IEC
Institute of Electronics
and Computer

Journal of Artificial Intelligence and Systems, 2023, 5, 139–146
https://iecscience.org/journals/AIS
ISSN Online: 2642–2859

A Novel Extremum Seeking Hyperparameter Design for Joint State and Parameter Estimation of Nonlinear Systems

Selami Beyhan

Izmir Demokrasi University, Electrical and Electronics Engineering, Uckuyular Dist., Karabaglar 35140, Izmir, Turkey.
Email: selami.beyhan@idu.edu.tr
*Corresponding Author: Selami Beyhan, Email: selami.beyhan@idu.edu.tr

How to cite this paper: Selami Beyhan (2023) A Novel Extremum Seeking Hyperparameter Design for Joint State and Parameter Estimation of Nonlinear Systems. Journal of Artificial Intelligence and Systems, 5, 139–146.
https://doi.org/10.33969/AIS.2023050110

Received: August 5, 2023
Accepted: October 12, 2023
Published: January 3, 2024

Abstract

This paper introduces a novel hyperparameter design based on extremum seeking (ES) method to enhance the convergence speed of Extended-Kalman Filter (EKF). ES method produces a real-time optimization output based on the second-order gradient of performance function so that the estimation performance of EKF is simultaneously optimized. In addition to the convergence speed, the proposed hyperparameter reduces the effect of initial covariance matrices, and improves the accuracy of estimation for the fast changes without any knowledge about system dynamics. In numerical applications, EKF with and without the proposed hyperparameter were first used to estimate the unknown parameters of a linear time-varying system. Second, on a real-time collected data, they were applied for the joint estimation of velocity and payload mass of a real-time nonlinear servo-system where the performance improvement is provided almost 30%. Performance measurements are given in terms of the root-mean squared-error (RMSE) of estimation.

Keywords

Extremum seeking, hyperparameter, EKF, velocity and payload estimation.

1. Introduction

Extremum seeking (ES) is a real-time optimization method for static and dynamic systems such that it does not use any explicit information of problem definition or input-out characteristics [1, 2]. Although ES was introduced very early, its stability with standard perturbation was proved in the late 1990s for nonlinear systems. A continuous-time [3] and discrete-time [4] adaptive control schemes were designed for the control of nonlinear systems. In addition, ES based constrained optimization was proposed in [5]. An experimental application of ES method has been presented for the control of microalgae Scenedesmus obliquus culture in a photobioreactor [6] .

In estimation theory, the accuracy, speed of convergence and stability are the prominent characteristics of the methods [7]. EKF is one of the conventional methods for nonlinear systems that can recursively estimate the unmeasurable states and unknown parameters under noise conditions [8]. With proper design of the parameters, EKF is a fast and accurate method so that preferable for real-time applications. Recently, model-based EKF has been designed for power systems to track power quality [9], and applied for minimization of covariance matrices for efficient estimation of bilinear systems [10].

On the other hand, recursive estimation methods have some hyperparameters such that one of them is called ”adaptive step-size” determined based on the user experience and grid-search of an interval. However, this hyperparameter mainly affects the success of estimation process so that it must be selected properly [11]. In this paper, an adaptive hyperparameter based on ES optimization is proposed for the state and parameter estimation. The possible improvement of the hyperparameter is first utilized for the EKF method as a basis for further studies. Proposed hyperparameter ensures that estimated state and parameters are adapted to minimize the cost function of estimation process i.e. at the same time, the cost function of estimation process is collaboratively minimized by EKF and ES dynamics. In numerical applications, valuable results have been recorded that significantly support the proposed method.

2. Extended Kalman Filter

EKF can be designed for the state and parameter estimation of discrete-time nonlinear systems as follows [8, 7].
1. The system dynamics are

$$\begin{aligned} \mathbf{x}(n) &= \mathbf{f}(\mathbf{x}(n-1), \mathbf{u}(n-1), \theta(n-1), w(n-1)) \\ \mathbf{y}(n) &= \mathbf{h}(\mathbf{x}(n), v(n)) \\ w &\sim \mathcal{N}(0, \mathbf{Q}) \\ v &\sim \mathcal{N}(0, \mathbf{R}) \end{aligned} \tag{1}$$

where $f(.)$ and $g(.)$ are nonlinear differentiable functions, $\mathbf{x}$ is the state, θ is the unknown parameter vector and $\mathbf{u}$ is the input of system. The v and w are normally distributed process and output noises, respectively, with $\mathbf{Q}$ and $\mathbf{R}$ the covariance matrices (1).
2. The linearized system dynamics around the current estimate are

$$\begin{aligned} \mathbf{F}_{[i,j]} &= \frac{\partial \mathbf{f}_{[i]}}{\partial \mathbf{x}_{[j]}}|_{(\hat{\mathbf{x}}_{[j]}(n-1), \mathbf{u}(n-1), 0)} \\ \mathbf{N}_{[i,j]} &= \frac{\partial \mathbf{f}_{[i]}}{\partial w_{[j]}}|_{(\hat{\mathbf{x}}_{[j]}(n-1), \mathbf{u}(n-1), 0)} \\ \mathbf{H}_{[i,j]} &= \frac{\partial \mathbf{h}_{[i]}}{\partial \mathbf{x}_{[j]}}|_{(\tilde{\mathbf{x}}_{[j]}(n-1), 0)} \\ \mathbf{M}_{[i,j]} &= \frac{\partial \mathbf{h}_{[i]}}{\partial v_{[j]}}|_{(\tilde{\mathbf{x}}_{[j]}(n-1), 0)} \end{aligned} \tag{2}$$

where i and j represents the indices for functions and estimations. The following matrices are used in the time and measurement update dynamics.

$$\begin{aligned} \mathbf{Q}_c &= \mathbf{N}\mathbf{Q}\mathbf{N}^T \\ \mathbf{R}_c &= \mathbf{M}\mathbf{R}\mathbf{M}^T \end{aligned} \tag{3}$$

3. Time update equations are

$$\begin{aligned} \hat{\mathbf{x}}^-(n) &= \mathbf{f}(\hat{\mathbf{x}}(n-1), \mathbf{u}(n-1), 0) \\ \mathbf{P}^-(n) &= \mathbf{F}(n)\mathbf{P}(n-1)\mathbf{F}^T(n) + \mathbf{Q}_c \end{aligned} \tag{4}$$

4. Measurement update equations are

$$
\begin{aligned}
\mathbf{K}(n) &= \mathbf{P}^{-}(n)\mathbf{H}^{T}(n)(\mathbf{H}(n)\mathbf{P}^{-}(n)\mathbf{H}^{T}(n)+\mathbf{R}_c)^{-1}\\
\hat{\mathbf{x}}(n) &= \hat{\mathbf{x}}^{-}(n)+\mathbf{K}(n)(\mathbf{y}(n)-\mathbf{h}(\hat{\mathbf{x}}^{-}(n),0))\\
\mathbf{P}(n) &= (\mathbf{I}-\mathbf{K}(n)\mathbf{H}(n))\mathbf{P}^{-}(n)
\end{aligned}
\tag{5}
$$

where $\hat{\mathbf{x}}^{-}(n), \mathbf{P}^{-}(n)$ are prior estimates and $\hat{\mathbf{x}}(n), \mathbf{P}(n)$ are the posterior estimates, respectively. The initial value of $\mathbf{P}$ is based on the estimated initial values of the system. Finally, state and parameters are estimated (5) with updated matrices.

3. Extremum Seeking Optimization Based Hyperparameter Design

Extremum seeking is a very powerful mathematical method so that unknown parameters can be optimized with very limited information [4]. ES method is here consulted to produce a prominent parameter that affect all estimation process. For a discrete-time dynamic system,

$$
\begin{aligned}
\psi(n+1) &= \mathbf{f}(\psi(n), u(n)),\\
J^{0}(n) &= h(\psi(n)),
\end{aligned}
\tag{6}
$$

where $\psi \in \mathbb{R}^{n_s}$ is the state, $u(n) \in \mathbb{R}$ is the input, $J^{0}(n) \in \mathbb{R}$ is the nominal cost function. $\mathbf{f} : \mathbb{R}^{n_s} \times \mathbb{R} \to \mathbb{R}^{n_s}$ and $h : \mathbb{R}^{n_s} \to \mathbb{R}$ are smooth functions. The control signal with optimization variable $(\eta(n))$ is

$$
u(n) = \beta(\mathbf{x}(n), \eta(n)), \tag{7}
$$

so that the ESO based closed-loop scheme

$$
\psi(n+1) = f(\psi(n), \beta(\psi(n), \eta(n))), \tag{8}
$$

has an extremum point parametrized by $\eta(n)$. Following assumptions are made for the closed-loop ES scheme.
A1. There exists a smooth function $l : \mathbb{R} \to \mathbb{R}^{n_s}$ such that

$$
f(\psi(n), u(\psi(n), \eta(n))) = 0 \text{ if and only if } \psi(n) = h(l(\eta)). \tag{9}
$$

A2: There exists $\eta^{*} \in \mathbb{R}$ such that

$$
\begin{aligned}
(h \circ l)'(\eta^{*}) &= 0,\\
(h \circ l)''(\eta^{*}) &< 0.
\end{aligned}
\tag{10}
$$

It is also assumed that the equilibrium point
$y = h(l(\eta))$ has a local maximum at $\eta = \eta^{*}$. Finding a local maximum is introduced for the above system dynamics via optimizing the η parameter. Extremum seeking optimization finds the optimal value of the η parameter as shown in Figure 1. The parameter update of the discrete-time stochastic ES [4] is

$$
\begin{aligned}
\hat{\eta}(n+1) &= \hat{\eta}(n) + \varepsilon\rho\zeta(n),\\
\zeta(n+1) &= \zeta(n) - \varepsilon w_1 \zeta(n) w_1 (J(n+1) - \zeta(n)) \sin(v(n+1)),\\
\zeta(n+1) &= \zeta(n) - \varepsilon w_2 \zeta(n) + \varepsilon w_2 J(n+1),\\
J(n+1) &= J^{0}(n) + w(n+1),
\end{aligned}
\tag{11}
$$

where $\rho > 0$, $w_1 > 0$, $w_2 > 0$, $\varepsilon > 0$ are design parameters. $v(n), n = 1, \ldots,$ is assumed to be i.i.d. Gaussian random variable sequence. $w(n)$ is the measurement noise with Gaussian distribution. The convergence theorem with averaging analysis has been given in [3] for the ESO of discrete-time dynamic process.

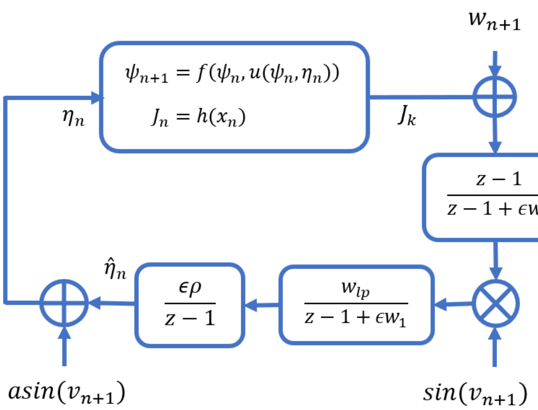


Figure 1. Discrete ES for a dynamic process.

The $\eta(n)$ variable is the optimization output of the ES scheme where its adaptation is only based on the information of cost function which is always trying to find a local maximum with second-order adaptation. We propose to use the $\eta(n)$ parameter as an adaptation term of estimation process that improves the estimation of all variables so called as "hyperparameter". The update of state and parameter estimation have been written as a total estimation process

$$\psi(n) = \psi(n-1) + \eta(n)\Delta\psi, \tag{12}$$

with $\psi(n) = \begin{bmatrix} \hat{\mathbf{x}}(n) \\ \hat{\theta}(n) \end{bmatrix}$ where $\eta(n)$ is obtained by ES scheme. The cost function of estimation process for ES scheme is introduced as the square of estimation error with $J(n) = -(y(n) - \hat{y}(n))^2$. Note that ES optimization run simultaneously to produce $\eta(n)$ parameter in the proposed scheme i.e. $\psi(n)$ is estimated and $\eta(n)$ are optimized at each sample index to optimize the same cost function. EKF estimation has already dynamics to minimize the cost function by true estimate of the parameters, but in the proposed method, ES scheme also try to optimize the same cost function in the embedded form. Therefore, the proposed hyperparameter improves the general performance of estimation.

4. Application Results

We present the application results to verify the effectiveness of proposed approach. Some of results are presented due to the limited space. For a quantitative comparison, the performances are measured in terms of $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\theta(n) - \hat{\theta}(n))^2}$.

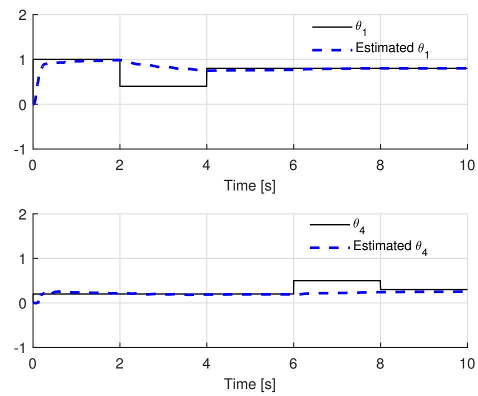
4.1. Identification of a Time-Varying System

A time-varying discrete-time system [12] has been used for system identification. It has a nonlinear term but its parameters are linear-in-parameters form as

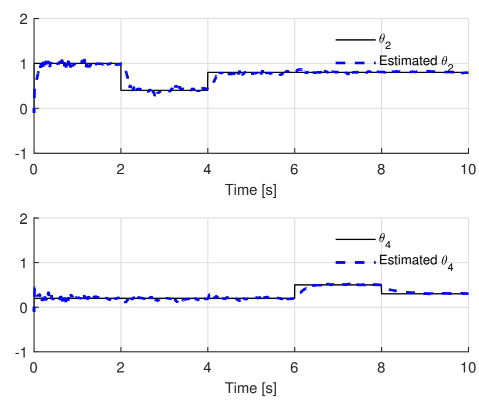
$$y(n) = \theta_1 y(n-1) + \theta_2 y(n-2) + \theta_3 u(n-1) + \theta_4 u(n-2)^2 + v(n) \tag{13}$$

where $v(n)$ is the process noise with normal distribution. It is assumed that θ_2 and θ_4 parameters are randomly changing parameters in time with a probability from a set of values as $\theta_2 \in [0.4, 0.8, 1]$ and $\theta_4 \in [0.2, 0.3, 0.5]$, respectively. On the other hand, the fixed parameters are $\theta_1 = -0.3$ and $\theta_3 = 0.1$, respectively.

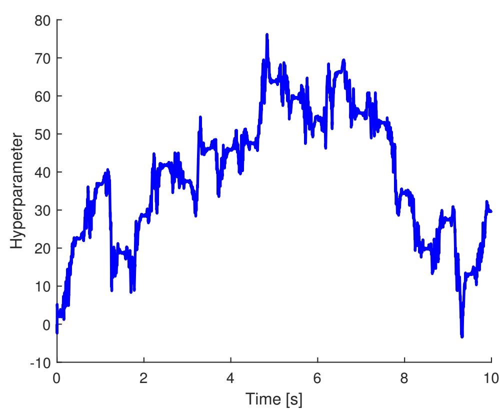
The input signal have four frequency components to improve the persistent excitation, and guarantee the convergence of estimated parameters as $u(n) = 0.7\cos(2\pi n/1000) - 0.3\cos(2\pi n/700) + 0.1\cos(2\pi n/900) - 0.5\cos(2\pi n/800)$. In addition, the process noise is formulated as $v(n) = 0.5\sin(2\pi n/1000)\text{rand}()$ to further increasing of the persistent excitation. Table 2 gives the RMSE performances and Figure 2 shows the estimation results.



(a) EKF based parameter estimation.



(b) ES-EKF based parameter estimation.



(c) Hyperparameter of ES-EKF.

Figure 2. Time varying system identification.

Accurate values of the parameters are obtained in a short-time under process noise.

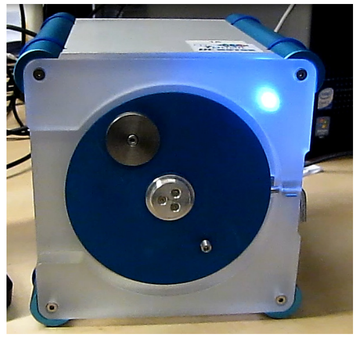


Figure 3. Setup of the nonlinear servo system.

4.2. Payload and Velocity Estimation of a Servo System

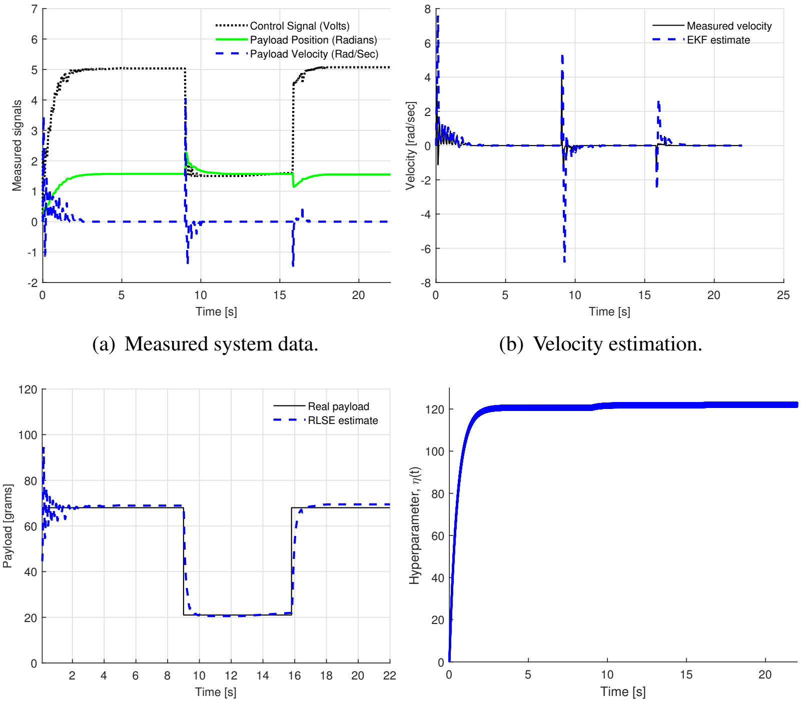
Servo systems are the most common actuators in the industry. In environments where these systems are changing in time, the control of servo systems might be troublesome. Therefore, payload estimation has important place for the accurate control. The dynamics of a nonlinear servo system shown in Figure 3 are given as

$$\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \frac{-K_m^2 - bR_m}{R_m J} x_2 - \frac{gL_m}{J}\sin(x_1) m_L + \frac{K_m}{R_m J} u, \\
y &= x_1,
\end{aligned} \tag{14}$$

where x_1 is the angular position in radians, and x_2 is the angular velocity in rad/sec of the payload, respectively. The payload mass value m_L is physically unknown and piecewise-constant in time. The parameters of servo system are listed in Table 1. Euler method is used for discretization due to the continuous-time dynamics of the nonlinear servo system. The angular position is assumed to be measured output whereas the velocity and payload mass value are estimated by discrete EKF and ES-EKF, respectively. In the estimation, unknown payload is considered as an external parameter and its change with respect to time is written as $\dot{x}_3 = 0$. We add normally distributed artificial noises w and v to the third state and output measurement with the variances $Q = 10^{-1}$ and $R = 10^{-3}$, respectively. For the estimation, the linearized system matrices are given as

$$\begin{aligned}
\mathbf{F} &= \begin{bmatrix} 0 & 1 & 0 \\ -\frac{gL_m}{J}\cos(\hat{x}_1)\hat{x}_3 & \frac{-K_m^2 - bR_m}{R_m J} & -\frac{gL_m}{J}\sin(\hat{x}_1) \\ 0 & 0 & 0 \end{bmatrix} \\
\mathbf{H} &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}
\end{aligned} \tag{15}$$

First, the data is collected when a simple feedback controller is applied to stabilize the payload position at $\pi/2$ where the variation of payload causes an abrupt change of the input voltage and velocity of the payload, respectively. Figure 4 presents the application results on the data where the velocity and payload value are accurately estimated with the help of proposed hyperparameter. The ES based hyperparameter converges to an optimal value. RMSE value of payload estimation is here improved as 30% seen Table 2. Although the nonlinear servo system is estimated in [13], the estimation performance of EKF has been improved in this study.



(a) Measured system data. (b) Velocity estimation.

(c) Payload estimation. (d) Hyperparameter.

Figure 4. ES-EKF estimation results with real-time collected data.

Table 1. Parameters of the servo system.

K_m	Electromotive force constant	0.0536 N m/A
b	Damping coefficient (viscous friction)	3×10^{-6} kg/s
R_m	Electric resistance	9.5
J	Moment of inertia of rotor and disk	1.91×10^{-4} kg m^2
g	Acceleration due to gravity	9.81 m/s^2
L_m	Payload distance from the center	0.042 m
m_L	Unknown payload	$0 - 150$ gr

Table 2. RMSE performances.

Method	System Identification		Method	Servo System Estimation	
	θ_2	θ_4		Velocity [rad/sec]	Payload [grams]
EKF	0.211	0.134	EKF	0.70	7.51
ES-EKF	0.174	0.071	ES-EKF	0.59	5.21

5. Conclusion

This paper focused on a hyperparameter design to improve the speed and accuracy of the estimation. The proposed hyperparameter is the output of extreme seeking method where it optimizes the estimation process in an embedded form. Numerical results with and without designed hyperparameter have been compared where a satisfactory level of enhancement is provided. Resulting that fast convergence of the estimations have been obtained for the applications of discrete-time time-varying system identification and joint velocity/payload estimation of a nonlinear servo system, respectively. As a conclusion, the advantage of proposed hyperparameter can be used with other methods for various estimation problems in future.

Conflicts of Interest

The authors declare that they do not have any conflict of interest regarding the work.

References

[1] M. Krstić, "Performance improvement and limitations in extremum seeking control," *Systems & Control Letters*, vol. 39, no. 5, pp. 313–326, 2000.

[2] K. B. Ariyur and M. Krstic, *Real-time optimization by extremum-seeking control*. John Wiley & Sons, 2003.

[3] S.-J. Liu and M. Krstic, "Stochastic averaging in continuous time and its applications to extremum seeking," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2235–2250, 2010.

[4] J.-Y. Choi, M. Krstic, K. B. Ariyur, and J. S. Lee, "Extremum seeking control for discrete-time systems," *IEEE Transactions on automatic control*, vol. 47, no. 2, pp. 318–323, 2002.

[5] D. DeHaan and M. Guay, "Extremum-seeking control of state-constrained nonlinear systems," *Automatica*, vol. 41, no. 9, pp. 1567–1574, 2005.

[6] C. G. Feudjio Letchindjio, L. Dewasme, and A. Vande Wouwer, "An experimental application of extremum seeking control to cultures of the microalgae scenedesmus obliquus in a continuous photobioreactor," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 7, pp. 1285–1297, 2021.

[7] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[8] G. Welch, G. Bishop, *et al.*, "An introduction to the kalman filter," *Open Source*, 1995.

[9] H. K. Sahoo, U. Subudhi, and S. K. Mishra, "Tracking of power quality disturbances using sparse model–based extended kalman filters," *International Journal of Adaptive Control and Signal Processing*, vol. 32, no. 10, pp. 1489–1507, 2018.

[10] X. Zhang, F. Ding, and E. Yang, "State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 7, pp. 1157–1173, 2019.

[11] S. Beyhan and M. Cetin, "Second-order hyperparameter tuning of model-based and adaptive observers for time-varying and unknown chaotic systems," *Chaos, Solitons & Fractals*, vol. 156, p. 111898, 2022.

[12] F. He, H.-L. Wei, and S. A. Billings, "Identification and frequency domain analysis of non-stationary and nonlinear systems using time-varying narmax models," *International Journal of Systems Science*, vol. 46, no. 11, pp. 2087–2100, 2015.

[13] S. Beyhan, Z. Lendek, M. Alcı, and R. Babuška, "Takagi-sugeno fuzzy observer and extended-kalman filter for adaptive payload estimation," in *2013 9th Asian Control Conference (ASCC)*. IEEE, 2013, pp. 1–6.