? **RQ1)** How are the performances of JUKF and JCKF under the same biomanufacturing conditions? **RQ2)** Can the SANTO approach improve the performance of JUKF and JCKF? Our main contributions are: 1) Theoretical analysis that proves the inability of DEKF to simultaneously estimate the unshared parameters and the state under the studied biomanufacturing conditions. 2) Empirical results, using a synthetic dataset of monoclonal Antibody (mAb) productions with new challenge scenery, showing that JUKF and JCKF performed better only with the SANTO approach, but with an unconventional behavior of Kalman gain in one of the studied cases. The codes and data used in this study are available in the gitHub repository to facilitate reproducibility [1].

2. Background

2.1. Conditions for fast and low-cost bioprocess monitoring

Fast and low-cost bioprocess monitoring can be defined as a set of methods designed to track and analyze the parameters and states (critical process parameters and quality attributes) of biomanufacturing in real time to minimize both capital and operational expenses (Iglesias Jr et al., 2023; Zimmerleiter et al., 2020; Fonseca & Zaiat, 2023). An example of fast and low-cost bioprocess monitoring involves the real-time estimation of nutrients, metabolites, and production formation based on online cell growth measurements. This strategy is low-cost because it requires only one device to measure cell growth (Xv) instead of multiple assays/devices to perform offline and online measurements of all state variables, as in traditional approaches to monitoring bioprocesses. In addition, it is fast because nutrients, metabolites, and production formation are estimated in real-time, while the conventional methods take hours or days (Iglesias Jr et al., 2022; 2023).

Given that, the biomanufacturing conditions for fast and low-cost bioprocess monitoring with a NKE are: 1) A state variables vector defined as $\mathbf{x} = [x_{msv}, x_2, \ldots, x_n]$. 2) x_{msv} as the unique measured state variable (MSV). 3) A UMM represented by a system of nonlinear differential equations of the form: $\mathbf{f}(\mathbf{x}) = [\frac{dx_{msv}}{dt} = f_1(x_{msv}, x_2, \ldots, x_{n-1}, \theta_1, \theta_2, \ldots, \theta_m), \frac{dx_2}{dt} = f_2(x_{msv}, x_2, \ldots, x_{n-1}, \theta_1, \theta_2, \ldots, \theta_m), \cdots, \frac{dx_n}{dt} = f_n(x_{msv}, \theta_{up})]$ where x_{msv} and $x_2, \ldots, x_n$ are the variables of the system, $f_1, f_2, \ldots, f_n$ are the functions defining the system represented by $\mathbf{f}(\cdot)$, and $\theta_1, \theta_2, \ldots, \theta_m$ are the parameters of the system, and θ_{up} an unshared parameter. 4) $\mathbf{R}$ as measurement noise variance of x_{msv}. 5) θ_{up} as the unshared parameter (UP) to be evolved (estimated) and presented in only one weak term. A weak term is a term of an ODE with a low percentage of variables of the state variable vector, and a "strong term" is one with a high percentage of variables of the state variable vector (Iglesias Jr & Bolic, 2024). 6) $\mathbf{P} = Diag([P_{x_{msv},x_{msv}}, P_{x_2,x_2}, \cdots, P_{n,n}, P_{\theta_{up},\theta_{up}}])$ and $\mathbf{Q} = Diag([Q_{x_{msv},x_{msv}}, Q_{x_2,x_2}, \cdots, Q_{n,n}, Q_{\theta_{up},\theta_{up}}])$ with uncorrelated elements due to limited data, meaning they are diagonal with nonzero diagonal elements (noise variances) and zero off-diagonal elements (Ohadi et al., 2015; Paquet-Durand et al., 2020; Yousefi-Darani et al., 2020b; Iglesias Jr et al., 2022).

Due to the limited space, the background information about UMM, NKE, and SANTO can be seen in the Appendix A.

3. Theoretical Analysis

If the process and measurement noises can be assumed to be additive, a state space model can be written as

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1}, \;\; \mathbf{q}_{k-1} \sim N(0, \mathbf{Q}_{k-1}) \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k, \;\; \mathbf{r}_k \sim N(0, \mathbf{R}_k) \end{aligned} \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state, where $\mathbf{y}_k \in \mathbb{R}^m$ is the measurement, $\mathbf{q}_{k-1} \sim N(0, \mathbf{Q}_{k-1})$ is the Gaussian process noise, $\mathbf{r}_k \sim N(0, \mathbf{R}_k)$ is the Gaussian measurement noise, and $\mathbf{h}(\cdot)$ is the measurement model function. Then, using the DEKF (see algorithm 1 Appendix B.5), we can estimate the state and parameters of the state space model (Equation 1). However, in the presence of biomanufacturing conditions for fast and low-cost bioprocess monitoring, we have the issue case described by the following theorem.

Theorem 3.1. *The DEKF cannot estimate an unshared parameter (parameter evolution) that is part of a weak term in a UMM if the unshared parameter is not part of the nonlinear function that models the unique state variable measured.*

Theorem 3.1 negatively answers the **RQ0**, and the proof can be seen in the Appendix C. It shows that the "failure" arises from the inability of the algorithm to derive any information about the unshared parameter from the measured state variable.

4. Empirical Evaluation

Here, the objective is to answer RQ1 and RQ2 by assessing the ability of the NKEs to simultaneously estimate state variables and unshared parameters in a synthetic monoclonal Antibody (mAb) production dataset and evaluating the improvement in performance with the SANTO approach.

[1] https://github.com/cristovaoiglesias/NKEs-SANTO

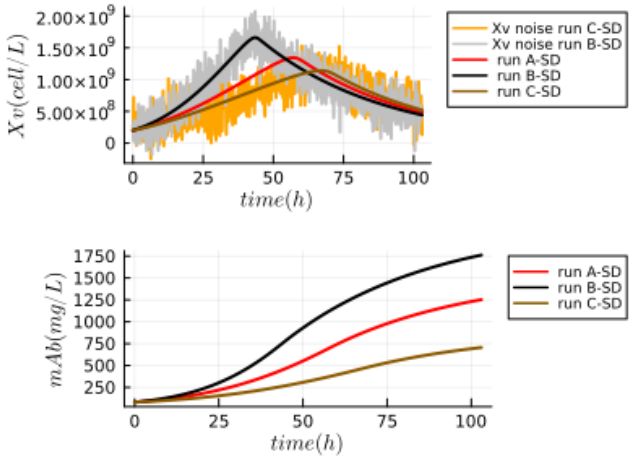

Figure 1. Synthetic dataset regarding the mAb production. The run A-SD (red lines) was generated using the original parameters proposed by (Liu & Gunawan, 2017). Run B-SD (black lines) has the maximum cell expansions and maximum of mAb (titer) production of SD. On the other hand, the run C-SD (brown lines) has the minimum cell expansions and mAb (titer) production. The X_v of B-SD and C-SD with noise is highlighted in grey and in orange in the first plot. These noises are used to evaluate the performance of the NKEs to estimate mAb and QmAb.

4.1. Synthetic Dataset - mAb production

The Synthetic dataset (SD) has data regarding Monoclonal Antibody (mAb) productions that represent the biomanufacturing of a protein widely used as diagnostic reagents and for therapeutic purposes (Jyothilekshmi & Jayaprakash, 2021). The SD is composed of three runs (A-SD, B-SD, and C-SD) with different cell expansions and different maximums of mAb (titer) production. The runs of SD can be seen in Figure 1, and the runs have a sample rate of 7.5 minutes during 103 hours of the process. It is important to point out that C-SD represents a challenging additional scenery to those studied in (Iglesias Jr & Bolic, 2024). The details about the development of the SD can be seen in Appendix E.1.

4.2. NKEs assessment with synthetic dataset to address RQ1 and RQ2

All NKEs (JEKF-Classic, JUKF-Classic, JCKF-Classic, JEKF-SANTO, JUKF-SANTO, and JCKF-SANTO) used the UMM described in Appendix D.1 as process model and the same initial concentration regarding the state variables, (see Table 4 in the Appendix). The NKEs were used to correct (estimate) the predictions regarding state variables (Xv and mAb) and to evolve the unshared parameter (QmAb) of the process model simultaneously. This was done using the Xv samples with the noise of the run B-SD and C-SD as the unique measured state variable and the parameters used to generate the run A-SD as initial parameters of the process model (see Table 3 in Appendix). This situation represents a joint estimation problem where the prediction and parameter of the process model should be corrected by the NKEs based on measured state variable Xv with noise. For example, the initial value used for QmAb is the value of run A-SD (QmAb = 7.21 $\times 10^{-9}$ $mg\ cells^{-1}h^{-1}$), and it should be evolved to the value of run B-SD (9.21 $\times 10^{-9}$ $mg\ cells^{-1}h^{-1}$) based on Xv with the noise of run B-SD. In the case of run C-SD, QmAb should evolve from 7.21 $\times 10^{-9}$ $mg\ cells^{-1}h^{-1}$ to 4.21 $\times 10^{-9}$ $mg\ cells^{-1}h^{-1}$. Furthermore, the Xv (without noise) and mAb samples of run B-SD and C-SD were used as ground truth too. In addition, the root mean square percentage error (RMSPE) was used as a metric to assess the similarity between NKEs estimations and the ground truth of run B-SD and C-SD. It is important to point out that in case of run B-SD, we applied the SANTO approach by adding a small positive quantity to $\mathrm{P}_{\mathrm{X_v,QmAb}} = \mathrm{P}_{\mathrm{QmAb,X_v}}$ of $\mathbf{P}(0)$, and in case of run C-SD we added a small negative quantity. This was done because $\mathrm{P}_{\mathrm{X_v,QmAb}}$ is an off-diagonal element of $\mathbf{P}(0)$, and it can be a positive or negative quantity. The details about the design of NKEs related to $\mathbf{R}$, $\mathbf{Q}$ and $\mathbf{P}(0)$ with SD can be found in Appendix E.2.

Table 1. RMSPE between NKEs estimations about mAb and ground truth of run B-SD and run C-SD

NKE	RMSPE (run B-SD)	RMSPE (run C-SD)
JEKF-SANTO	1.92%	1.83%
JUKF-SANTO	1.75%	1.80%
JCKF-SANTO	1.11%	1.12%
JEKF-Classic	18.6%	48.1%
JUKF-Classic	18.4%	48.1%
JCKF-Classic	18.4%	48.3%

4.3. Results and Discussion

Answer to RQ1. The estimations regarding the states of Xv and mAb, and the unshared parameter (QmAb) done by classic JEKF, JUKF, and JCKF can be seen in Figures 2, and 3. In the case of run B-SD (Figure 2), the classic JEKF was not able to evolve (update) the QmAb, because the estimations were constant and equal to the initial value of 7.21 $\times 10^{-9}$ $mg\ cells^{-1}h^{-1}$ (purple dash line in Plots B and C). On the other hand, the classic JUKF and JCKF evolved QmAb (green and blue dash line in Plot C) but not significantly to arrive close to ground truth (red dash line in Plot B). Consequently, the classic JEKF, JUKF, and JCKF estimation regarding mAb were far from the ground truth of run B-SD (red dash line in plot D), and they had the highest RMSPE values, see Table 1. It is essential to point out that the Kalman gain over time (related to QmAb) that was obtained by JEKF-Classic is constant and equal to zero (purple dash line in Plots E, F, and G) as described in (Iglesias Jr & Bolic, 2024). However, the Kalman gain values obtained by classic JUKF and JCKF presented fluctuation around zero. We created two plots in different scales to visualize the fluctuation of their Kalman gain values (green and blue dash

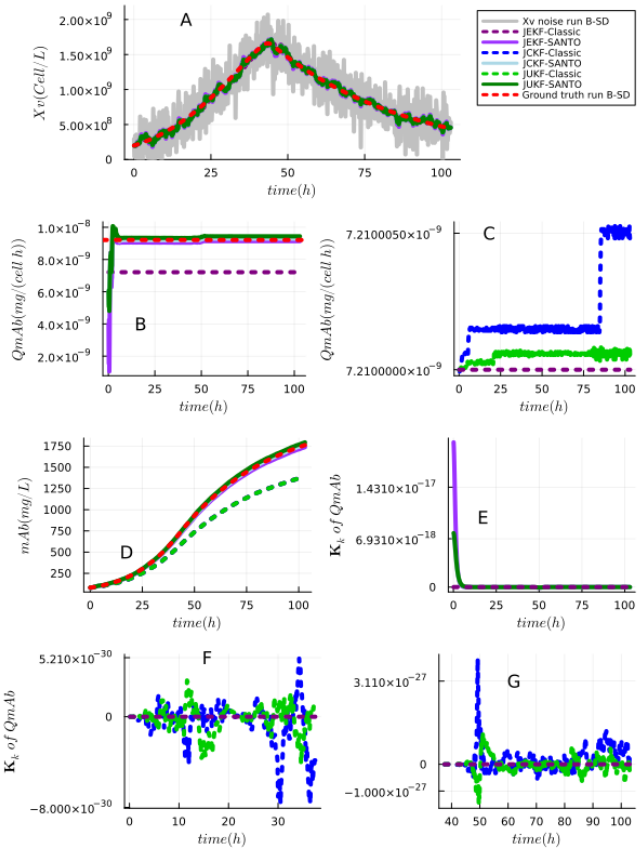

Figure 2. Empirical test with **run B-SD**. Plot A demonstrates noise levels and estimations of Xv, closely aligning with the ground truth. Plot B covers QmAb estimations by JEKF-Classic, JEKF-SANTO, JUKF-SANTO and JCKF-SANTO, while Plot C focuses on the classic versions of JUKF-Classic and JCKF-Classic. Plot D shows the mAb estimation done by all NKEs . Plots E, F and G show the Kalmain Gain over time for the NKEs.

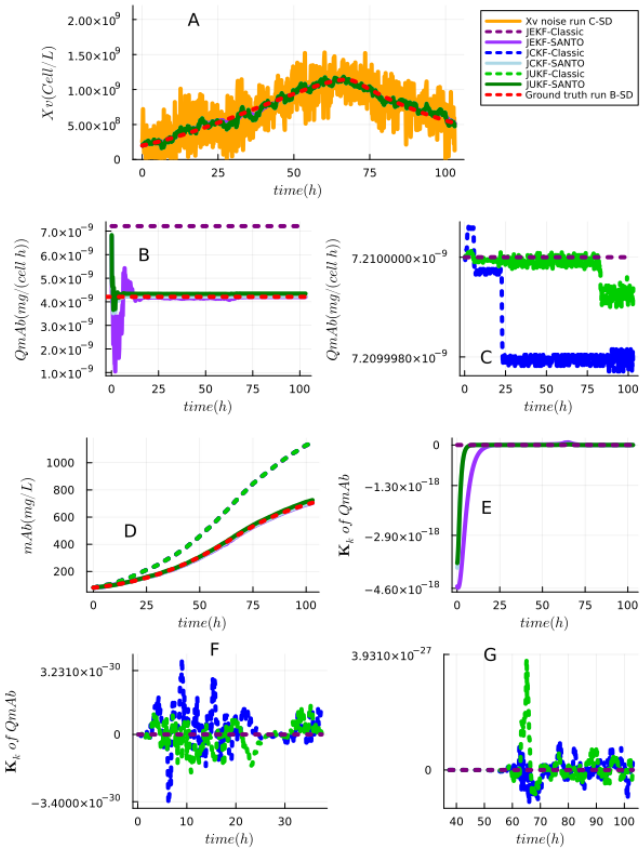

Figure 3. Empirical test with **run C-SD**. Plot A demonstrates noise levels and estimations of Xv, closely aligning with the ground truth. Plot B covers QmAb estimations by JEKF-Classic, JEKF-SANTO, JUKF-SANTO and JCKF-SANTO, while Plot C focuses on the classic versions of JUKF-Classic and JCKF-Classic. Plot D shows the mAb estimation done by all NKEs . Plots E, F and G show the Kalmain Gain over time for the NKEs.

lines in Plots F and G). Furthermore, the obtained results with run C-SD (Figure 3) by classic JEKF, JUKF, and JCKF were similar to the ones obtained with run B-SD (Figure 2). They also could not simultaneously estimate the Xv, mAb, and QmAb (the estimations were far from the red dash line in Plots B and C), and the Kalman gain values also had the same behavior.

Answer to RQ2. The estimations regarding Xv, mAb, and QmAb done by the NKEs (JEKF, JUKF, and JCKF) with the SANTO approach can be seen in Figures 2, and 3. In the case of run B-SD (Figure 2), all NKEs with the SANTO approach evolved the QmAb from the initial value to the ground truth (red dash line in Plot B), and consequently, they estimated the mAb close to the ground truth of run B-SD (red dash line Plot D). This is confirmed by their smallest RMSPE values; see Table 1. Furthermore, the Kalman gain values obtained by all NKEs with SANTO presented a normal and stable behavior starting from a positive value and converging to a small value close to zero, see plot E. However, in the case of run C-SD (Figure 3), despite all NKEs with SANTO estimate simultaneously mAb and QmAb close to the ground truth (Plots B and D), their Kalman gain values presented an unconventional behavior. They started from a negative value and converged to a small value close to zero; see plot E. This is a consequence of adding a small negative quantity to $\mathrm{P}_{\mathrm{Xv,QmAb}}$ of $\mathbf{P}(0)$ that includes a negative component in the Kalman gain computation.

5. Conclusion and Future works

Our analysis and empirical results demonstrated that the DEKF, JUKF, and JCKF are inefficient under the conditions studied. The SANTO approach, however, improved the performance of JUKF and JCKF, achieving lower RMSPE values than the classic versions. Nevertheless, the empirical test using run C-SD showed unconventional Kalman gain values for all NKEs with the SANTO approach, starting from a negative value and converging to near zero. Future work will focus on developing a hybrid NKEs and an autotune approach for all NKE components, including process

model parameters, to address cases like run C-SD.

References

Akca, A. and Efe, M. Ö. Multiple model kalman and particle filters and applications: A survey. *IFAC-PapersOnLine*, 52(3):73–78, 2019.

Arasaratnam, I. and Haykin, S. Cubature kalman filters. *IEEE Transactions on automatic control*, 54(6):1254–1269, 2009.

Aswal, N., Sen, S., and Mevel, L. Switching kalman filter for damage estimation in the presence of sensor faults. *Mechanical Systems and Signal Processing*, 175:109116, 2022.

Biomatik. How much does it cost to make a custom antibody?, 2022. url - https://www.biomatik.com/blog/how-much-does-it-cost-to-make-a-custom-antibody/, Accessed: 2024-05-30.

Capra, E., Godfre, A., Loche, A., and Smith, J. Gene-therapy innovation: Unlocking the promise of viral vectors. *Recuperate by: https://www. mckinsey. com/industries/life-sciences/our-insights/gene-therapy-innovat ion-unlocking-the-promise-of-viral-vectors*, 2021.

Chopda, V., Gyorgypal, A., Yang, O., Singh, R., Ramachandran, R., Zhang, H., Tsilomelekis, G., Chundawat, S. P., and Ierapetritou, M. G. Recent advances in integrated process analytical techniques, modeling, and control strategies to enable continuous biomanufacturing of monoclonal antibodies. *Journal of Chemical Technology & Biotechnology*, 97(9):2317–2335, 2022.

Christakis, I., Tsakiridis, O., Kandris, D., and Stavrakas, I. A kalman filter scheme for the optimization of low-cost gas sensor measurements. *Electronics*, 13(1):25, 2023.

Fonseca, R. F. and Zaiat, M. Development of a low-cost electrochemical sensor for monitoring components in wastewater treatment processes. *Environmental Technology*, 44(25):3883–3896, 2023.

Forum, W. E. The bio revolution: Innovations transforming economies, societies, and our lives, 2020. https://www.mckinsey.com/industries/life-sciences/our-insights/the-bio-revolution-innovations-transforming-economies-societies-and-our-lives.

Haykin, S. S. and Haykin, S. S. *Kalman filtering and neural networks*, volume 284. Wiley Online Library, 2001.

Hernandez, I., Bott, S. W., Patel, A. a., Wolf, C. G., Hospodar, A. R., Sampathkumar, S., and Shrank, W. H. Pricing of monoclonal antibody therapies: higher if used for cancer. *Am J Manag Care*, 24(2):109–112, 2018.

Herwig, C., Pörtner, R., and Möller, J. *Digital Twins: tools and concepts for smart biomanufacturing*. Springer, 2021a.

Herwig, C., Pörtner, R., and Möller, J. *Digital Twins: Applications to the Design and Optimization of Bioprocesses*. Springer, 2021b.

Hetzler, Z., Lott, N., Poonam, A. D., Dalgan, S., and Wei, Q. Single-use biosensors for biomanufacturing: Perspective on the state-of-the-art. *Current Opinion in Biomedical Engineering*, pp. 100512, 2023.

Iglesias, Cristovão Freitas, J. X. X. V. M. M. A. A. V.-S. N. B. A. K. and Bolic, M. Monitoring the recombinant adeno-associated virus production using extended kalman filter. *Processes*, 10(11):2180, 2022.

Iglesias Jr, C. F. and Bolic, M. How not to make the joint extended kalman filter fail with unstructured mechanistic models. *Sensors*, 24(2):653, 2024.

Iglesias Jr, C. F., Xu, X., Mehta, V., Akassou, M., Venereo-Sanchez, A., Belacel, N., Kamen, A., and Bolic, M. Monitoring the recombinant adeno-associated virus production using extended kalman filter. *Processes*, 10(11):2180, 2022.

Iglesias Jr, C. F., Ristovski, M., Bolic, M., and Cuperlovic-Culf, M. raav manufacturing: The challenges of soft sensing during upstream processing. *Bioengineering*, 10 (2):229, 2023.

Ji, Z. and Brown, M. Joint state and parameter estimation for biochemical dynamic pathways with iterative extended kalman filter: comparison with dual state and parameter estimation. *The Open Automation and Control Systems Journal*, 2(1), 2009.

Julier, S. J. and Uhlmann, J. K. New extension of the kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI*, volume 3068, pp. 182–193. Spie, 1997.

Jyothilekshmi, I. and Jayaprakash, N. Trends in monoclonal antibody production using various bioreactor systems. 2021.

Khodarahmi, M. and Maihami, V. A review on kalman filter models. *Archives of Computational Methods in Engineering*, pp. 1–21, 2022.

Khuat, T. T., Bassett, R., Otte, E., Grevis-James, A., and Gabrys, B. Applications of machine learning in antibody discovery, process development, manufacturing and formulation: Current trends, challenges, and opportunities. *Computers & Chemical Engineering*, pp. 108585, 2024.

Kopp, R. E. and Orford, R. J. Linear regression applied to system identification for adaptive control systems. *Aiaa Journal*, 1(10):2300–2306, 1963.

Kyriakopoulos, S., Ang, K. S., Lakshmanan, M., Huang, Z., Yoon, S., Gunawan, R., and Lee, D.-Y. Kinetic modeling of mammalian cell culture bioprocessing: the quest to advance biomanufacturing. *Biotechnology Journal*, 13 (3):1700229, 2018.

Li, A. H., Wu, P., and Kennedy, M. Replay overshooting: Learning stochastic latent dynamics with the extended kalman filter. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 852–858. IEEE, 2021.

Liu, Y. and Gunawan, R. Bioprocess optimization under uncertainty using ensemble modeling. *Journal of biotechnology*, 244:34–44, 2017.

Ljung, L. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1):36–50, 1979.

Luo, Y., Kurian, V., and Ogunnaike, B. A. Bioprocess systems analysis, modeling, estimation, and control. *Current Opinion in Chemical Engineering*, 33:100705, 2021.

Mandenius, C.-F. and Gustavsson, R. Mini-review: Soft sensors as means for pat in the manufacture of biotherapeutics. *Journal of Chemical Technology & Biotechnology*, 90(2):215–227, 2015.

Mears, L., Stocks, S. M., Albaek, M. O., Sin, G., and Gernaey, K. V. Mechanistic fermentation models for process design, monitoring, and control. *Trends in biotechnology*, 35(10):914–924, 2017.

Moser, A., Appl, C., Brüning, S., and Hass, V. C. Mechanistic mathematical models as a basis for digital twins. In *Digital Twins*, pp. 133–180. Springer, 2020.

Murugan, C. Soft sensors for biomass monitoring during low cost cellulase production. In *Biomass*. IntechOpen, 2021.

Narayanan, H., Behle, L., Luna, M. F., Sokolov, M., Guillén-Gosálbez, G., Morbidelli, M., and Butté, A. Hybrid-ekf: Hybrid model coupled with extended kalman filter for real-time monitoring and control of mammalian cell culture. *Biotechnology and Bioengineering*, 117(9):2703–2714, 2020.

Narayanan, H., von Stosch, M., Feidl, F., Sokolov, M., Morbidelli, M., and Butté, A. Hybrid modeling for biopharmaceutical processes: advantages, opportunities, and implementation. *Frontiers in Chemical Engineering*, 5: 1157889, 2023.

Nelson, L. and Stear, E. The simultaneous on-line estimation of parameters and states in linear systems. *IEEE Transactions on automatic Control*, 21(1):94–98, 1976.

Niazi, S. and Lokesh, S. Understanding bioprocessing. In *Biopharmaceutical Manufacturing, Volume 2*, 2053-2563, pp. 1–1 to 1–42. IOP Publishing, 2022. ISBN 978-0-7503-3179-1. doi: 10.1088/978-0-7503-3179-1ch1. URL https://dx.doi.org/10.1088/978-0-7503-3179-1ch1.

Ohadi, K., Legge, R. L., and Budman, H. M. Development of a soft-sensor based on multi-wavelength fluorescence spectroscopy and a dynamic metabolic model for monitoring mammalian cell cultures. *Biotechnology and bioengineering*, 112(1):197–208, 2015.

Papathanasiou, M. M., Burnak, B., Katz, J., Shah, N., and Pistikopoulos, E. N. Assisting continuous biomanufacturing through advanced control in downstream purification. *Computers & Chemical Engineering*, 125:232–248, 2019.

Paquet-Durand, O., Zettel, V., Yousefi-Darani, A., and Hitzmann, B. The supervision of dough fermentation using image analysis complemented by a continuous discrete extended kalman filter. *Processes*, 8(12):1669, 2020.

Park, S.-Y., Park, C.-H., Choi, D.-H., Hong, J. K., and Lee, D.-Y. Bioprocess digital twins of mammalian cell culture for advanced biomanufacturing. *Current Opinion in Chemical Engineering*, 33:100702, 2021.

Rathore, A. S., Nikita, S., Thakur, G., and Mishra, S. Artificial intelligence and machine learning applications in biopharmaceutical manufacturing. *Trends in Biotechnology*, 41(4):497–510, 2023.

Research, G. V. Biotechnology market size, share & trends analysis report by application (health, agriculture, food, natural resources & environment, industrial processing bioinformatics), by technology, and segment forecasts, 2020 - 2027, 2020. URL https://www.grandviewresearch.com/industry-analysis/biotechnology-market.

Reyes, S. J., Durocher, Y., Pham, P. L., and Henry, O. Modern sensor tools and techniques for monitoring, controlling, and improving cell culture processes. *Processes*, 10 (2):189, 2022.

Rocha, K. D. and Terra, M. H. Robust kalman filter for systems subject to parametric uncertainties. *Systems & Control Letters*, 157:105034, 2021.

Särkkä, S. and Svensson, L. *Bayesian filtering and smoothing*, volume 17. Cambridge university press, 2023.

Sinner, P., Daume, S., Herwig, C., and Kager, J. *Usage of Digital Twins Along a Typical Process Development Cycle*, pp. 71–96. Springer International Publishing, Cham, 2021. ISBN 978-3-030-71660-8. doi: 10.1007/10_2020_149. URL https://doi.org/10.1007/10_2020_149.

Sokolov, M., von Stosch, M., Narayanan, H., Feidl, F., and Butté, A. Hybrid modeling—a key enabler towards realizing digital twins in biopharma? *Current Opinion in Chemical Engineering*, 34:100715, 2021.

SU Support. Costs of monoclonal antibody production, 2024. URL https://www.susupport.com/knowledge/monoclonal-antibodies/costs-monoclonal-antibody-production. Accessed: 2024-05-30.

Tsopanoglou, A. and del Val, I. J. Moving towards an era of hybrid modelling: advantages and challenges of coupling mechanistic and data-driven models for upstream pharmaceutical bioprocesses. *Current Opinion in Chemical Engineering*, 32:100691, 2021.

Tulsyan, A., Schorner, G., Khodabandehlou, H., Wang, T., Coufal, M., and Undey, C. A machine-learning approach to calibrate generic raman models for real-time monitoring of cell culture processes. *Biotechnology and Bioengineering*, 116(10):2575–2586, 2019.

Tulsyan, A., Khodabandehlou, H., Wang, T., Schorner, G., Coufal, M., and Undey, C. Spectroscopic models for real-time monitoring of cell culture processes using spatiotemporal just-in-time gaussian processes. *AIChE Journal*, 67 (5):e17210, 2021.

Wan, E. A. and Van Der Merwe, R. The unscented kalman filter for nonlinear estimation. In *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, pp. 153–158. Ieee, 2000.

Yousefi-Darani, A., Paquet-Durand, O., and Hitzmann, B. The Kalman Filter for the Supervision of Cultivation Processes. *Advances in biochemical engineering/biotechnology*, 177:95–125, 2020a. ISSN 07246145. doi: 10.1007/10_2020_145. URL https://link.springer.com/chapter/10.1007/10_2020_145.

Yousefi-Darani, A., Paquet-Durand, O., and Hitzmann, B. The kalman filter for the supervision of cultivation processes. *Digital Twins*, pp. 95–125, 2020b.

Yousefi-Darani, A., Paquet-Durand, O., Hinrichs, J., and Hitzmann, B. Parameter and state estimation of backers yeast cultivation with a gas sensor array and unscented kalman filter. *Engineering in Life Sciences*, 21(3-4):170–180, 2021a.

Yousefi-Darani, A., Paquet-Durand, O., and Hitzmann, B. *The Kalman Filter for the Supervision of Cultivation Processes*, pp. 95–125. Springer International Publishing, Cham, 2021b. ISBN 978-3-030-71656-1. doi: 10.1007/10_2020_145. URL https://doi.org/10.1007/10_2020_145.

Zhang, D., Del Rio-Chanona, E. A., Petsagkourakis, P., and Wagner, J. Hybrid physics-based and data-driven modeling for bioprocess online simulation and optimization. *Biotechnology and bioengineering*, 116(11):2919–2930, 2019.

Zhao, B., Li, X., Sun, W., Qian, J., Liu, J., Gao, M., Guan, X., Ma, Z., and Li, J. Biodt: An integrated digital-twin-based framework for intelligent biomanufacturing. *Processes*, 11(4):1213, 2023.

Zimmerleiter, R., Kager, J., Nikzad-Langerodi, R., Berezhinskiy, V., Westad, F., Herwig, C., and Brandstetter, M. Probeless non-invasive near-infrared spectroscopic bioprocess monitoring using microspectrometer technology. *Analytical and Bioanalytical Chemistry*, 412(9):2103–2109, 2020.

Author Contributions

Conceptualization, C.F.I.J.; Implemented algorithms and conducted the experiments, C.F.I.J. and L.P; Performed analysis on experimental results and wrote the manuscript, C.F.I.J.; Provided insightful discussions, reviewed the results and revised the manuscript, C.M. and M.B.; Supervision, M.B.; Project administration, M.B.; funding acquisition, M.B.

A. Background Extension

B. Motivation

Biomanufacturing has significant economic importance (Forum, 2020). The estimated value of the worldwide biotechnology market stood at USD 1.37 trillion in 2022. It is projected to expand at a compound annual growth rate (CAGR) of 13.96% from 2023 through 2030 (Research, 2020). To achieve this target, it is necessary to employ more cost-effective solutions and smart manufacturing while simultaneously not sacrificing the process robustness and product quality when aiming for higher process productivity, more affordable end-products, and shorter production times (Khuat et al., 2024). Therefore, there is constant pressure for bioprocess monitoring strategies that are fast (efficient) and low-cost (cost-effective) (Niazi & Lokesh, 2022; Khuat et al., 2024; Hetzler et al., 2023), particularly for the upstream stages of viral vectors and monoclonal antibodies productions (Khuat et al., 2024). Some reasons for this include the following.

First, bioprocess monitoring in areas like cell and gene therapy relies more heavily on offline measurements (Hetzler et al., 2023; Reyes et al., 2022). Commercial manufacturing still heavily relies on time-consuming offline analytics and manual control strategies (Khuat et al., 2024). For example, nutrients (glucose, glutamine) and metabolites (lactate, ammonium) can be measured using inline near-infrared (NIR) sensors. However, due to potential cross-interference with other molecular species, the current industrial standard method of measuring glucose, glutamine, lactate, and ammonium is still based on offline measurement (Reyes et al., 2022; Hetzler et al., 2023). Online sensing involves high-frequency measurements from the sensors, and offline sensing involves collecting samples from the biomanufacturing process and analyzing them in a separate laboratory setting. It often takes several hours or days for the result to arrive (Chopda et al., 2022; Iglesias & Bolic, 2022; Papathanasiou et al., 2019). Thus, offline sensing typically leads to a lower level of process control and a higher demand for labor (Hetzler et al., 2023). Typically, these offline techniques, which involve extracting cells from the bioreactor, are often time-intensive, labor-intensive, and lead to waste due to the use of expensive and potentially harmful reagents and samples (Tulsyan et al., 2019; Khuat et al., 2024). Furthermore, offline measurements are performed infrequently (e.g., every 12–24 hours), resulting in low-resolution process monitoring and the risk of missing metabolic changes in cells that could indicate process alterations or operational problems, making control strategies based on offline analytical measurements often ineffective (Tulsyan et al., 2021; Khuat et al., 2024). Additionally, each sample taken from the bioreactor carries a risk of contamination or batch loss (Tulsyan et al., 2019). Second, factors such as labor costs, raw materials, equipment, facility expenses, the type and quantity of in-process control analyses, and overall output all influence the bioprocess economy (Niazi & Lokesh, 2022). Additionally, there are often overlooked costs associated with lifecycle management, a lack of process agility when deviations occur, and regulatory compliance updates. For instance, the production cost of monoclonal antibodies varies significantly due to the factors presented previously, and it is expensive. An annual mAb therapy in oncology can cost around $100,000. Producing one mAb can range from $6,000 to $15,000 (Biomatik, 2022; Hernandez et al., 2018; SU Support, 2024). Another example is the production of recombinant Adeno-Associated Virus. Currently, a standard production run of a recombinant Adeno-Associated Virus vector treatment costs around $100,000 per dose (considering approximately 1×10^{17} vg per batch) (Capra et al., 2021). Therefore, any technological innovation that reduces the cost per dose would be immediately beneficial (Capra et al., 2021; Niazi & Lokesh, 2022).

Fast (efficient) and low-cost (cost-effective) bioprocess monitoring can be defined as a set of methods designed to track and analyze the parameters and states (critical process parameters and quality attributes) of biomanufacturing in real time to minimize both capital and operational expenses (Iglesias Jr et al., 2022; 2023; Niazi & Lokesh, 2022). An example illustrating an ”extreme case” of fast and low-cost bioprocess monitoring involves the real-time estimation of nutrients (glucose, glutamine), metabolites (lactate, ammonium), and production formation (titer) based on online viable cell density (Xv) measurements and initial conditions (Yousefi-Darani et al., 2020a; 2021a). This strategy is low-cost because it requires only one device to measure Xv instead of multiple assays/devices to perform offline and online measurements of all state variables, as in traditional approaches for bioprocess monitoring. In addition, it is fast because nutrients, metabolites, and production formation are estimated in real-time, while the offline methods take hours or days to deliver results, as discussed before (Iglesias Jr et al., 2022; 2023; Niazi & Lokesh, 2022). It is important to note that the information about product formation, such as titer, is typically available only after the end of the process, such as viral vectors and monoclonal antibody productions, which takes several days through offline measurements.

The literature indicates that soft sensors based on nonlinear Kalman estimators (NKE) with unstructured mechanistic models (UMM) can enable fast and low-cost bioprocess monitoring and achieve the goals of several initiatives, such as Quality by Design, Process Analytical tools, and Pharma 4.0 (Reyes et al., 2022; Christakis et al., 2023; Mandenius & Gustavsson, 2015; Murugan, 2021; Sokolov et al., 2021). Soft sensors based on NKE with UMM can enable real-time monitoring of CPP or

CQA that are difficult to measure directly or that can only be measured at low sampling frequencies in a bioprocess (Sinner et al., 2021; Yousefi-Darani et al., 2020a). UMMs allow modeling the macro-scale of the bioprocess phenomenon, and they can be divided into two groups. Specific UMMs provide a detailed and customized description of a particular bioprocess, while Generic UMMs offer a more generalized framework that can be adapted to various bioprocesses with the help of additional computational tools. NKE are nonlinear state estimators based on the Kalman Filter (KF) framework. NKEs use a two-step recursive algorithmic process: prediction and update (Khodarahmi & Maihami, 2022; Särkkä & Svensson, 2023). NKEs are widely applicable in various fields today, and the most popular NKEs are (Akca & Efe, 2019): Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), and Cubature Kalman Filter (CKF). The most used NKE in bioprocess monitoring are EKF and UKF (Yousefi-Darani et al., 2020a). It is essential to note that other nonlinear state estimators exist, such as particle filtering (PF). However, due to the simple structure (practical implementation) and low computational effort of NKEs, these methods have gained more interest. Many research studies have been dedicated to the implementation of such filters for state and parameter estimation in bioprocess technologies based on micro-controllers which requires numerically economical and robust algorithms, such as NKEs (Zhao et al., 2023; Reyes et al., 2022; Yousefi-Darani et al., 2020a; Herwig et al., 2021b;a; Aswal et al., 2022; Nelson & Stear, 1976). The optimal performance of NKEs relies on the proper tuning of five components: i) the process noise covariance matrix ($\mathbf{Q}$), ii) the measurement noise covariance matrix ($\mathbf{R}$), iii) the initial state noise covariance matrix ($\mathbf{P}(0)$), iv) initial condition of state variables ($\mathbf{x}_0$) and v) parameters ($\boldsymbol{\theta}$) of the UMM (Li et al., 2021; Khodarahmi & Maihami, 2022; Särkkä & Svensson, 2023). In general, to enhance the performance of NKE as soft sensors, they are performed with joint and dual strategies to estimate the states and parameters of a UMM. The joint NKE (JNKE) considers a single joint state variable vector, which includes both the states and parameters $\boldsymbol{\theta}$ of the UMM (Ljung, 1979; Kopp & Orford, 1963; Haykin & Haykin, 2001), and the dual NKE (DNKE) employs two consecutive NKE, separating the estimation of the system states and parameters (Ji & Brown, 2009). JNKE and DNKE are motivated by the need to correct the prediction of a UMM regarding state variables and to update the UMM by evolving its parameters based on the corrections made (Haykin & Haykin, 2001). However, although JNKE and DNKE have been used in many different bioprocess monitoring applications (Yousefi-Darani et al., 2020a; 2021a; Paquet-Durand et al., 2020; Ji & Brown, 2009; Herwig et al., 2021a;b; Iglesias Jr et al., 2022; Iglesias & Bolic, 2022), the classic NKEs with UMM are limited to fast and low-cost bioprocess monitoring, and improvements are needed to handle different biomanufacturing conditions.

B.1. Biopharma is a data-limited industry

Obtaining data in biopharmaceutical studies is costly and time-consuming due to the necessity of conducting biological experiments, which require significant time and resources. Additionally, these data are often linked to strategic products and confidential information of pharmaceutical companies, making it challenging to publicly share them with the research community. This lack of publicly available benchmark datasets hinders the reproducibility of machine learning (ML) solutions for bioprocess data, essential for assessing effectiveness and performing comparative analyses between different algorithms. The bioprocess data are expensive to obtain, requiring time-consuming, labor-intensive, and costly biological experiments (Khuat et al., 2024). Furthermore, these data are associated with strategic products and business secrets, making them unlikely to be publicly shared, preventing the explosive growth of ML solutions in biopharmaceuticals compared to fields like natural language processing and computer vision. Artificial intelligence and machine learning (AI/ML) techniques are suitable for learning patterns in data from high-dimensional design spaces with complex non-linear interactions (Narayanan et al., 2023; Khuat et al., 2024). However, the effective use of AI/ML requires high-quality data in significant quantities to develop useful and applicable models for biopharmaceutical applications. The biopharma industry is particularly data-limited regarding actively generated data, as each experiment and analysis is resource-intensive, limiting the number of experiments and corresponding data that can be generated (Narayanan et al., 2023). While the industry possesses substantial ”historical data,” these were not collected for AI/ML training purposes and contain biases in the design spaces explored and information recorded, affecting overall data quality. However, it is important to point out that hybrid dynamic models and NKE with UMM can overcome this need for large amount of data and enable the development of soft sensor for fast and low-cost bioprocess monitoring.

B.2. Unstructured Mechanistic Model (UMM)

UMMs also known as Unstructured Mechanistic Kinetic Models, are pivotal in modeling the temporal progression of bioprocesses like the production of therapeutic monoclonal antibodies (mAbs), projected to generate USD 300 billion by 2025, and rAAV production, a leading viral vector technology for gene therapy (Kyriakopoulos et al., 2018; Luo et al., 2021; Iglesias Jr et al., 2022; 2023). These models, grounded in fundamental principles, are key to understanding and simulating bioprocess dynamics at the macro-scale, such as cell density, viability, and nutrient/metabolite concentrations. Despite their

Table 2. Glossary

JEKF	Joint estimation of states and parameters with Extended Kalman Filter
JUKF	Joint estimation of states and parameters with Unscented Kalman Filter
JCKF	Joint estimation of states and parameters with Cubature Kalman Filter
NKE	Nonlinear Kalman Estimator
UMM	Unstructured Mechanistic Model
MRDE	Matrix Ricatti Differential Equation
MSV	Measured State Variable
UP	Unshared parameter
SANTO	Specific initiAl coNdiTiOn
CD-EKF	Continuous-Discrete EKF
SD	Synthetic dataset
mAb	Monoclonal Antibody

critical role in digital twin (DT) development and soft sensors, the industrial application of UMMs is still nascent (Moser et al., 2020; Park et al., 2021; Mears et al., 2017; Reyes et al., 2022). In contrast to Structured Mechanistic Models (SMMs), which delve into the intracellular details of a homogeneous cell population and are more complex, requiring extensive expertise for development, UMMs are less detailed but more practical for dynamic control in common biomanufacturing bioreactors (Luo et al., 2021; Tsopanoglou & del Val, 2021). SMMs are better suited for cell-line development, focusing on genomic-level alterations for desired process behaviors. However, the predictive capability of simple UMMs is limited, often failing to accurately estimate process states across different operating conditions (Zhang et al., 2019). To enhance their predictive accuracy, UMMs are frequently integrated with the Kalman filter and its nonlinear variants like the extended Kalman filter, effectively predicting unobserved states.

B.3. Nonlinear Kalman Estimators

The two-step recursive algorithmic process used by Nonlinear Kalman Estimators (EKF, UKF, and CKF) is summarized as follows (Akca & Efe, 2019; Khodarahmi & Maihami, 2022):

- **Prediction step**: This step is where the state and error are propagated forward in time. In this step the predicted mean of state, $\hat{\mathbf{x}}_{k/k-1}$, and predicted error covariance matrix of state $\mathbf{P}_{k/k-1}$ are obtained using a process model (nonlinear system dynamics), a initial condition ($\hat{\mathbf{x}}_0$ and $\mathbf{P}_0$) and $\mathbf{Q}$. See the Figure 4. Here is a description of the prediction step performed by each of the Nonlinear Kalman Estimators:
  - EKF: Linearizes the system's dynamics around the current state estimate to predict the next state (Julier & Uhlmann, 1997).
  - RKF: Similar to EKF but incorporates mechanisms to handle model uncertainties and outliers. It adapts to model uncertainties and outliers, enhancing resilience to deviations from nominal assumptions (Rocha & Terra, 2021).
  - UKF: Uses a set of deterministically chosen sample points (sigma points) to capture the mean and covariance of the state estimate and propagates these through the nonlinear system dynamics. UKF uses a deterministic sampling to approximate nonlinear transformations without linearization (Wan & Van Der Merwe, 2000).
  - CKF: Employs cubature rules to compute the integral of the state transition function over the state distribution, effectively predicting the next state without linearization (Arasaratnam & Haykin, 2009).
- **Update step**: This step is the same for all Nonlinear Kalman Estimators (EKF, UKF, RKF, and CKF). In this step the Predictions ($\hat{\mathbf{x}}_{k/k-1}$ and $\mathbf{P}_{k/k-1}$) are combined with the measured values ($\mathbf{y}_k$) to provide updated states and errors ($\hat{\mathbf{x}}_{k/k}$ and $\mathbf{P}_{k/k}$) using $\mathbf{R}$. This involves calculating the Kalman gain, determining how much the state prediction should be corrected based on the new measurement, and updating the error covariance to reflect the reduced uncertainty after incorporating the measurement. In addition, the updated state becomes the initial condition for the next prediction performed by the prediction step, $\hat{\mathbf{x}}_0 = \hat{\mathbf{x}}_{k/k}$ and $\mathbf{P}_0 = \mathbf{P}_{k/k}$. See the Figure 4.

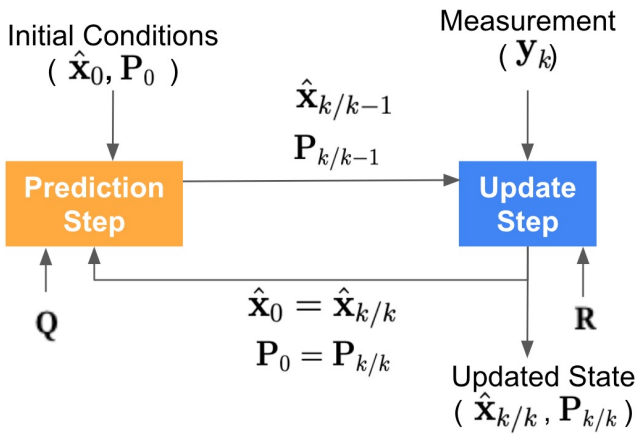

Figure 4. Two-step recursive algorithmic process used by Nonlinear Kalman Estimators such as EKF, UKF, and CKF. The state predictions are updated with new measurements, and this updated state becomes the basis for the next prediction.

B.4. JEKF, JUKF, and JCKF Overview

B.4.1. JEKF (Joint Extended Kalman Filter)

JEKF is a Bayesian filter-based approach for joint estimation in nonlinear dynamical systems. It concatenates states x_i and parameters θ of a process model into a single joint state vector. The state variables vector $\psi(t)$ in JEKF is extended as:

$$\psi(t) = [x_1, x_2, ..., x_n, \theta_1, ..., \theta_n]^T. \tag{2}$$

In JEKF, the learning involves both states x_i and parameters θ_i of a discrete-time nonlinear system (e.g., UMM). It corrects system states and model parameters simultaneously based on observed noisy signals Z_k. JEKF is recognized for parameter evolution, where parameters are treated as random variables with noise added at each timestep:

$$\theta(t_k) = \theta(t_{k-1}) + \text{noise}, \tag{3}$$

This approach is efficient for updating process model parameters, especially when near optimal parameters for specific conditions. In JEKF, parameter estimation refers to this ongoing evolution of parameters.

B.4.2. JUKF (Joint Unscented Kalman Filter)

JUKF extends the UKF to joint estimation scenarios. Unlike JEKF, JUKF does not linearize the process and measurement models but instead uses a deterministic sampling technique (the unscented transform) to capture the mean and covariance estimates. This makes JUKF more accurate in capturing the true state of a nonlinear system. JUKF also concatenates states and parameters into a single state vector and simultaneously estimates them using the unscented transform and Kalman filter equations.

B.4.3. JCKF (Joint Cubature Kalman Filter)

Similar to JUKF, JCKF is designed for joint estimation of states and parameters in nonlinear systems. JCKF employs the cubature Kalman filter, which uses cubature rules to approximate the integrals in the state and covariance propagation. This approach avoids linearization errors and is computationally more efficient than JUKF. JCKF, like the other joint estimation methods, concatenates states and parameters into a single state vector for simultaneous estimation.

Each of these filters - JEKF, JUKF, and JCKF - has unique characteristics and is suitable for different nonlinear estimation scenarios in processes such as biomanufacturing.

B.5. DEKF algorithm

This Dual EKF algorithm is a robust method for simultaneous state and parameter estimation in systems where model dynamics are influenced by unknown or time-varying parameters. It iteratively predicts and corrects estimates based on new measurements, adapting over time to improve accuracy. This approach is widely used in control systems, robotics, and areas requiring real-time estimation under uncertain conditions. Here is a detailed breakdown and description of each step within the DEKF Algorithm1:

Initialization: 1 - State Initialization: The initial state estimate, denoted as $\mathbf{x}_{0|0}$, is set to the expected value $E(\mathbf{x}_0)$. 2 - State Covariance Initialization: The initial covariance matrix $\mathbf{P}_{\mathbf{x}_{0|0}}$ of the state estimate is computed. It represents the expected error in the initial state estimate. 3 - Parameter Covariance Initialization: Similarly, $\mathbf{P}_{\boldsymbol{\theta}_0}$ is initialized, representing the covariance of the initial parameter estimates, indicating the uncertainty or variability in these initial parameter estimates.

Prediction Step for Parameters: 1 - The parameters $\boldsymbol{\theta}_{k|k-1}$ at step k given information up to step $k-1$ are assumed to be equal to the estimates from the previous step, reflecting a model where parameters do not evolve over time. 2 - The covariance of the parameter estimates, $\mathbf{P}^{\boldsymbol{\theta}}_{k|k-1}$, is updated by scaling the previous covariance $\mathbf{P}^{\boldsymbol{\theta}}_{k-1|k-1}$ by a factor λ^{-}, which adjusts the estimate's uncertainty.

Prediction Step for States: 1 - The state prediction $\mathbf{x}_{k|k-1}$ is calculated using the state transition function $\mathbf{f}$, which predicts the next state based on the previous state and current parameter estimates. 2 - The covariance of the state prediction, $\mathbf{P}^{\mathbf{x}}_{k|k-1}$, is updated to reflect the uncertainty due to the process noise and the propagation of the previous state's uncertainty.

Update Step for States: 1 - The Kalman Gain for the states, $\mathbf{K}^{\mathbf{x}}_k$, is computed, which determines how much the measurements taken at time k, denoted by ϵ_k, should impact the state estimate. 2 - The state estimate $\mathbf{x}_{k|k}$ is updated using the Kalman Gain and the measurement residual ϵ_k. 3 - The covariance of the updated state estimate, $\mathbf{P}^{\mathbf{x}}_{k|k}$, is calculated, indicating reduced uncertainty after incorporating the measurement.

Update Step for Parameters: 1 - Similarly to the states, a Kalman Gain for the parameters, $\mathbf{K}^{\boldsymbol{\theta}}_k$, is computed. 2 - The parameter estimates $\boldsymbol{\theta}_{k|k}$ are updated using this Kalman Gain and the measurement residual. 3 - The covariance of the updated parameter estimates, $\mathbf{P}^{\boldsymbol{\theta}}_{k|k}$, is updated to reflect the new level of uncertainty post-measurement.

Additional Mathematical Details: i) ϵ_k represents the innovation or measurement residual, which is the difference between the actual measurement and the predicted measurement. ii) Various Jacobian matrices ($\mathbf{F}_k$, $\mathbf{H}_k$, $\mathbf{H}^{\boldsymbol{\theta}}_k$) are used to linearize the nonlinear state transition and measurement functions around the current estimates, which is a key aspect of the EKF methodology.

Algorithm 1 DEKF Algorithm

1)Initialize
$\mathbf{x}_{0|0} = E(\mathbf{x}_0)$
$\mathbf{P}_{\mathbf{x}_{0|0}} = E[(\mathbf{x}_0 - \mathbf{x}_{0|0})(\mathbf{x}_0 - \mathbf{x}_{0|0})^T]$
$\mathbf{P}_{\boldsymbol{\theta}_0} = E[(\boldsymbol{\theta}_0 - \boldsymbol{\theta}_{0|0})(\boldsymbol{\theta}_0 - \boldsymbol{\theta}_{0|0})^T]$
2) Prediction step of parameters
$\boldsymbol{\theta}_{k|k-1} = \boldsymbol{\theta}_{k-1|k-1}$
$\mathbf{P}^{\boldsymbol{\theta}}_{k|k-1} = \lambda^{-}\mathbf{P}^{\boldsymbol{\theta}}_{k-1|k-1}$
3) Prediction step of states
$\mathbf{x}_{k|k-1} = \mathbf{f}(\mathbf{x}_{k-1|k-1}, \boldsymbol{\theta}_{k|k-1})$
$\mathbf{P}^{\mathbf{x}}_{k|k-1} = \mathbf{F}_k\mathbf{P}^{\mathbf{x}}_{k-1|k-1} + \mathbf{P}^{\mathbf{x}}_{k-1|k-1}\mathbf{F}_k^T + \mathbf{Q}$
4) Update step of states
$\mathbf{K}^{\mathbf{x}}_k = \mathbf{P}^{\mathbf{x}}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}^{\mathbf{x}}_{k|k-1}\mathbf{H}_k^T + \mathbf{R})^{-}$
$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}^{\mathbf{x}}_k\epsilon_k$
$\mathbf{P}^{\mathbf{x}}_{k|k} = (\mathbf{I} - \mathbf{K}^{\mathbf{x}}_k\mathbf{H}_k)\mathbf{P}^{\mathbf{x}}_{k|k-1}$
5) Update step of parameters
$\mathbf{K}^{\boldsymbol{\theta}}_k = \mathbf{P}^{\boldsymbol{\theta}}_{k|k-1}\mathbf{H}_k^{\boldsymbol{\theta}T}(\mathbf{H}^{\boldsymbol{\theta}}_k\mathbf{P}^{\boldsymbol{\theta}}_{k|k-1}\mathbf{H}_k^{\boldsymbol{\theta}T} + \mathbf{R}^{\boldsymbol{\theta}})^{-}$
$\boldsymbol{\theta}_{k|k} = \boldsymbol{\theta}_{k|k-1} + \mathbf{K}^{\boldsymbol{\theta}}_k\epsilon_k$
$\mathbf{P}^{\boldsymbol{\theta}}_{k|k} = (\mathbf{I} - \mathbf{K}^{\boldsymbol{\theta}}_k\mathbf{H}^{\boldsymbol{\theta}}_k)\mathbf{P}^{\boldsymbol{\theta}}_{k|k-1}$
where $\epsilon_k = \mathbf{y}_k - \mathbf{h}(\mathbf{x}_{k|k-1}) = \mathbf{y}_k - \bar{\mathbf{H}}\mathbf{x}_{k|k-1}$;
$\bar{\mathbf{H}} = [1\ 0\ ...\ 0]$
$\mathbf{F}_k = \frac{\partial \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}_{k|k-1})}{\partial \mathbf{x}} \mid_{\mathbf{x}_{k-1|k-1}}$
$\mathbf{H}_k = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} \mid_{\mathbf{x}_{k|k-1}} = \bar{\mathbf{H}}$
$\mathbf{H}^{\boldsymbol{\theta}}_k = -\frac{\partial \epsilon_k}{\partial \boldsymbol{\theta}} = \frac{\partial \mathbf{h}(\mathbf{x}_{k|k-1})}{\partial \boldsymbol{\theta}} = \bar{\mathbf{H}}\frac{\partial \mathbf{x}_{k|k-1}}{\partial \boldsymbol{\theta}} \mid_{\boldsymbol{\theta}_{k|k-1}}$

B.6. Failure Case: Biomanufacturing conditions

The following conditions are prevalent in biomanufacturing and should be taken into consideration while developing Joint or Dual NKE applications for this area:

- **ODEs of UMM with Unshared Parameters:** Parameters unique to one term of an ODE and not shared with other ODEs in the UMM are typical for modeling product formation dynamics in biomanufacturing.
- **P and Q with Uncorrelated Elements:** Often, limited data leads to assuming error covariance matrices **P** (process error covariance) and **Q** (measurement error covariance) with uncorrelated elements, meaning they are diagonal with nonzero diagonal elements (noise variances) and zero off-diagonal elements. This results in two scenarios: 1) *Using **P** with uncorrelated elements for MRDE construction and as **P**(t=0) initial condition:* Here, the MRDE ODEs depend only on noise variances of $\mathrm{P}_{i,i}$ and $\mathrm{Q}_{i,i}$, and elements of Jacobian $\mathbf{J}_t^{\phi}$. 2) *Using **P** with correlated elements for MRDE construction and **P**(t=0) with uncorrelated elements as initial condition:* This can reduce time-invariant ODEs predicting state error covariance between two state variables.
- **ODEs of UMM with Weak Terms:** Weak terms in an ODE have less impact on the predicted state error covariance $\mathbf{P}(t_{k|k-1})$ compared to strong terms. The Jacobian $\mathbf{J}_t^{\phi}$ is more influenced by strong terms.
- **ODEs of UMM with Weak Variables:** Weak variables, present only in the first member of an ODE, do not contribute to the computation of predicted error covariance $\mathbf{P}(t_{k|k-1})$, as their first-order partial derivatives in Jacobian $\mathbf{J}_t^{\phi}$ are zero. In contrast, strong variables significantly influence $\mathbf{P}(t_{k|k-1})$ computation.
- **Only One Measured State Variable:** In some JEKF applications, only a single state variable is measured. This variable determines the column of predicted state error covariance $\mathbf{P}(t_{k|k-1})$ used for Kalman gain computation. If a row in this column is zero (no covariance between the measured and state variable), the Kalman gain for the state variable represented by that row cannot be computed.

These conditions emphasize the complexities and limitations in applying classical JEKF in biomanufacturing, where unique parameter characteristics and measurement constraints can impact its effectiveness (Iglesias Jr & Bolic, 2024).

B.7. SANTO approach

The SANTO approach is designed to address the failure case of the classical Joint Extended Kalman Filter (JEKF) in biomanufacturing scenarios. This approach specifically targets the initial condition of the Matrix Riccati Differential Equation (MRDE), which is the initial state error covariance matrix $\mathbf{P}_0 = \mathbf{P}(t = 0)$. In typical situations where $\mathbf{P}_0$ is composed of uncorrelated elements ($\mathrm{P}_{i,j} = 0$), certain initial conditions of time-invariant ODEs in the MRDE are zero, leading to zero solutions for these ODEs from t_{k-1} to t_k. In the presence of biomanufacturing conditions, the Kalman gain for the unshared parameter (K_{UP}) and the predicted state error covariance between the measured state variable and the unshared parameter ($P_{MSV,UP}(t_{k|k-1})$) are also zero, resulting in $K_{UP} = 0$ and $P_{MSV,UP}(t_{k|k-1}) = 0$. This indicates an unrealistic scenario where the prediction regarding the unshared parameter is considered perfect, without the need for measurement influence in the JEKF correction step.

To avoid this failure, the SANTO approach modifies the initial condition of MRDE. Instead of considering all off-diagonal elements of $\mathbf{P}(t = 0)$ as zero, a key off-diagonal element, specifically $P_{MSV,UP}(t = 0)$, is assigned an initial value different from zero ($P_{MSV,UP}(t = 0) \neq 0$). This value can be a positive or negative quantity, λ, reflecting the covariance between two variables. The choice of λ is crucial: it must be small enough to not significantly affect the filter's estimates but large enough to prevent the failure case.

Theorem of SANTO: The introduction of a positive quantity λ to the $P_{MSV,UP}(t = 0)$ in $\mathbf{P}(t = 0)$, initializing the MRDE with a specific initial condition, can prevent the Kalman gain from being constantly zero throughout the JEKF execution, thereby averting the failure case (Iglesias Jr & Bolic, 2024).

C. Proof of Theorem 3.1

Theorem 3.1 *The Dual estimation of states and parameters with EKF cannot estimate an unshared parameter (parameter evolution) that is part of a weak term in a UMM if the unshared parameter is not part of the nonlinear function that models the unique state variable measured.*

The proof of Theorem 3.1 is in the following.

Proof. Let's consider the following:

- A general UMM with an unshared parameter in an weak term represented by a system of nonlinear differential equations of the form:

$$\begin{aligned} \frac{dx_{msv}}{dt} &= f_1(x_{msv}, x_2, \ldots, x_{n-1}, \theta_1, \theta_2, \ldots, \theta_m) \\ \frac{dx_2}{dt} &= f_2(x_{msv}, x_2, \ldots, x_{n-1}, \theta_1, \theta_2, \ldots, \theta_m) \\ &\vdots \\ \frac{dx_n}{dt} &= f_n(x_{msv}, \theta_{up}) \end{aligned} \tag{4}$$

  where x_{msv} and $x_2, \ldots, x_n$ are the variables of the system, $f_1, f_2, \ldots, f_n$ are the functions defining the system represented by $\mathbf{f}(\cdot)$, and $\theta_1, \theta_2, \ldots, \theta_m$ are the parameters of the system and θ_{up} an unshared parameter.

- A state variables vector defined as

$$\mathbf{x} = [x_{msv}, x_2, \ldots, x_n]^T. \tag{5}$$

- x_{msv} as the unique measured state variable (MSV) and $\bar{\mathbf{H}}$ = [1 0 ... 0 0].

- $\mathbf{R}$ as measurement noise variance of x_{msv}.
- $\boldsymbol{\theta} = [\theta_{up}]$, where θ_{up} as the unshared parameter (UP) to be evolved (estimated) and presented in only one weak term.
- Euler–Maruyama discretization of general UMM (Equations 4)

$$\mathbf{x}_{k|k-1} = \begin{pmatrix} x_{msv,k|k-1} \\ x_{2,k|k-1} \\ \vdots \\ x_{n,k|k-1} \end{pmatrix} = \begin{pmatrix} x_{msv,k-1|k-1} + f_1(x_{msv,k-1|k-1}, x_{2,k-1|k-1}, \ldots, x_{n-1,k-1|k-1}, \theta_1, \theta_2, \ldots, \theta_m)\Delta t \\ x_{2,k-1|k-1} + f_2(x_{msv,k-1|k-1}, x_{2,k-1|k-1}, \ldots, x_{n-1,k-1|k-1}, \theta_1, \theta_2, \ldots, \theta_m)\Delta t \\ \vdots \\ x_{n,k-1|k-1} + f_n(x_{msv,k-1|k-1}, \theta_{up,k-1|k-1})\Delta t \end{pmatrix} \tag{6}$$

Given these conditions above and the DEKF Algorithm 1, we have that

$$\begin{aligned} \mathbf{H}_k^{\boldsymbol{\theta}} &= -\frac{\partial \epsilon_k}{\partial \boldsymbol{\theta}} = \frac{\partial \mathbf{h}(\mathbf{x}_{k|k-1})}{\partial \boldsymbol{\theta}} = \bar{\mathbf{H}} \frac{\partial \mathbf{x}_{k|k-1}}{\partial \boldsymbol{\theta}} \mid_{\boldsymbol{\theta}_{k|k-1}} \\ \mathbf{H}_k^{\boldsymbol{\theta}} &= [1\ 0\ \ldots\ 0] \begin{bmatrix} \frac{d(x_{msv,k-1|k-1} + f_1(x_{msv,k-1|k-1}, x_{2,k-1|k-1}, \ldots, x_{n-1,k-1|k-1}, \theta_1, \theta_2, \ldots, \theta_m)\Delta t)}{d\theta_{up}} = 0 \\ \frac{d(x_{2,k-1|k-1} + f_2(x_{msv,k-1|k-1}, x_{2,k-1|k-1}, \ldots, x_{n-1,k-1|k-1}, \theta_1, \theta_2, \ldots, \theta_m)\Delta t)}{d\theta_{up}} = 0 \\ \vdots \\ \frac{d(x_{n,k-1|k-1} + f_n(x_{msv,k-1|k-1}, \theta_{up,k-1|k-1})\Delta t)}{d\theta_{up}} = \Delta t \frac{d(f_n(x_{msv,k-1|k-1}, \theta_{up,k-1|k-1}))}{d\theta_{up}} \end{bmatrix}^T \\ \mathbf{H}_k^{\boldsymbol{\theta}} &= [1\ 0\ \ldots\ 0] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \Delta t \frac{d(f_n(x_{msv,k-1|k-1}, \theta_{up,k-1|k-1}))}{d\theta_{up}} \end{bmatrix}^T = 0. \end{aligned} \tag{7}$$

Therefore,

$$\begin{aligned} \mathbf{K}_k^{\boldsymbol{\theta}} &= \mathbf{P}_{k|k-1}^{\boldsymbol{\theta}} \mathbf{H}_k^{\boldsymbol{\theta}T} (\mathbf{H}_k^{\boldsymbol{\theta}} \mathbf{P}_{k|k-1}^{\boldsymbol{\theta}} \mathbf{H}_k^{\boldsymbol{\theta}T} + \mathbf{R}^{\boldsymbol{\theta}})^{-} \\ \mathbf{K}_k^{\boldsymbol{\theta}} &= 0 \end{aligned} \tag{8}$$

Consequently, there is no update because $\boldsymbol{\theta}_{k|k} = \boldsymbol{\theta}_{k|k-1}$,

$$\begin{aligned} \boldsymbol{\theta}_{k|k} &= \boldsymbol{\theta}_{k|k-1} + \mathbf{K}_k^{\boldsymbol{\theta}} \epsilon_k, \\ \boldsymbol{\theta}_{k|k} &= \boldsymbol{\theta}_{k|k-1} + 0, \\ \boldsymbol{\theta}_{k|k} &= \boldsymbol{\theta}_{k|k-1}. \end{aligned} \tag{9}$$

□

The proof for Theorem 3.1 explains why the Dual Extended Kalman Filter (DEKF) cannot estimate an unshared parameter in a weak term within an unstructured mechanistic model (UMM), particularly if the unshared parameter does not affect the unique state variable measured. It shows that the "failure" arises from the inability of the algorithm to derive any information about the unshared parameter from the measured state variable. The Jacobian ($\mathbf{H}^{\boldsymbol{\theta}}$) is derived from the relationship between the measurement model and the parameter. Since the measurement model only references the MSV, the partial derivative

with respect to θ_{up} is zero. Consequently, the Kalman gain ($\mathbf{K}^{\boldsymbol{\theta}}$) for this parameter is also zero because it is calculated using the covariance and the Jacobian. Given that the Kalman gain is zero, updating the parameter estimate becomes impossible, leading to no change in the estimate. The parameter estimate remains the same, indicating that the DEKF cannot correct or evolve the unshared parameter due to its lack of influence on the MSV.

D. UMM for Empirical Evaluation

D.1. UMM for monoclonal antibody (mAb) production

The ODE system 10 is a UMM used for Mab production (Liu & Gunawan, 2017). This system represents the cell growth, uptake of substrates, metabolism, and production process with 16 parameters described in the Table 3. It is important to point out that $QmAb$ denotes the specific mAb production rate, and is an example of unshared parameter. More details can be found in (Liu & Gunawan, 2017).

$$
\begin{aligned}
\frac{d\,X_V}{dt} &= (\mu - \mu_d)X_V \\
\frac{d\,X_t}{dt} &= \mu X_V - k_{lysis}(X_t - X_V) \\
\mu &= \mu_{max}.\frac{[GLC]}{K_{glc}+[GLC]}.\frac{[GLN]}{K_{gln}+[GLN]}.\frac{K_{Ilac}}{K_{Ilac}+[LAC]}.\frac{K_{Iamm}}{K_{Iamm}+[AMM]} \\
\mu_d &= \frac{\mu_{d,max}}{1+(K_{d,amm}+[AMM])^2} \\
\frac{d\,[GLC]}{dt} &= -Q_{glc}Xv \\
\frac{d\,[GLN]}{dt} &= -Q_{gln}Xv - K_{d,gln}[GLN] \\
\frac{d\,[LAC]}{dt} &= Q_{lac}Xv \\
\frac{d\,[AMM]}{dt} &= Q_{amm}Xv + K_{d,gln}[GLN] \\
Q_{glc}Xv &= \frac{\mu}{Y_{x,glc}} + m_{glc} \\
Q_{gln}Xv &= \frac{\mu}{Y_{x,gln}} + m_{gln} = \frac{\mu}{Y_{x,gln}} + \frac{\alpha_2[GLN]}{\alpha_2+[GLN]} \\
Q_{lac}Xv &= Y_{lac,glc}Q_{glc} \\
Q_{amm}Xv &= Y_{amm,gln}Q_{gln} \\
\frac{d\,[mAb]}{dt} &= (2-\gamma\mu)Q_{mAb}.X_V
\end{aligned}
\tag{10}
$$

Let's break down the components of this ODE system:

1. **Cell Growth and Death Dynamics:**

   • $\frac{dX_V}{dt} = (\mu - \mu_d)X_V$: This equation models the rate of change of viable cell density (X_V) over time. The growth rate (μ) minus the death rate (μ_d) is multiplied by the current viable cell density.

   • $\frac{dX_t}{dt} = \mu X_V - k_{lysis}(X_t - X_V)$: This equation describes the total cell density (X_t), considering both viable and non-viable cells. The rate of total cell density change is determined by the growth of viable cells and the lysis (breakdown) of cells, where k_{lysis} is the lysis rate constant.

2. Growth Rate (μ) and Death Rate (μ_d):

   • μ: Defined as a function of substrate concentrations ([GLC] for glucose and [GLN] for glutamine) and inhibitors ([LAC] for lactate and [AMM] for ammonium). This function reflects how cell growth rate is influenced by the availability of nutrients and the presence of metabolic byproducts.

• μ_d: The death rate, modeled as a function of the ammonium concentration, with $\mu_{d,max}$ representing the maximum death rate and $K_{d,amm}$ as a constant.

3. Substrate Consumption and Metabolite Production:

   • The following set of equations ($\frac{d[GLC]}{dt}$, $\frac{d[GLN]}{dt}$, $\frac{d[LAC]}{dt}$, $\frac{d[AMM]}{dt}$) represent the rates of change in concentrations of glucose, glutamine, lactate, and ammonium, respectively. These are key substrates and metabolites in the cell culture. The terms $Q_{glc}, Q_{gln}, Q_{lac}, Q_{amm}$ denote specific consumption/production rates of these components, and $K_{d,gln}$ is the degradation constant for glutamine.

4. Balancing Equations for Substrate Consumption and Product Formation:

   • The equations relating $Q_{glc}Xv, Q_{gln}Xv, Q_{lac}Xv, Q_{amm}Xv$ establish relationships between growth rate, substrate consumption, and metabolite production rates. These are based on yield coefficients ($Y_{x,glc}, Y_{x,gln}, Y_{lac,glc}, Y_{amm,gln}$) and maintenance coefficients ($m_{glc}, m_{gln}, \alpha_2$).

5. Monoclonal Antibody (mAb) Production:

   • $\frac{d[mAb]}{dt} = (2 - \gamma\mu)Q_{mAb}.X_V$: This equation models the rate of mAb production. The specific mAb production rate (Q_{mAb}) is multiplied by the viable cell density and a factor considering the growth rate, where γ is a constant.

The model's strength lies in its ability to capture the interplay between cell growth, nutrient consumption, metabolite accumulation, and product formation, which are crucial for optimizing and monitoring biomanufacturing processes. The parameter Q_{mAb}, representing the specific mAb production rate, is particularly notable as it's an unshared parameter, meaning its value is unique to this process and not shared with other models or components within this system.

E. Empirical Evaluation - Extension

E.1. Synthetic dataset development - mAb production

The Synthetic dataset (SD) is composed of three runs (A-SD, B-SD, and C-SD). The runs have different samples regarding the state variables Xv, GLC, GLN, LAC, AMM, and mAb and were generated using the UMM D.1 with three set of different parameters (Table 3), and the same initial condition (Table 4). The runs were generated using the UMM proposed by (Liu & Gunawan, 2017) (see Section D.1) with small variations in the parameters μ_{max} (Maximum growth rate) and QmAb (mAb specific production rate) (see Table 3), but with the same initial concentrations of states variables (viable cell density (Xv), glucose (GLC), glutamine (GLN), lactate (LAC), ammonium (AMM) and mAb), and with different conditions of pH and Temperature as done in the synthetic dataset of (Narayanan et al., 2020). The run A-SD (red lines in plots of Figure 1) was generated using the original parameters proposed by (Liu & Gunawan, 2017), that are the parameters μ_{max} = $5.8 \times 10^{-9}(h^-)$ and QmAb = 7.21 ($\times 10^{-9}\ mg\ cells^{-1}h^{-1}$). Run B-SD (black lines in plots of Figure 1) has the maximum cell expansions and maximum of mAb (titer) production of SD, and they were obtained with the parameters μ_{max} = $7.5 \times 10^{-9}(h^-)$ and QmAb = 9.21 ($\times 10^{-9}\ mg\ cells^{-1}h^{-1}$). On the other hand, the run C-SD (brown lines in plots of Figure 1) has the minimum cell expansions and minimum mAb (titer) production, and they were obtained with the parameters μ_{max} = $5 \times 10^{-9}(h^-)$ and QmAb = 4.21 ($\times 10^{-9}\ mg\ cells^{-1}h^{-1}$). Furthermore, the runs B-SD and C-SD have samples regarding X_V (cell/L) with Gaussian white noise, and they were created by adding the Gaussian white noise with standard deviation of 20×10^7 to the data represented in blue and green lines. The X_v of B-SD and C-SD with noise is highlighted in light grey and in orange in the first plot. It is important to point out that X_V samples with Gaussian white noise represent a possible online measurement with sensor including noises. These noises are used to evaluate the performance of the NKEs to estimate mAb and QmAb.

Table 3. Initial parameters used in UMM case D.1 to generate the runs A-SD, B-SD and C-SD of Synthetic Dataset (SD).

Parameter	**Name**	**run A-SD**	**run B-SD**	**run C-SD**
$\mu_{max}(h^-)$	Maximum growth rate	5.8×10^{-2}	7.5×10^{-2}	5×10^{-2}
$k_{glc}(mM)$	Monod constant glucose	7.5×10^{-1}	7.5×10^{-1}	7.5×10^{-1}
$k_{gln}(mM)$	Monod constant glutamine	7.5×10^{-2}	7.5×10^{-2}	7.5×10^{-2}
$k_{Ilac}(mM)$	Monod constant lactate for inhibition	1.72×10^{2}	1.72×10^{2}	1.72×10^{2}
$k_{Iamm}(mM)$	Monod constant ammonium for inhibition	2.85×10^{1}	2.85×10^{1}	2.85×10^{1}
$\mu_{d,max}(h^-)$	Maximum death rate	3.0×10^{-2}	3.0×10^{-2}	3.0×10^{-2}
$K_{d,amm}(mM)$	Monod constant ammonium for death	1.76	1.76	1.76
$K_{lysis}(h^-)$	Breakdown of cell membranes	5.51×10^{-2}	5.51×10^{-2}	5.51×10^{-2}
$Y_{X,glc}(cells\ mmol^-)$	Yield coefficient cell conc./glucose	1.06×10^{8}	1.06×10^{8}	1.06×10^{8}
$m_{glc}(mmol/cells\ h)$	Glucose maintenance coefficient	4.85×10^{-14}	4.85×10^{-14}	4.85×10^{-14}
$Y_{X,gln}(cells/mmmol)$	Yield coefficient cell conc./glutamine	5.57×10^{8}	5.57×10^{8}	5.57×10^{8}
$\alpha_1(mmols\ cells^-\ h^-)$	Coefficient for m_{gln}	3.40×10^{-13}	3.40×10^{-13}	3.40×10^{-13}
$\alpha_2(mM)$	Coefficient for m_{gln}	4.0	4.0	4.0
$k_{d,gln}(h^-)$	Monod constant glutamine for death	9.6×10^{-3}	9.6×10^{-3}	9.6×10^{-3}
$Y_{lac/glc}(1)$	Yield coefficient lactate/glucose	1.4	1.4	1.4
$Y_{amm/gln}(1)$	Yield coefficient ammonium/glutamine	4.27×10^{-1}	4.27×10^{-1}	4.27×10^{-1}
γ	constant parameter	4.27×10^{-1}	4.27×10^{-1}	4.27×10^{-1}
$Q_{mAb}(mg\ cells^-\ h^-)$	mAb specific production rate	7.21×10^{-9}	9.21×10^{-9}	4.21×10^{-9}

Table 4. Initial conditions of state variables of UMM case D.1.

State Variable	**Name**	**Value**
Xv	Viable cells density	2×10^8 c/mL
Xt	total cells density	2×10^8 c/mL
GLC	Glucose	29.1 mM
GLN	Glutamine	4.9 mM
LAC	Lactate	0 mM
AMM	Ammonium	0.31 mM
mAb	Monoclonal Antibody (titer)	80.6 mg/L
QmAb	Specific production rate of mAb	$7.21 \times 10^{-9}\ mg\ cells^{-1}h^{-1}$

E.2. NKEs design to address RQ1 and RQ2

The process model (based on UMM case D.1) and joint state variable vector used by NKEs (JEKF-Classic, JUKF-Classic, JCKF-Classic, JEKF-SANTO, JUKF-SANTO and JCKF-SANTO) are the following:

$$\boldsymbol{\psi}(t)_{case4} = [X_V, X_t, GLC, GLN, LAC, AMM, mAb, QmAb]^T, \tag{11}$$

and

$$\frac{d}{dt}\begin{bmatrix} X_V \\ X_t \\ GLC \\ GLN \\ LAC \\ AMM \\ mAb \\ QmAb \end{bmatrix} = \begin{bmatrix} f_{X_V} \\ f_{GLC} \\ f_{GLN} \\ f_{LAC} \\ f_{AMM} \\ f_{mAb} \\ 0 \end{bmatrix} + \boldsymbol{\omega}(t). \tag{12}$$

The **P**(t=0) that were used by the NKEs with run B of Synthetic Dataset are in Tables 5, 6, and 7. Furthermore, the **P**(t=0) that were used by the NKEs with run C of Synthetic Dataset are in Tables 10, 11, and 12. It is important to point out that in case of run B-SD, we applied the SANTO approach by adding a small positive quantity to $\mathrm{P_{X_v,QmAb}}$ and $\mathrm{P_{QmAb,X_v}}$, and in case of run C-SD we added a small negative quantity to $\mathrm{P_{X_v,QmAb}}$ and $\mathrm{P_{QmAb,X_v}}$, see Tables 5, 6, 7, 10, 11, and 12.

The **R** and **Q** used by the NKEs (for runs B of Synthetic Dataset) are presented in Tables 9 and 8. Furthermore, The **R** and **Q** used by the NKEs (for runs C of Synthetic Dataset) are presented in Tables 14 and 13.

It is important point out that all NKEs used a **R**, **P**(t=0), and **Q** that were obtained by by trial and error until achieve positive results in the Normalized Innovations Squared Chi-square Test.

Table 5. Standard initial state error covariance matrix (standard **P**(t=0)) for JEKF-Classic, and JEKF-SANTO with run B of Synthetic Dataset.

Parameter	**Name**	$\mathbf{P}_{i,i}$ for JEKF-Classic	$\mathbf{P}_{i,i}$ for JEKF-SANTO
$\mathrm{P_{X_v,X_v}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00	0.00
$\mathrm{P_{X_t,X_t}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00	0.00
$\mathrm{P_{GLC,GLC}}$ $(\mathrm{mM^2})$	Glucose	0.00	0.00
$\mathrm{P_{GLN,GLN}}$ $(\mathrm{mM^2})$	Glutamine	0.00	0.00
$\mathrm{P_{LAC,LAC}}$ $(\mathrm{mM^2})$	Lactate	0.00	0.00
$\mathrm{P_{AMM,AMM}}$ $(\mathrm{mM^2})$	Ammonium	0.00	0.00
$\mathrm{P_{mAb,mAb}}$ $(\mathrm{mg/L})^2$	Monoclonal Antibody (titer)	0.00	0.00
$\mathrm{P_{QmAb,QmAb}}$ $(\mathrm{g\ cells^{-1}h^{-1}})^2$	Specific production rate of mAb	3.9e-18	3.9e-18
$\mathrm{P_{X_v,QmAb}}$ $(\mathrm{c^2/mL^2})(\mathrm{g\ cells^{-1}h^{-1}})$	Initial $Cov(X_v, QmAb)$	0.0	0.8404

Table 6. Standard initial state error covariance matrix (standard **P**(t=0)) for JUKF-Classic, and JUKF-SANTO with run B of Synthetic Dataset.

Parameter	**Name**	$\mathbf{P}_{i,i}$ for JCKF-Classic	$\mathbf{P}_{i,i}$ for JCKF-SANTO
$\mathrm{P_{X_v,X_v}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00001	0.00001
$\mathrm{P_{X_t,X_t}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00001	0.00001
$\mathrm{P_{GLC,GLC}}$ $(\mathrm{mM^2})$	Glucose	0.00001	0.00001
$\mathrm{P_{GLN,GLN}}$ $(\mathrm{mM^2})$	Glutamine	0.00001	0.00001
$\mathrm{P_{LAC,LAC}}$ $(\mathrm{mM^2})$	Lactate	0.00001	0.00001
$\mathrm{P_{AMM,AMM}}$ $(\mathrm{mM^2})$	Ammonium	0.00001	0.00001
$\mathrm{P_{mAb,mAb}}$ $(\mathrm{mg/L})^2$	Monoclonal Antibody (titer)	0.00001	0.00001
$\mathrm{P_{QmAb,QmAb}}$ $(\mathrm{g\ cells^{-1}h^{-1}})^2$	Specific production rate of mAb	0.01	10000.01
$\mathrm{P_{X_v,QmAb}}$ $(\mathrm{c^2/mL^2})(\mathrm{g\ cells^{-1}h^{-1}})$	Initial $Cov(X_v, QmAb)$	0.0	3.1e-1

Table 7. Standard initial state error covariance matrix (standard **P**(t=0)) for JCKF-Classic, and JCKF-SANTO with run B of Synthetic Dataset.

Parameter	**Name**	$\mathbf{P}_{i,i}$ for JCKF-Classic	$\mathbf{P}_{i,i}$ for JCKF-SANTO
$\mathrm{P_{X_v,X_v}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00001	0.00001
$\mathrm{P_{X_t,X_t}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00001	0.00001
$\mathrm{P_{GLC,GLC}}$ $(\mathrm{mM^2})$	Glucose	0.00001	0.00001
$\mathrm{P_{GLN,GLN}}$ $(\mathrm{mM^2})$	Glutamine	0.00001	0.00001
$\mathrm{P_{LAC,LAC}}$ $(\mathrm{mM^2})$	Lactate	0.00001	0.00001
$\mathrm{P_{AMM,AMM}}$ $(\mathrm{mM^2})$	Ammonium	0.00001	0.00001
$\mathrm{P_{mAb,mAb}}$ $(\mathrm{mg/L})^2$	Monoclonal Antibody (titer)	0.00001	0.00001
$\mathrm{P_{QmAb,QmAb}}$ $(\mathrm{g\ cells^{-1}h^{-1}})^2$	Specific production rate of mAb	0.01	10000.01
$\mathrm{P_{X_v,QmAb}}$ $(\mathrm{c^2/mL^2})(\mathrm{g\ cells^{-1}h^{-1}})$	Initial $Cov(X_v, QmAb)$	0.0	2.9e-1

Table 8. Measurement noise variance **R** and error covariance matrix of process model (**Q**) for the JEKF-Classic, JUKF-Classic and JCKF-Classic with run B of Synthetic Dataset.

Parameter	**Name**	JEKF-Classic	JUKF-Classic	JCKF-Classic
$\mathrm{R^2}$ $(\mathrm{c^2/mL^2})$	Viable cells MNV [1]	$(20\times 10^7)^2$	$(20\times 10^7)^2$	$(20\times 10^7)^2$
$\mathrm{Q_{X_v,X_v}}$ $(\mathrm{c^2/mL^2})$	Viable cells PNV [2]	$(20\times 10^6)^2$	$(20\times 10^6)^2$	$(20\times 10^6)^2$
$\mathrm{Q_{X_t,X_t}}$ $(\mathrm{c^2/mL^2})$	Viable cells PNV [2]	0.001	0.00001	0.00001
$\mathrm{Q_{GLC,GLC}}$ $(\mathrm{mM^2})$	Glucose PNV	0.001	0.00001	0.00001
$\mathrm{Q_{GLN,GLN}}$ $\mathrm{mM^2}$	Glutamine PNV	0.001	0.00001	0.00001
$\mathrm{Q_{LAC,LAC}}$ $(\mathrm{mM^2})$	Lactate PNV	0.001	0.00001	0.00001
$\mathrm{Q_{AMM,AMM}}$ $(\mathrm{mM^2})$	Ammonium PNV	0.001	0.00001	0.00001
$\mathrm{Q_{mAb,mAb}}$ $(\mathrm{VG^2/mL^2})$	Monoclonal Antibody (titer) PNV	0.001	00001	0.001
$\mathrm{Q_{QmAb,QmAb}}$ $(\mathrm{h^{-2}})$	Specific production rate of mAb	1×10^{-18}	0.01	0.1

[1] MNV—measurement noise value; [2] PNV—process noise value.

Table 9. Measurement noise variance **R** and error covariance matrix of process model (**Q**) for the JEKF-SANTO, JUKF-SANTO and JCKF-SANTO with run B of Synthetic Dataset.

Parameter	**Name**	JEKF-SANTO	JUKF-SANTO	JCKF-SANTO
$\mathrm{R}^2\ (\mathrm{c}^2/\mathrm{mL}^2)$	Viable cells MNV [1]	$(20\times 10^7)^2$	$(20\times 10^7)^2$	$(20\times 10^7)^2$
$\mathrm{Q}_{\mathrm{X_v,X_v}}\ (\mathrm{c}^2/\mathrm{mL}^2)$	Viable cells PNV [2]	$(80\times 10^6)^2$	$(20\times 10^6)^2$	$(20\times 10^6)^2$
$\mathrm{Q}_{\mathrm{X_t,X_t}}\ (\mathrm{c}^2/\mathrm{mL}^2)$	Viable cells PNV [2]	0.001	0.00001	0.00001
$\mathrm{Q}_{\mathrm{GLC,GLC}}\ (\mathrm{mM}^2)$	Glucose PNV	0.001	0.00001	0.00001
$\mathrm{Q}_{\mathrm{GLN,GLN}}\ \mathrm{mM}^2$	Glutamine PNV	0.001	0.00001	0.00001
$\mathrm{Q}_{\mathrm{LAC,LAC}}\ (\mathrm{mM}^2)$	Lactate PNV	0.001	0.00001	0.00001
$\mathrm{Q}_{\mathrm{AMM,AMM}}\ (\mathrm{mM}^2)$	Ammonium PNV	0.001	0.00001	0.00001
$\mathrm{Q}_{\mathrm{mAb,mAb}}\ (\mathrm{VG}^2/\mathrm{mL}^2)$	Monoclonal Antibody (titer) PNV	0.001	0.001	0.001
$\mathrm{Q}_{\mathrm{QmAb,QmAb}}\ (\mathrm{h}^{-2})$	Specific production rate of mAb	0.001	0.001	0.001

[1] MNV—measurement noise value; [2] PNV—process noise value.

Table 10. Standard initial state error covariance matrix (standard **P**(t=0)) for JEKF-Classic, and JEKF-SANTO with run C of Synthetic Dataset.

Parameter	**Name**	$\mathbf{P}_{i,i}$ for JEKF-Classic	$\mathbf{P}_{i,i}$ for JEKF-SANTO
$\mathrm{P}_{\mathrm{X_v,X_v}}\ (\mathrm{c}^2/\mathrm{mL}^2)$	Viable cells	0.00	0.00
$\mathrm{P}_{\mathrm{X_t,X_t}}\ (\mathrm{c}^2/\mathrm{mL}^2)$	Viable cells	0.00	0.00
$\mathrm{P}_{\mathrm{GLC,GLC}}\ (\mathrm{mM}^2)$	Glucose	0.00	0.00
$\mathrm{P}_{\mathrm{GLN,GLN}}\ (\mathrm{mM}^2)$	Glutamine	0.00	0.00
$\mathrm{P}_{\mathrm{LAC,LAC}}\ (\mathrm{mM}^2)$	Lactate	0.00	0.00
$\mathrm{P}_{\mathrm{AMM,AMM}}\ (\mathrm{mM}^2)$	Ammonium	0.00	0.00
$\mathrm{P}_{\mathrm{mAb,mAb}}\ (\mathrm{mg/L})^2$	Monoclonal Antibody (titer)	0.00	0.00
$\mathrm{P}_{\mathrm{QmAb,QmAb}}\ (\mathrm{g\ cells}^{-1}\mathrm{h}^{-1})^2$	Specific production rate of mAb	8.9e-18	8.9e-18
$\mathrm{P}_{\mathrm{X_v,QmAb}}\ (\mathrm{c}^2/\mathrm{mL}^2)(\mathrm{g\ cells}^{-1}\mathrm{h}^{-1})$	Initial $Cov(X_v, QmAb)$	0.0	-0.1805

Table 11. Standard initial state error covariance matrix (standard **P**(t=0)) for JUKF-Classic, and JUKF-SANTO with run C of Synthetic Dataset.

Parameter	**Name**	$\mathbf{P}_{i,i}$ for JUKF-Classic	$\mathbf{P}_{i,i}$ for JUKF-SANTO
$\mathrm{P}_{\mathrm{X_v,X_v}}\ (\mathrm{c}^2/\mathrm{mL}^2)$	Viable cells	0.00001	0.00001
$\mathrm{P}_{\mathrm{X_t,X_t}}\ (\mathrm{c}^2/\mathrm{mL}^2)$	Viable cells	0.00001	0.00001
$\mathrm{P}_{\mathrm{GLC,GLC}}\ (\mathrm{mM}^2)$	Glucose	0.00001	0.00001
$\mathrm{P}_{\mathrm{GLN,GLN}}\ (\mathrm{mM}^2)$	Glutamine	0.00001	0.00001
$\mathrm{P}_{\mathrm{LAC,LAC}}\ (\mathrm{mM}^2)$	Lactate	0.00001	0.00001
$\mathrm{P}_{\mathrm{AMM,AMM}}\ (\mathrm{mM}^2)$	Ammonium	0.00001	0.00001
$\mathrm{P}_{\mathrm{mAb,mAb}}\ (\mathrm{mg/L})^2$	Monoclonal Antibody (titer)	0.00001	0.00001
$\mathrm{P}_{\mathrm{QmAb,QmAb}}\ (\mathrm{g\ cells}^{-1}\mathrm{h}^{-1})^2$	Specific production rate of mAb	0.01	10000.01
$\mathrm{P}_{\mathrm{X_v,QmAb}}\ (\mathrm{c}^2/\mathrm{mL}^2)(\mathrm{g\ cells}^{-1}\mathrm{h}^{-1})$	Initial $Cov(X_v, QmAb)$	0.0	-1.518e-1

Table 12. Standard initial state error covariance matrix (standard **P**(t=0)) for JCKF-Classic, and JCKF-SANTO with run C of Synthetic Dataset.

Parameter	**Name**	$\mathbf{P}_{i,i}$ for JCKF-Classic	$\mathbf{P}_{i,i}$ for JCKF-SANTO
$\mathrm{P_{X_v,X_v}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00001	0.00001
$\mathrm{P_{X_t,X_t}}$ $(\mathrm{c^2/mL^2})$	Viable cells	0.00001	0.00001
$\mathrm{P_{GLC,GLC}}$ $(\mathrm{mM^2})$	Glucose	0.00001	0.00001
$\mathrm{P_{GLN,GLN}}$ $(\mathrm{mM^2})$	Glutamine	0.00001	0.00001
$\mathrm{P_{LAC,LAC}}$ $(\mathrm{mM^2})$	Lactate	0.00001	0.00001
$\mathrm{P_{AMM,AMM}}$ $(\mathrm{mM^2})$	Ammonium	0.00001	0.00001
$\mathrm{P_{mAb,mAb}}$ $(\mathrm{mg/L})^2$	Monoclonal Antibody (titer)	0.00001	0.00001
$\mathrm{P_{QmAb,QmAb}}$ $(\mathrm{g\ cells^{-1}h^{-1}})^2$	Specific production rate of mAb	0.01	10000.01
$\mathrm{P_{X_v,QmAb}}$ $(\mathrm{c^2/mL^2})(\mathrm{g\ cells^{-1}h^{-1}})$	Initial $Cov(X_v, QmAb)$	0.0	-1.581e-1

Table 13. Measurement noise variance **R** and error covariance matrix of process model (**Q**) for the JEKF-Classic, JUKF-Classic and JCKF-Classic with run C of Synthetic Dataset.

Parameter	**Name**	JEKF-Classic	JUKF-Classic	JCKF-Classic
$\mathrm{R^2}$ $(\mathrm{c^2/mL^2})$	Viable cells MNV [1]	$(20\times 10^7)^2$	$(20\times 10^7)^2$	$(20\times 10^7)^2$
$\mathrm{Q_{X_v,X_v}}$ $(\mathrm{c^2/mL^2})$	Viable cells PNV [2]	$(20\times 10^6)^2$	$(20\times 10^6)^2$	$(20\times 10^6)^2$
$\mathrm{Q_{X_t,X_t}}$ $(\mathrm{c^2/mL^2})$	Viable cells PNV [2]	0.001	0.00001	0.00001
$\mathrm{Q_{GLC,GLC}}$ $(\mathrm{mM^2})$	Glucose PNV	0.001	0.00001	0.00001
$\mathrm{Q_{GLN,GLN}}$ $\mathrm{mM^2}$	Glutamine PNV	0.001	0.00001	0.00001
$\mathrm{Q_{LAC,LAC}}$ $(\mathrm{mM^2})$	Lactate PNV	0.001	0.00001	0.00001
$\mathrm{Q_{AMM,AMM}}$ $(\mathrm{mM^2})$	Ammonium PNV	0.001	0.00001	0.00001
$\mathrm{Q_{mAb,mAb}}$ $(\mathrm{VG^2/mL^2})$	Monoclonal Antibody (titer) PNV	0.001	001	0.001
$\mathrm{Q_{QmAb,QmAb}}$ $(\mathrm{h^{-2}})$	Specific production rate of mAb	0.001	0.01	0.1

[1] MNV—measurement noise value; [2] PNV—process noise value.

Table 14. Measurement noise variance **R** and error covariance matrix of process model (**Q**) for the JEKF-SANTO, JUKF-SANTO and JCKF-SANTO with run C of Synthetic Dataset.

Parameter	**Name**	JEKF-SANTO	JUKF-SANTO	JCKF-SANTO
$\mathrm{R^2}$ $(\mathrm{c^2/mL^2})$	Viable cells MNV [1]	$(20\times 10^7)^2$	$(20\times 10^7)^2$	$(20\times 10^7)^2$
$\mathrm{Q_{X_v,X_v}}$ $(\mathrm{c^2/mL^2})$	Viable cells PNV [2]	$(20\times 10^6)^2$	$(20\times 10^6)^2$	$(20\times 10^6)^2$
$\mathrm{Q_{X_t,X_t}}$ $(\mathrm{c^2/mL^2})$	Viable cells PNV [2]	0.001	0.00001	0.00001
$\mathrm{Q_{GLC,GLC}}$ $(\mathrm{mM^2})$	Glucose PNV	0.001	0.00001	0.00001
$\mathrm{Q_{GLN,GLN}}$ $\mathrm{mM^2}$	Glutamine PNV	0.001	0.00001	0.00001
$\mathrm{Q_{LAC,LAC}}$ $(\mathrm{mM^2})$	Lactate PNV	0.001	0.00001	0.00001
$\mathrm{Q_{AMM,AMM}}$ $(\mathrm{mM^2})$	Ammonium PNV	0.001	0.00001	0.00001
$\mathrm{Q_{mAb,mAb}}$ $(\mathrm{VG^2/mL^2})$	Monoclonal Antibody (titer) PNV	0.001	0.001	0.001
$\mathrm{Q_{QmAb,QmAb}}$ $(\mathrm{h^{-2}})$	Specific production rate of mAb	0.001	0.001	0.001

[1] MNV—measurement noise value; [2] PNV—process noise value.