1 Introduction

Autonomous and nonautonomous operations of any electromechanical systems, in particular robots, have received considerable attention with respect to both the stability and, more importantly, the efficiency issues. Both the autonomous and the nonautonomous methods, which utilize online and offline/blind optimization and control schemes, respectively, have revealed some advantages and disadvantages. We focus here on the non-autonomous energy-efficient operation of the Baxter manipulator which would subsequently be used in nonlinear control schemes. The offline optimization/control of the robot will be gradually examined with respect to the autonomous practice to yield the most reliable and optimal configuration.

Robots are widely utilized in industry due to their reliable, fast, and precise motions although they are not energy-efficient, and hence, consume a significant lumped amount of energy. The energy consumption and subsequently cost of operation considerably increase when thousands of robots are working together, for example, in a factory, to carry out a network-based task for thousands of cycles. Based on the recent statistics published, industries are among the largest consumers of energy in which the robots take the biggest share of consumption [1]. It is worth mentioning that the robots used in auto industry consume more than half of the total energy required to produce a vehicle body.

The importance of the optimal operation can be visualized through a network of robots operating simultaneously to carry out a specific task defined. The robot manipulator, which is being analyzed in this research work, is operated for thousands of cycles in industries and even at homes as a reliable servant. A considerable lumped amount of energy is expectedly consumed in a trajectory given, as a part of the network-based operation, and it hence needs to be minimized resulting in a significant reduction of operational cost. Therefore, the issue of energy consumed by robots has become a major challenge for researchers and robot manufacturers.

The total mechanical energy consumed by the robot is expectedly affected by the required torque of each joint in addition to the joints’ angular velocities. The high level of energy consumption is typically caused by jerky motions of robots. Many research efforts addressed design optimization [2,3], path planning [4–12], and minimizing joints’ torques [13–16].

Garg and Kumar [16] developed an optimum path requiring the minimum amount of torque which expectedly leads to a minimal amount of energy consumption. They utilized genetic and simulated annealing algorithms by comparing their performance. Other efforts reported in Refs. [17–19] investigated smooth and time-optimal trajectories. Some researchers have focused on path smoothness and/or minimizing the execution time, which may not necessarily yield a minimal amount of energy consumption [20–22].

Extremum seeking (ES) is a model-free optimization approach [23–25] for systems with unknown dynamics and with a measurable output which has been applied to a wide range of technical applications [26–30]. The first proof of stability for an extremum seeking feedback scheme was provided by Krstić and Wang [24]. They utilized the tools of averaging and singular perturbations in revealing that solutions of the closed-loop system converge to a small neighborhood of the extremum of the equilibrium map. Note that the ES approach can yield fast convergence, in spite of being simple to implement by utilizing iterative (batch-to-batch) optimization of the cost function. Frihauf et al. [31] carried out optimization of a single-input discrete-time linear system using discrete-time ES.

Discrete-time extremum seeking with stochastic perturbation was studied without measurement noise in Ref. [32]. Stanković and Stipanović [33] investigated discrete-time extremum seeking with sinusoidal perturbation including measurement noise. Liu and Krstić [34] and Choi et al. [35] employed discrete-time ES for one-variable static system with an extremum using stochastic and sinusoidal perturbations, respectively.

Rotea [36] and Walsh [37] studied multivariable extremum seeking schemes for time-invariant plants. Ariyur and Krstić [38]
investigated, for the first time, the multivariable extremum seeking scheme for general time-varying parameters. Li et al. [39] utilized the multivariable ES in optimizing the cooling power of a tunable thermoacoustic cooler. Other multivariable ES applications can be found in Refs. [40–42].

Through this research effort, the time-invariant multivariable optimization of all joints’ trajectories is presented in detail. To the best of our knowledge, the multivariable ES has not yet been utilized for the minimization of the energy consumed by robotic manipulators. Note that time-energy optimization would not be beneficial by yielding, as expected, a lower amount of energy saved; although the operation time may potentially decrease which is not the immediate concern of this research work [17–19].

The contribution of our work is in employing the multivariable gradient-based discrete-time ES scheme as follows:

1. The scheme is being numerically applied for a seven degrees-of-freedom (7DOF) manipulator and the results implemented experimentally.
2. The scheme’s computational burden is significantly less than other optimization methods including genetic algorithm (GA) which we examine here.

In order to carry out the operational optimization, 14th-order dynamic equations using the Lagrangian method are derived and the results of experimental work are presented to examine the accuracy of the modeling process. Then, the cost function is formulated as the lumped amount of mechanical energy consumption enforcing operational time and torque saturation constraints. The effects of changes of optimization variables on the cost function are studied using global design sensitivity analysis in order to select the most effective ones, and a nominal “S-Shaped” trajectory is fitted for every joint for a collision-free trajectory given. We utilize both extremum seeking and genetic algorithms to improve the dynamic characteristics of the fitted (nominal) trajectories along with minimizing the energy consumption. The optimal trajectory is experimentally implemented and thoroughly compared with the inefficient one.

2 Mathematical Modeling

The redundant manipulator, which is being studied here, has 7DOF as shown in Figs. 1 and 2. The Baxter manipulator’s Denavit–Hartenberg parameters are shown in Table 1 provided by the manufacturer.

The Euler–Lagrange equation expectedly leads to a set of seven second-order ordinary differential equations

\[ M(q)\ddot{q} + C(q, \dot{q})\dot{q} + \phi(q) = \tau \]  

(1)

where \( q, \dot{q}, \) and \( \ddot{q} \in \mathbb{R}^7 \) are rotation angles, angular velocities, and angular accelerations of the joints, respectively, and \( \tau \in \mathbb{R}^7 \) indicates the vector of joints’ driving torques. Also, \( M(q) \in \mathbb{R}^{7 \times 7} \), \( C(q, \dot{q}) \in \mathbb{R}^{7 \times 7} \), and \( \phi(q) \in \mathbb{R}^7 \) are the mass, Coriolis, and gravitational matrices, respectively. This coupled nonlinear dynamic model of the robot is used in the optimization process [43,44].

3 Experimental Validation

The experimental validation of such a coupled nonlinear mathematical model is a necessity to be carried out in order to examine the accuracy of the formulation and then possibly refine the model. We hence recorded the joints’ torques to be compared with the ones computed through the interconnected equations. The results are shown in Fig. 3 for four different joints which, in particular, need significant higher and lower amounts of torques, respectively, to be operated. Note that non-zero torques at the initial point, \( t = 0 \), stand for holding torques against gravity, while the links are stationary leading to zero angular velocities/accelerations at the initial point.

Shown in Figs. 3(a)–3(d) are the experimental and analytically computed torques used in driving the joints \( S_6, E_10, W_6, \) and \( W_1, \)

![Fig. 1 The 7DOF Baxter’s arm](image1)

**Table 1** Baxter’s Denavit–Hartenberg parameters

<table>
<thead>
<tr>
<th>Link/Joint</th>
<th>( a_i )</th>
<th>( d_i )</th>
<th>( z_i )</th>
<th>( \theta_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/s6</td>
<td>0.069</td>
<td>0.27035</td>
<td>−π/2</td>
<td>( \theta_1 )</td>
</tr>
<tr>
<td>2/s6</td>
<td>0</td>
<td>0</td>
<td>π/2</td>
<td>( \theta_2 + \pi/2 )</td>
</tr>
<tr>
<td>3/E6</td>
<td>0.069</td>
<td>0.36435</td>
<td>−π/2</td>
<td>( \theta_3 )</td>
</tr>
<tr>
<td>4/E6</td>
<td>0</td>
<td>0</td>
<td>π/2</td>
<td>( \theta_4 )</td>
</tr>
<tr>
<td>5/W6</td>
<td>0.010</td>
<td>0.37429</td>
<td>−π/2</td>
<td>( \theta_5 )</td>
</tr>
<tr>
<td>6/W6</td>
<td>0</td>
<td>0</td>
<td>π/2</td>
<td>( \theta_6 )</td>
</tr>
<tr>
<td>7/W6</td>
<td>0</td>
<td>0.3945</td>
<td>0</td>
<td>( \theta_7 )</td>
</tr>
</tbody>
</table>

respectively, which reveal an acceptable consistency giving us the confidence to utilize the model developed in the optimization process. Note that the negligible differences potentially root on the unmodeled friction and backlash of the joints.

4 Trajectory Optimization

Baxter uses a simple proportional–integral–derivative (PID) controller for every joint with respect to the initial and end points given in the joint space (while its gains are not adjustable), which yields the jerky motions and subsequently, the inefficient
operation of the robot. The undesirable responses can be observed through the experimental work which we have carried out in our Dynamic Systems and Control Laboratory (DSCL). We observed that the robot collides with other objects close to the end point making the motion unreliable and inefficient. This is counted as a harmful dynamical behavior for both the industrial and home applications. As mentioned earlier, the robot operates using the PID controller which in turn generates energy-inefficient trajectories. Note that Baxter, which is being analyzed here, has been designed for research purposes, and hence, has no predefined nominal trajectory. Therefore, the coupled trajectory optimization of the robot, as a part of the nonautonomous approach, is a necessity to be carried out in order to considerably reduce the mechanical energy consumption along with removing the jerky motions to avoid such a harmful collision discussed earlier.

The optimization needs to be formulated enforcing the operational time and torque saturation constraints to avoid the expected singularities. The feasible joints’ ranges along with the initial and end points are listed in Table 2. Note that one of the physical constraints, which needs to be implemented in the optimization formulation, is zero angular velocity/acceleration at the initial and zero angular velocity at the end points, indicating that the manipulator would remain stationary at those points. This constraint leads us to the well-known “S-Shaped” trajectories which would yield the robot’s smooth dynamical behavior by mitigating the effects of jerky motions. Such a smooth trajectory obviously satisfies the initial/end points’ zero angular velocity condition.

Note that the immediate focus of this research work is on reducing cumbersome computational burden of optimization efforts associated with high-DOF robots including Baxter. We hence propose and utilize the S-shaped trajectory (here \( \tanh(\cdot) \)), based on our previous efforts for operational optimization of smart valves network [45–48], which is highly compatible with the actual motion of Baxter’s joints. From another aspect, using conventional trajectories, including Spline and Bézier in addition to polynomials generating S-shaped ones, expectedly imposes more variables to be optimized with respect to the joint-space optimization; we inevitably need high-order polynomials to generate such a smooth S-shaped trajectory. On the other hand, we only have to provide initial and end points, for each joint, based on Baxter’s operation configuration by employing the S-shaped trajectory.

<table>
<thead>
<tr>
<th>Joints’ name</th>
<th>Range</th>
<th>Initial point</th>
<th>End point</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_0 )</td>
<td>-97.5 to 90</td>
<td>-87.0532</td>
<td>-25.6510</td>
</tr>
<tr>
<td>( S_1 )</td>
<td>-80 to 60</td>
<td>-50.0156</td>
<td>5.0300</td>
</tr>
<tr>
<td>( E_0 )</td>
<td>-170 to 170</td>
<td>-10.1733</td>
<td>41.0350</td>
</tr>
<tr>
<td>( E_1 )</td>
<td>0 to 150</td>
<td>20.1435</td>
<td>65.1590</td>
</tr>
<tr>
<td>( W_0 )</td>
<td>-170 to 170</td>
<td>-30.1357</td>
<td>-85.2770</td>
</tr>
<tr>
<td>( W_1 )</td>
<td>-90 to 115</td>
<td>9.2920</td>
<td>-46.2050</td>
</tr>
<tr>
<td>( W_2 )</td>
<td>-170 to 170</td>
<td>-60.0735</td>
<td>12.0360</td>
</tr>
</tbody>
</table>

Fig. 3 Comparison between the experimentally measured and nominal analytical torques used in driving the joints (a) \( S_0 \), (b) \( E_0 \), (c) \( W_0 \) and (d) \( W_1 \); the non-zero torques at the initial point \( t = 0 \) stand for holding torques against gravity.

Table 2 The ranges of joints’ angles (deg)
Hence, utilizing the polynomials subsequently leads to an undesirable higher computational cost. Also, the logistic functions (for example \( \theta_i(k) = (L_i/1 + e^{-6\theta_i(k)})) \) can also be used as the S-shaped trajectories, but we logically expect to obtain almost the same computational cost with that of the following one (Eq. (2)), whereas seven \( B_i \)'s are again optimized with respect to the practically imposed bounds and constant \( L_i \) for each joint; \( L_i \) like \( A_i \) and \( D_i \) of Eq. (2), is constant. We fit the following nonlinear functions (nominal trajectories) to the joints' actual trajectories which are generated with respect to the initial/end points given in Table 2 using Baxter’s PID controller:

\[
\theta_i(k) = A_i \tan(b_i(k\Delta t)^C_i) + D_i \quad i = 1, \ldots, 7
\]  

(2)

where \( k = 0, 1, \ldots, N, \Delta t \) indicates constant time-step, \( \tau_y = N\Delta t \) (operation time), and \( A_i, B_i, C_i, \) and \( D_i \)'s are calculated utilizing the least square method for the trajectory fitting process listed in Table 3. Note that we discretized the functions due to the discrete-time nature of the problem.

Note that \( A_i \)'s and \( D_i \)'s are constant/unique parameters for each joint and are easily calculated as “End Points – Initial Points” and “Initial Points,” respectively. The parameters \( B_i \)'s and \( C_i \)'s are the optimization variables although a crucial issue to address is the number of parameters expectedly leading to a cumbersome computational cost. Therefore, the sensitivity of the optimization process to the variables of \( B_i \)'s and \( C_i \)'s needs to be carefully addressed.

4.1 Analysis. A global sensitivity analysis has to be carried out in order to examine the roles of \( B_i \)'s and \( C_i \)'s in the optimization process. Typically, the local and global sensitivity analyses are used in determining the effects of changes of the optimization variables on the cost function defined. The local sensitivity analysis (one-at-a-time method) evaluates the effect of one variable on the cost function at a time while keeping the other variables constant, although the global sensitivity analysis utilizes a set of random samples to search the design space with respect to the bounds defined. The global analysis would be an efficient approach as the change of each variable affects the dynamic characteristics of all the joints/links, through the interconnected dynamic equations, and subsequently, the lumped cost function.

To carry out the global sensitivity analysis, we need to numerically calculate the gradient of the cost function with respect to the optimization variables as follows:

\[
\nabla E = \left[ \frac{\partial E}{\partial B_i}, \frac{\partial E}{\partial C_i} \right]^T \quad i = 1, \ldots, 7
\]  

where,

\[
E = \sum_{i=1}^{7} E_i = \sum_{i=1}^{7} \sum_{k=0}^{N-1} \tau_i(k) \theta_i(k) / \Delta t
\]  

(3)

Figures 4 and 5 reveal that the roles of \( B_i \)'s are more drastic than \( C_i \)'s. On the other hand, the effects of \( C_i \)'s are negligible in comparison with those of \( B_i \)'s on the changes of energy consumption. Therefore, all \( B_i \)'s are logically chosen to be optimized using both the ES and the GA. The parameters \( B_i \)'s are optimized subject to the following lower and upper bounds determined through the constraints:

\[
\gamma = [B_1, B_2, B_3, B_4, B_5, B_6, B_7]
\]  

(4)

\[
\gamma_{\min} = [68, 69, 68.5, 69, 66.5, 69.3, 69] \times 10^{-4}
\]  

(5)

\[
\gamma_{\max} = [1385, 1368, 1372, 1368, 1383, 1390, 1386] \times 10^{-4}
\]  

(6)

The lower bound indicates the operational time, which we are willing to keep within \( \tau_y = 8 \) s as a case study. Note that decreasing the lower bound would yield much slower motion which is not desirable and logical, in particular for the industrial applications. We use a semi-analytical approach to determine the upper bound. The Baxter manufacturer provides tables of allowable maximum angular velocities/torques of the joints to avoid the robot catastrophic failure/damage. We utilize the allowable maximum angular velocity and torque to determine the maximum slope of the S-shaped trajectory and the time derivative of slope (\( \theta_i \)), respectively, which can be easily translated to the upper bound of \( B_i \)'s; increasing the upper bound would yield abrupt torques leading to both the motors’ failures and considerably fast motion.

It is worth mentioning that Eq. (2) is the nonlinear function of the \( B_i \) leading to

\[
q(B, k) = [\theta_1(B_1, k), \ldots, \theta_7(B_7, k)]^T
\]  

(7)

\[
\dot{q}(B, k) = [\dot{\theta}_1(B_1, k), \ldots, \dot{\theta}_7(B_7, k)]^T
\]  

(8)

\[
\ddot{q}(B, k) = [\ddot{\theta}_1(B_1, k), \ldots, \ddot{\theta}_7(B_7, k)]^T
\]  

(9)

Therefore, the joint’s torques (Eq. (1)), and subsequently, the mechanical energy consumed are the nonlinear functions of \( B_i \)’s

\[
E(B) = \sum_{i=1}^{7} \sum_{k=0}^{N-1} \tau_i(k) \theta_i(B, k) / \Delta t
\]  

= \sum_{i=1}^{7} \sum_{k=0}^{N-1} \left( D_i(B, k) \dot{\theta}_i(B, k) + C_i(B, k) \dot{\theta}_i(B, k) \right) / \Delta t
\]  

\]  

(10)

where \( D_i(B, k) \) and \( C_i(B, k) \) are the ith rows of \( D \) and \( C \) matrices, respectively.

Therefore, the optimization problem is a constrained one, enforcing the mentioned lower and upper bounds, with the following cost function defined as the lumped amount of mechanical energy consumed in the robot:

\[
\min E(B)
\]  

Subject to: The Interconnected Equations &

\[
\gamma_{\min} \leq \gamma \leq \gamma_{\max}
\]  

(11)

We hence need to optimize seven interconnected variables using both the ES and the GA. One issue to consider is the small values of the variables resulting in serious numerical errors. We fixed this problem by conditioning them using a normalization scheme as follows:

\[
\gamma_n = \gamma \times 10^4
\]  

(12)

Two optimization schemes, including both the ES and the GA, are utilized to avoid being trapped in several possible local minima. The genetic method was developed based on a heuristic search to mimic the process of natural selection [49]. The genetic
algorithm is typically more robust than other conventional schemes. It does not break down easily in the presence of slight changes of inputs and noise. For large-scale state-space equations, the algorithm may potentially exhibit significantly better performance than typical optimization techniques.

Note that for the GA method, we have utilized: (1) “PopulationSize” of 200 for the size of population, (2) “Generations” of 400 which indicates the maximum number of iterations before the algorithm halts, (3) “MigrationFraction” of 0.2 specifying the fraction of individuals in each subpopulation that migrates to a different subpopulation, (4) “MigrationInterval” of 20 standing for the number of generations that take place between migrations of individuals between subpopulations, and (5) Function (TolFun) and constraints (TolCon) tolerances of $10^{-6}$.

5 Multivariable Optimization Using Gradient-Based Extremum Seeking

Our objective is to develop a feedback mechanism minimizing the energy consumed ($E$), where its nonlinear static map is known to have an extremum. We utilize the multivariable extremum seeking scheme [50–53], developed from Krstić and Wang efforts [24], in obtaining optimal values $B_i = [B_1, ..., B_7]^T$.

The extremum seeking scheme estimates the gradient of cost function defined in addition to driving it to zero. The gradient is estimated using a zero-mean external periodic perturbation (or dither signal) and a series of filtering and modulation operations. The convergence of the gradient algorithm is dictated by the second derivative (Hessian) of the cost function. The minimizer is the optimal parameters $B^*$ obtained by driving the system with $B(l) = [B_1(l), ..., B_7(l)]^T$ to determine the cost value $E(l)$ and then iterating the discrete-time extremum seeking to produce $B(l + 1)$; where $l$ denotes the $l$th iteration of the algorithm [31]. Shown in Fig. 6 is a schematic of the discrete-time ES algorithm. It is worth mentioning that the measured output (Fig. 6) passing through a washout (high-pass) filter $W(z) = (z^{-1}/z + h)$, by having zero DC gain, expectedly helps better performance [31,35].

Note that there is a map from $B_i$’s to the energy consumed ($E$) through Eq. (10). The extremum seeking-based optimization shown in Fig. 6 is governed by the following equations:

$$\hat{B}(l) = -\frac{uK}{z - 1} \zeta(l)$$  \hspace{1cm} (13)

$$\zeta(l) = M(l) \frac{z - 1}{z + h} [E(l)]$$  \hspace{1cm} (14)

$$B(l) = \hat{B}(l) + S(l)$$  \hspace{1cm} (15)

where $\hat{B}(l) = [\hat{B}_1(l), ..., \hat{B}_7(l)]^T$, $u$ is a small positive parameter, $K$ is a positive diagonal matrix, and $0 < h < 1$. The notation $P(z)[q(l)]$ is used to denote the signal in the iteration domain that is the output of the transfer function $P(z)$ driven by $q(l)$, where $P(z)$ operates with respect to the iteration domain. Note that $q(l)$ is a signal in iteration domain, where $l$ denotes the $l$th iteration. The perturbation signals $M(l)$ and $S(l)$ are given by

Journal of Dynamic Systems, Measurement, and Control  
Using the Taylor series expansion of the cost function around the local minimum $B^*$ ($\nabla E(B^*) = 0$), the cost function can be written as

$$E(B) = E(B^*) + \frac{1}{2} (B - B^*)^T H (B - B^*)$$

(18)

where $H$ is a positive definite Hessian matrix ($H := (\partial^2 E / \partial B^2)$). Note that cubic and higher order terms are eliminated since they are negligible for local stability analysis via averaging [35]. We then define

$$\tilde{B}(l) = B(l) - B^* = B(l) - S(l) - B^*$$

(19)

Substituting Eq. (19) into Eq. (18) yields

$$E(\tilde{B}(l)) = E(B^*) + \frac{1}{2} (\tilde{B}(l) + S(l))^T H (\tilde{B}(l) + S(l))$$

(20)

$$= E(B^*) + \frac{1}{2} \tilde{B}^T H \tilde{B} + S^T H S + \frac{1}{2} S^T H S$$

(21)

Equation (13) can be therefore rewritten as

$$\tilde{B}(l) = -\frac{\varepsilon K}{z-1} [\zeta(l)] - B^*$$

(22)

which leads to the following difference equation by utilizing Eq. (21) and ($W(z) = z - 1/z + h$):
Due to the fact that which can be rewritten as follows by applying Lemma 3:

\[
\Re \{e^{i(\omega_1 t - \phi_1)} P(e^{i\omega}) \} \Re \{e^{i(\omega_2 t - \phi_2)} Q(z, e^{i\omega})[\nu(t)]\}
\]

\[
= \frac{1}{2} \Re \{e^{i(\omega_1 t - \phi_1) + \omega_2 t - \phi_2)} Q(z, e^{i\omega})[\nu(t)]\}
\]

\[
+ \frac{1}{2} \Re \{e^{i(\omega_1 t - \phi_1) + \omega_2 t - \phi_2} P(e^{i\omega}) Q(z, e^{i\omega})[\nu(t)]\}
\]

(25)

Then, using the result of Eqs. (32) and (33) results in

\[
M(l)W(z)[S(l)^T \hat{H}B(l)] = C^- (l) \Re \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
- S^- (l) \Im \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
+ C^+ (l) \Re \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
- S^+ (l) \Im \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

(34)

where $C^-, S^-, C^+$, and $S^+$ are $7 \times 7$ matrices whose $k$th rows are given by

\[
C_k^- (l) = \left[ \frac{a_1}{a_k} \cos((\omega_1 - \omega_k)l + \phi_k), \ldots, \frac{a_l}{a_k} \cos((\omega_l - \omega_k)l + \phi_k) \right]
\]

(35)

\[
S_k^- (l) = \left[ \frac{a_1}{a_k} \sin((\omega_1 - \omega_k)l + \phi_k), \ldots, \frac{a_l}{a_k} \sin((\omega_l - \omega_k)l + \phi_k) \right]
\]

(36)

\[
C_k^+ (l) = \left[ \frac{a_1}{a_k} \cos((\omega_1 + \omega_k)l - \phi_k), \ldots, \frac{a_l}{a_k} \cos((\omega_l + \omega_k)l - \phi_k) \right]
\]

(37)

\[
S_k^+ (l) = \left[ \frac{a_1}{a_k} \sin((\omega_1 + \omega_k)l - \phi_k), \ldots, \frac{a_l}{a_k} \sin((\omega_l + \omega_k)l - \phi_k) \right]
\]

(38)

The diagonal elements of $C^-$ and $S^-$ are time-invariant. We define diagonal matrices $C_D^-$ and $S_D^-$ containing the diagonal elements of $C$ and $S$, respectively, to highlight these time-invariant terms. Then, Eq. (34) can be rewritten as

\[
M(l)W(z)[S(l)^T \hat{H}B(l)] = \Re \{\tilde{\Omega}(e^{i\omega})\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
+ (C^- (l) - C_D^-) \Re \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
- (S^- (l) - S_D^-) \Im \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
+ C^+ (l) \Re \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
- S^+ (l) \Im \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

(39)

where $\tilde{\Omega}(e^{i\omega}) = \text{diag}[e^{i\phi_1}, \ldots, e^{i\phi_l}]$. Substituting Eq. (39) into Eq. (23) yields the following error dynamics:

\[
\hat{B}(l + 1) - \hat{B}(l) = a(L(z)\hat{H}B) + \Psi^- (l) + \Psi^+ (l)
\]

\[
+ \delta(l)
\]

(40)

where

\[
L(z) = -\frac{K}{2} \left( \tilde{\Omega}(e^{i\omega})\Omega(z, e^{i\omega}) + \tilde{\Omega}(e^{i\omega})\Omega(z, e^{i\omega}) \right)
\]

(41)

\[
\Psi^- (l) = KS^- (l) - S_D^- \Im \{\tilde{\Omega}(z, e^{i\omega})[\hat{H}B]\}
\]

\[
- KC^- (l) \Re \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

(42)

\[
\Psi^+ (l) = KS^+ (l) \Re \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

\[
- KC^+ (l) \Im \{\Omega(z, e^{i\omega})[\hat{H}B]\}
\]

(43)

\[
\Psi_2 (l) = -\frac{1}{2} KM(l)W(z) \left[ \hat{B}(l)^T \hat{H}B(l) \right]
\]

(44)

\[
\delta(l) = -eKM(l)W(z) [E(B^*)]
\]

\[
- \frac{1}{2} eKM(l)W(z) \left[ S(l)^T HS(l) \right]
\]

(45)
It is straightforward to observe that the error dynamics evolve according to a sum of a linear time-invariant term, \( L(z)HB \), linear time-varying functions, \( \Psi^I_1(l) \) and \( \Psi^I_2(l) \), a nonlinear time-varying function, \( \Psi_3(l) \), and a time-varying function, \( \delta(l) \), which is independent of the error of the optimization variable \( B \).

We extend the results reported in Ref. [35] in order to establish the convergence of the homogeneous error system regardless of the convergence of the multivariable discrete-time extremum seeking scheme. We state the local exponential convergence of the averaged homogeneous error system and then consider the full system (Eq. (40)), including the \( \delta(l) \) term, to establish its convergence.

5.1 Convergence of Homogeneous Error System. The following theorem states a sufficient condition for local exponential convergence of the error system (Eq. (40)) regardless of the \( \delta(l) \) term. We deal with the homogeneous error system which is periodic in time \( l \)

\[
\bar{B}(l + 1) - \bar{B}(l) = \varepsilon(L(z)HB) + \Psi^I_1(l) + \Psi^I_2(l) + \Psi_3(l) \quad \text{(46)}
\]

**Theorem 1.** Consider the homogeneous error system (Eq. (46)) with modulation frequencies that satisfy \( \omega_k \neq \omega_l \) for all distinct \( i,k \in \{1,2,\ldots,7\} \) and phase values \( \phi_k \) selected such that \( \text{Re}\{\omega_k \phi_k W(e^{i\omega_k})\} > 0 \) for all \( k \in \{1,2,\ldots,7\} \). There exists a positive constant \( \varepsilon^* \) such that the state-space realization is locally exponentially stable at the origin for all \( 0 < \varepsilon < \varepsilon^* \). This theorem is proved as in Ref. [31].

Note that the eigenvalues of \( H \), denoted by \( \lambda(H) \), play important roles in the convergence rate of the extremum seeking [31].

5.2 Convergence of Full Error System. The established exponential stability of the averaged homogeneous error system helps us investigate the full system (Eq. (40)). We state the convergence properties of \( \delta(l) \) through the following lemma proved in Ref. [31].

**Lemma 4.** The time-varying function \( \delta(l) \) exponentially converges to an \( \mathcal{O}(\varepsilon|a|) \) of zero,

\[
|\delta(l)| \leq \varepsilon^{-1} + c_1\varepsilon|a| \quad \text{(47)}
\]

where \( c_1 \) is a constant and \( a = [a_1,\ldots,a_7] \).

Utilizing the perturbed averaged system [31]

\[
\bar{B}^{\text{ave}}(l + 1) = (I - \varepsilon KH) \bar{B}^{\text{ave}}(l) + \delta(l) \quad \text{(48)}
\]

and Lemma 4, we notice that \( \bar{B}^{\text{ave}}(l) \) converges exponentially to an \( \mathcal{O}(|a|) \)-neighborhood of the origin since \( |\delta(l)| \leq \varepsilon^{-1} + c_1\varepsilon|a| \). From Ref. [54], the exponential convergence rate of \( \bar{B} \) in Eq. (40) tends to the rate of \( \bar{B}^{\text{ave}} \) in the average system as \( \varepsilon \) tends to zero.

We can hence state the convergence of the overall \( B \)-system.

**Theorem 2.** Consider the full system (Eq. (40)) with the conditions of Theorem 1 satisfied. For sufficiently small \( \varepsilon_k \), \( k \in \{1,2,\ldots,7\} \), there exists \( 0 < \varepsilon'_k \leq \varepsilon^* \), such that the error
variable $\tilde{B}$ locally exponentially converges to an $O(|a|)$-neighborhood of the origin for all $0 < \varepsilon \leq \varepsilon_0$.

**Corollary 1.** With the conditions of Theorem 2 satisfied, the cost value $E$ locally exponentially converges to an $O(|a|^2)$-neighborhood of the optimal cost $E(B^*)$.

**Proof.** By defining

$$
\tilde{E}(B) = E(B) - E(B^*)
\quad = \frac{1}{2} \tilde{B}^T \tilde{H} \tilde{B} + \tilde{S}^T \tilde{H} \tilde{B} + \frac{1}{2} \tilde{S}^T \tilde{H} \tilde{S}
$$

Table 4 Optimal trajectories’ coefficients

<table>
<thead>
<tr>
<th>Joints’ name</th>
<th>Optimized $B$ using ES</th>
<th>Optimized $B$ using GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_0$</td>
<td>0.0078</td>
<td>0.00703</td>
</tr>
<tr>
<td>$S_1$</td>
<td>0.0071</td>
<td>0.00700</td>
</tr>
<tr>
<td>$E_0$</td>
<td>0.1354</td>
<td>0.1400</td>
</tr>
<tr>
<td>$E_1$</td>
<td>0.00703</td>
<td>0.00700</td>
</tr>
<tr>
<td>$W_0$</td>
<td>0.01306</td>
<td>0.1398</td>
</tr>
<tr>
<td>$W_1$</td>
<td>0.0703</td>
<td>0.0650</td>
</tr>
<tr>
<td>$W_2$</td>
<td>0.1222</td>
<td>0.1212</td>
</tr>
</tbody>
</table>

Fig. 8 The optimal values of $B$’s using the GA

Fig. 9 The (a) actual and (b) mean value of energy optimized using the ES
From Theorem 2, $\dot{B}$ locally exponentially converges to an $O(|a|)$-neighborhood of the origin. Thus, $E(B)$ locally exponentially converges to an $O(|a|^2)$-neighborhood of the origin.

6 Experimental Results

We used both the analytical (ES) and the numerical (GA) approaches to obtain the optimal values of $B_i$’s shown in Figs. 7 and 8, respectively. The optimal values of $B_i$’s are listed in Table 4 indicating negligible differences between the methods.

It is straightforward to observe that the optimal values of $B_1$, $B_2$, and $B_3$ shown in Figs. 7(a), 7(b), and 7(d), respectively, are lower than the nominal ones indicating that their corresponding links move slower than those of the nominal trajectories. This subsequently leads to a significant reduction in the energy consumed. Note that the joint $S_1$, as expected, takes the biggest share of energy consumption, and therefore, its lower angular velocities
would lead to a lower amount of the cost function defined. Although, the optimal values of $B_1$, $B_3$, $B_5$, and $B_7$ presented in Figs. 7(c) and 7(e–g), respectively, are higher than those of the nominal ones resulting in higher angular velocities of the optimal trajectories than the nominal ones.

Shown in Figs. 9 and 10 are the energy consumptions minimized using both the ES and the GA, respectively. Figure 9(a) presents the energy optimization process versus time while the energy consumed sharply decreases to almost 37 ($J$) and then gradually converges to the optimal value of 36.627 ($J$) (at $t = 84.75s$). Shown in Fig. 9(a) reveals that the optimization of energy consumption fluctuates stochastically, as all the seven parameters ($B_i$’s) are oscillating with seven different frequencies satisfying the mentioned conditions. Therefore, the value of optimal energy is not transparent to be compared with that of the GA one. We hence calculated its mean value over a running average window of one cycle of the specified fundamental low frequency (Fig. 9(b)) to obtain the amount of energy saved

$$\Delta E_{ES} = \frac{E_{nominal} - E_{optimal}}{E_{nominal}} \times 100 = 19.44\% \quad (50)$$

Shown in Fig. 10(a) is the energy consumption minimized using the GA, and its best value is 36.631 ($J$) shown in Fig. 10(b)

$$\Delta E_{GA} = \frac{E_{nominal} - E_{optimal}}{E_{nominal}} \times 100 = 19.21\% \quad (51)$$

From another aspect, Fig. 10 presents considerable computational cost (iterations) of 46,400 for the GA which looks logical with respect to the scale of the coupled dynamic equations resulting in a significant computational time of 2876 s in comparison with 137 s of the ES method. Although Figs. 9 and 10 reveal a negligible difference (less than 1%) for the energy savings of both the schemes, the ES yields the better performance. Such a superior performance of the ES can be justified as follows. The ES carries out optimization by continuously sliding on the cost function in gradient direction rather than finding optimal points discretely with a certain step size of the GA.

Fig. 12 The experimental nominal and optimal trajectories using the ES in sample times of (a) $t = 1$ s, (b) $t = 3$ s, (c) $t = 5$ s, and (d) $t = 6$ s, the robot’s end effector through the nominal trajectory collides with another object due to the jerky motion, while the optimal one avoids such a collision throughout the whole operational time. The shadow frames present the optimal trajectory.
energy, and causes the collision at $t = 6s$. For the optimal case, the manipulator is fast enough, moves toward the end point safely, and no jerky motion can be observed.

7 Conclusion

Through this paper, we presented the interconnected trajectory optimization of a 7DOF Baxter manipulator using both the extremum seeking and heuristic based methods to avoid being trapped in several possible local minima. The coupled dynamic equations of the robot were derived utilizing the Lagrangian method and then validated through the experimental work. We then optimized the joints’ trajectories to generate smooth paths to avoid being exposed to the jerky motions of the nominal ones in addition to minimizing the energy consumption.

The design sensitivity analysis was then carried out to evaluate the effects of changes of the optimization variables on the cost function defined leading to select the most effective ones. Based on the sensitivity analysis, $B_i$’s were optimized to considerably decrease the operation’s energy consumed and also to address the crucial issue of jerky motion. Finally, the optimal trajectory was experimentally implemented and compared with the actual (inefficient) one.

The principal results of this research work can be summarized as follows:

- Using the multivariable discrete-time Extremum Seeking results in a significant decrease in computational cost, an almost twenty-fold reduction relative to Genetic Algorithm.
- A considerable amount of energy is saved (upward of 19%).
- The jerky motion and the subsequent collision between the robot’s end effector and another object close to the end point are removed using the optimal trajectory, which is noted in experimental results.

We are currently focusing our efforts on developing an adaptive controller (to be validated experimentally) to carry an unknown mass to the desired position using an optimal trajectory.

References


