Decision-Directed Adaptive Estimation and Guidance for an Interception Endgame

Dany Dionne* and Hannah Michalska†‡
McGill University, Montreal, Quebec H3A 2A7, Canada

and

Josef Shinar† and Yaakov Oshman§
Technion—Israel Institute of Technology, 32000 Haifa, Israel

A new integrated estimation and guidance design approach is presented as a computationally effective procedure for interception of maneuvering targets. This is an adaptive approach that uses the following elements: banks of state estimators, and guidance laws, a maneuver detector for the onset of the target’s maneuver, and a hierarchical decision law for online selection of the estimator/guidance law pair to be employed. Simulation results confirm that the adaptive approach leads to a reduction in the miss distance as compared with cases in which only a single estimator/guidance law combination is available.

I. Introduction

ALTHOUGH the study presented in this paper was motivated by a future ballistic missile defense (BMD) scenario, it addresses a more general problem, namely, the interception of a randomly maneuvering target by a guided missile in an environment of noise-corrupted measurements. The missile guidance endgame is an imperfect information terminal control problem with a very short horizon and it requires a design approach different from the one used in other control systems. In the classical approach, a linearized model of the dynamic process about a nominal set point is first derived. For this linearized model, the estimator and the control law are designed independently. The separate design is based on assuming the validity of the certainty equivalence principle (CEP) and the associated separation theorem (ST).†

Realistic interceptor guidance, characterized (in addition to noise-corrupted measurements) by bounded controls and saturated state variables, as well as non-Gaussian random disturbances, does not belong to the class of problems for which CEP and ST were proved. Nevertheless, in the 50-yr-long history of guided missiles it has been common practice to design estimators and missile guidance laws separately. This convenient (but suboptimal) design approach has been acceptable because it succeeded in satisfying the performance requirements. The guided missiles had a substantial maneuverability advantage over their manned aircraft targets. Miss distances on the order of a few meters, compatible with the lethal radius (LR) of the missile warhead and its proximity fuse, were considered admissible due to the aircraft structural vulnerability.

This situation was changed by the threat of tactical ballistic missiles (TBM), reintroduced in the 1991 Gulf War and presenting a great challenge to the guided missile community. Successful interception of a TBM, potentially carrying an unconventional warhead, requires a very small miss distance, or even a direct hit. This challenge motivated intensive development of several ballistic missile defense (BMD) systems. All of them were designed using state-of-the-art technology, but with conventional guidance and estimation concepts. Against nonmaneuvering targets, flying on predictable ballistic trajectories, these systems succeeded in demonstrating hit-to-kill performance.‡—§

Re-entering TBMs, as well as modern anti-ship missiles, fly at very high speeds and their atmospheric maneuvering potential is comparable to that of interceptors. Because nonmaneuvering targets can be easily intercepted, the designer of such anti-surface missiles will have to use the option of activating this inherent maneuver potential, which requires only modest technical effort. If an anti-surface missile is maneuvering in a fixed direction, or not maneuvering at all, its trajectory can be considered predictable, thus allowing successful interception. The optimal evasive maneuver in the deterministic sense is a well-timed change (switch) in the direction of the maximum available lateral acceleration. In a realistic interception scenario, the target has no information about the interceptor’s state; therefore it has to maneuver randomly.

The interception endgame has commonly been formulated as a deterministic optimal control problem.† Because target maneuvers are independently controlled and are not known in advance, this classical approach is not adequate. The mathematical framework suitable for the analysis of conflicts controlled by two independent agents is to be found in the area of dynamical games.¶ An interception of a maneuverable target is naturally formulated as a perfect-information zero-sum pursuit–evasion game, where the two independent agents (the players) are the interceptor (pursuer) and the maneuvering target (evader). The game solution provides simultaneously the missile guidance law (the optimal pursuer strategy), the worst target maneuver (the optimal evader strategy), and the resulting guaranteed miss distance (the value of the game). This formulation dates back to the seminal book of Rufus Isaacs.¶ A detailed comparison study§ based on extensive simulations demonstrated the superiority of an interceptor guidance law derived from a perfect-information differential game formulation with bounded controls§ (denoted in recent literature as DGL/1) over those obtained using deterministic optimal control theory.

The optimality of guidance laws derived from a perfect-information differential game formulation, as well as from deterministic optimal control theory, is guaranteed only in a noise-free environment and if the assumption of full state observation, including the knowledge of the target acceleration, is valid. Because acceleration cannot be measured by another moving object, the guidance system needs an observer, even in a noise-free environment. In the case of noise-corrupted measurements, this task is performed by a state estimator whose aim is also to filter out the noise. It is a common
The remainder of this paper is organized as follows. In the next section, the mathematical model of the terminal phase of the interception is presented. It includes the linearized set of the equations of motion, their discretized versions, and the formulation of a relevant stochastic control problem. Section III presents the major elements of the proposed decision-directed adaptive estimation and guidance scheme: the maneuver detector, the bank of estimators, and the bank of guidance laws. In Section IV, the decision-directed adaptive estimation and guidance scheme is described. Section V presents the results of a Monte Carlo simulation study implementing the adaptive scheme in an interception endgame. The last section of the paper summarizes the conclusions drawn from the study.

II. Problem Statement

An interception endgame is a short-horizon terminal control problem describing the pursuit of a maneuverable target by a guided missile. The information structure in such scenarios is generally imperfect, characterized by noise-corrupted measurements acquired by the guided missile (pursuer) on the relative position of the target (evader). The evader has no information about the pursuer, but (being aware that an interception may occur) is likely to perform evasive maneuvers. Optimal control and differential game formulations of the problem,\(^{15,18}\) as well as extensive simulation studies (Ref. 19, p. 104), indicate that the most effective evasion is achieved by well-timed directional reversal of a maximum effort maneuver (also called bang-bang maneuver). In a recent paper,\(^{20}\) it has been shown that spiral maneuvers of the same amplitude generate smaller miss distances if the interceptor missile uses an appropriate estimator.

In this section, the interception endgame is first formulated in a deterministic setting. A stochastic formulation of the problem is subsequently presented.

A. Deterministic Endgame Dynamics

The interception endgame is essentially a three-dimensional nonlinear problem that can be linearized about a nominal collision trajectory, determined by the initial line of sight and by the initial velocity vector of the pursuer. The pursuer heading angle, \(\phi_{p,0}\), required for collision, is determined by

\[
\sin(\phi_{p,0}) = \left(\frac{V_E}{V_P} \sin(\phi_E(0))\right)
\]  (1)

where \(V_P\) and \(V_E\) are the pursuer and evader velocities, respectively, and \(\phi_E(0)\) is the initial heading angle of the evader. The linearization admits a decoupling of the three-dimensional model into two identical sets of planar equations, independently valid in two perpendicular planes.\(^{21}\) Thus, without loss of generality, a single model of linearized planar motion can be considered.

A schematic view of the planar endgame geometry is displayed in Fig. 1. The X-axis is aligned with the initial line of sight that serves as the reference direction. Note that the respective velocity vectors are generally not aligned with the reference line of sight, but they remain close to the directions of the collision course indicated by Eq. (1).

It is assumed that both the pursuer and the evader move with constant speeds and have bounded lateral accelerations \(|a_p| \leq (a)_{\text{max}}\), \(j \in \{E, P\}\). It is further assumed that the dynamics of both players

![Fig. 1 Planar engagement geometry.](image-url)
can be approximated by first-order transfer functions with time constants $\tau_P$ and $\tau_E$, respectively. Because gravity affects both vehicles similarly, it is neglected in the equations of motion.

The (deterministic) nonlinear equations of the planar interception are

$$
\dot{x}_P = V_P \cos(\phi_P), \quad \dot{y}_P = V_P \sin(\phi_P),
$$

$$
\dot{\phi}_P = a_P / V_P, \quad \dot{\phi}_E = a_E / V_E,
$$

$$
\dot{a}_P = (a'_P - a_P) / \tau_P, \quad \dot{a}_E = (a'_E - a_E) / \tau_E,
$$

where $x_P$ and $y_P$ are the positions of the pursuer along the X and Y axes, $x_E$ and $y_E$ are the positions of the evader along the X and Y axes, and $a_P$ and $a_E$ are the lateral accelerations of the pursuer and evader, respectively.

To facilitate linearization, it is assumed that the heading angles, $\phi_P$ and $\phi_E$, are close to the directions of the nominal collision course (1). Let

$$
x = [x_1, x_2, x_3, x_4]^T \triangleq \left[ y \frac{dy}{dt} a_E, a_P \right]^T
$$

be the state vector of the linearized problem, where $y \triangleq y_E - y_P$ is the lateral separation between the evader and the pursuer and $dy / dt$ is the relative lateral velocity. The corresponding linearized differential equations of relative planar motion normal to the reference line and the respective initial conditions are

$$
\dot{x}_1 = x_2, \quad x_1(0) = 0
$$

$$
\dot{x}_2 = x_3 - x_4, \quad x_2(0) = \frac{dy}{dt} \bigg|_{t=0}
$$

$$
\dot{x}_3 = \frac{a'_E - x_3}{\tau_E}, \quad x_3(0) = 0
$$

$$
\dot{x}_4 = \frac{a'_E - x_4}{\tau_E}, \quad x_4(0) = 0
$$

The nonzero initial condition in Eq. (4b) represents the difference between the respective initial velocity components that are not aligned with the initial (reference) line of sight. Because of the assumption of small deviations from the collision geometry, this difference is small compared to the components along the line of sight. Such linearization also yields a constant closing velocity,

$$
V_c = V_P \cos(\phi_{P,\text{col}}) + V_E \cos(\phi_E(0))
$$

making it possible to compute the final time of the interception, $t_f$, for a given initial distance, $X_0$, as

$$
t_f = X_0 / V_c
$$

Based on Eqs. (4), the continuous, time-invariant model of the system becomes

$$
\dot{x}(t) = Ax(t) + B_1 a'_P(t) + B_2 a'_E(t)
$$

where

$$
A = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & -1/\tau_E & 0 & 0 \\
0 & 0 & 0 & -1/\tau_E
\end{bmatrix}, \quad B_1 = \begin{bmatrix}
0 \\
0 \\
0 \\
1/\tau_P
\end{bmatrix}, \quad B_2 = \begin{bmatrix}
0 \\
0 \\
0 \\
1/\tau_E
\end{bmatrix}
$$

The cost function in the deterministic formulation, to be minimized by the pursuer against any admissible evader maneuver command, is defined as the miss distance $M = |x_1(t_f)|$. Thus, the minimization problem for the pursuer is stated as

$$
\inf_{a'_P, a'_E} J, \quad J = |x_1(t_f)|
$$

$$
A'_P \triangleq \left\{ a'_P \in \mathcal{P} \mid a'_P(t) \leq (a'_P)^\text{max} \quad \text{a.e.} \quad t \in [0, t_f] \right\}
$$

where $A'_P$ is the set of admissible acceleration commands (control strategies) of the pursuer and $\mathcal{P}$ denotes the family of piecewise continuous functions.

The measurements performed by the pursuer comprise the range $r(t)$ and the angle of the line of sight $\lambda(t)$ with respect to an inertial reference (the initial line of sight). It is assumed that the range is perfectly measured, but that the measurement of $\lambda(t)$ is corrupted by additive noise. The presence of this noise calls for a stochastic reformulation of the problem.

B. Stochastic Formulation

The information structure of the engagement is imperfect. Because of lack of information about the state of the pursuer, the evader cannot accurately time its required direction reversal. Because no maneuvering, or maneuvering in a fixed direction, may lead to certain interception, the evasive strategy of the evader has to be random. Thus, the target’s acceleration command, $a'_E$, is assumed to be adequately represented by a bounded random process, $\lambda$, subject to additive abrupt changes of unknown magnitude,

$$
a'_E(t) \approx \zeta(t), \quad |\zeta| \leq \zeta^\text{max}
$$

After discretization, the stochastic linear system of the terminal interception problem becomes

$$
x(k + 1) = F x(k) + G_1 a'_P(k) + G_2 \zeta(k) + w(k)
$$

$$
w(k) \sim \mathcal{N}(0, Q_x(k))
$$

$$
y_m(k) = H x(k) + \eta(k)
$$

where $\eta$ is the linearized measurement noise.

The matrices $F$, $G_1$, and $G_2$ of the discrete-time representation of the linear system over a sampling time interval $\Delta$ are (Ref. 22, p. 192)

$$
F = \Phi(\Delta) = \mathcal{L}^{-1}\left( (s I - A)^{-1} \right) \big|_\Delta
$$

$$
\begin{bmatrix}
1 & \Delta & \tau_P(\Delta - \Psi_E) & -\tau_P(\Delta - \Psi_P) \\
0 & 1 & \Psi_E & -\Psi_P \\
0 & 0 & e^{-\Delta/\tau_E} & 0 \\
0 & 0 & 0 & e^{-\Delta/\tau_P}
\end{bmatrix}
$$

$$
G_1 = \int_0^\Delta \Phi(\Delta - \tau) B_1 \, d\tau = \begin{bmatrix}
\tau_P(\Delta - \Psi_P) - \Delta^2/2 \\
\Psi_P - \Delta \\
1 - e^{-\Delta/\tau_E} \\
-\Psi_E + \Delta \\
1 - e^{-\Delta/\tau_P}
\end{bmatrix}
$$

$$
G_2 = \int_0^\Delta \Phi(\Delta - \tau) B_2 \, d\tau = \begin{bmatrix}
-\tau_E(\Delta - \Psi_E) + \Delta^2/2 \\
-\Psi_E + \Delta \\
1 - e^{-\Delta/\tau_E} \\
0
\end{bmatrix}
$$

where

$$
\Psi_{i} \triangleq \tau_i (1 - e^{-\Delta/\tau_i}) \quad i \in \{P, E\}
$$

Using on-board sensors, the two measurements available to the pursuer are the relative angular position, $\lambda$, of the evader with respect to an inertially fixed reference (e.g., the initial line of sight) and
the range, \( r \). It is assumed that the range is measured perfectly but the measurement of the relative angular position is corrupted by a normally distributed additive noise, \( \mu \). Using the small angle approximation, the linearized measurement of the lateral separation, \( y_m \), is

\[
y_m(t) = r(t) \sin(\Delta(t)) + \mu(t)) \approx \mu(t) R_k + r(t) \mu(t)
\]

(13)

where \( r(t) \) denotes the distance to the evader (assumed to be measured perfectly by the pursuer) and \( \mu(t) \) is the angular measurement noise. Thus, the measurement matrix \( H \), and the linearized measurement noise \( \eta \) are

\[
H = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad \eta(t) \overset{\Delta}{=} r(t) \mu \sim \mathcal{N}(0, \sigma^2_r t^2)
\]

(14)

Because of the measurement noise and the presence of random maneuvers, the miss distance, \( M \), becomes a random variable and the cost function needs to be suitably redefined in a probabilistic setting.

A realistic lethality model of the interceptor’s warhead and its detection, as introduced in Ref. 17, is an extension of the original GLR detector, where the concept of a binary indicator \( \hat{E} \) is adopted. While an abrupt change is detected \( \hat{E}(k) = 1 \), otherwise \( \hat{E}(k) = 0 \). An adaptive-\( H_0 \) GLR detector is selected to address the problem of maneuver detection.\(^{(17)}\) The adaptive-\( H_0 \) GLR detector, as introduced in Ref. 17, is an extension of the original GLR detector of Willsky and Johns,\(^{(21)}\) in that it applies to linear systems in which the value of the input variable \( z \) is unknown both before and after an abrupt change; see Eqs. (11).

The GLR detector employed here addresses the basic problem of detecting changes in the mean value of an independent Gaussian sequence when both the onset time and the value of the change are unknown. The main ingredients of the GLR detector are parametric families of input functions, referred to as the hypotheses \( \{\mathcal{H}_i, i = 0, \ldots, w \} \) that describe the unknown input process \( z \). Each family \( \mathcal{H}_i \) is parameterized by the value of the change, \( \theta_i \), and is characterized by a specific, a priori selected onset time instant for the occurrence of the change, \( \kappa_i \). The GLR detector translates the parametric family of input functions into parametric families of distributions for the observations. The distributions of the observations are then estimated online as members of these families. Based on the estimated distributions, a decision concerning the occurrence (or absence) of a maneuver is made, and the characteristics of the maneuver (onset time and value) are derived. The basic tool employed by the GLR detector to estimate the distribution of the observations is the log-likelihood ratio, defined as

\[
L(\mathcal{H}_i, \mathcal{H}_0) \triangleq \log \frac{p(Y^k|\mathcal{H}_i, \theta_i)}{p(Y^k|\mathcal{H}_0, \theta_0)}
\]

(19)

where \( \theta_0, \theta_i \) are parameters and \( \mathcal{Y}^k \) is the \( \sigma \)-algebra generated by the measurements, \( Y^k \equiv \{ y_m(s) : 1 \leq s \leq k \} \). The parameter \( \theta_0 \) describes the input signal to the system before the change. The specific value of \( \theta_0 \), which is assumed known by the GLR detector of Willsky and Johns,\(^{(21)}\) is only estimated by the adaptive-\( H_0 \) GLR detector.\(^{(17)}\)

The adopted statistical approach to such detection relies on maximizing the likelihood ratio twice, first with respect to the parameter \( \theta_i \), and then with respect to the time instant of the change; that is,

\[
g_k = \max_{1 \leq j \leq w} L(\mathcal{H}_j, \mathcal{H}_0)
\]

(20)

where \( g_k \) is the decision function to be used. The precise statement of the conditions on the probability densities under which this double maximization can be performed can be found in Ref. 24. The detection of a change is proclaimed whenever the value of the decision function \( g_k \) reaches or exceeds a given threshold, \( h \), and is performed as follows:

\[
\begin{align*}
& \hat{E}(k) = 0 \quad \forall k \leq h \\
& \hat{E}(k) = 1 \\
\end{align*}
\]

(21)
The scheme defined by Eqs. (20) and (21) belongs to the class of sequential probability ratio tests (SPRT). For a jump-Gaussian linear system, the value of $h$ is chosen to satisfy

$$\alpha = \int_h^{\infty} \chi^2(u) \, du \quad (22)$$

where $\alpha$ is an a priori selected probability of false alarm and $\chi^2$ denotes the central chi-square distribution with one degree of freedom.

The application of the GLR detector to the linear system in Eq. (11) requires the generation of residuals that reflect the changes of interest. The residuals are the (Gaussian distributed) innovations of a Kalman filter (the reference Kalman filter) that is matched to the onset time instant of the change and of the realization of the linear system employed by the Kalman filter augmented with a Wiener process acceleration model. The matrices $Q_a, Q_0,$ and $Q_1$ are different for each member, referred to as $E_0$ and $E_1.$ Both estimators have the same general form of a Kalman filter augmented by a shaping filter. The shaping filter is used as a finite-dimensional linear approximation to the input random process $z.$ (The detector provides an estimate, $\hat{z}_{ML},$ of the process $z,$ but a shaping filter is used to estimate $z$ independently because the value of $\hat{z}_{ML}$ is affected by the detection delay and by possible false detections.) The shaping filter is employed by augmenting the system with a Wiener process acceleration model (Ref. 22, p. 264) in which

$$\text{d}z \approx w_u \, dt$$

where $w_u$ is a zero-mean white Gaussian noise process with power spectral density $Q_u.$ $Q_u$ is referred to as the jerk process intensity. The approximation (25) preserves the autocorrelation function of the random process $z$ whenever a single evasive maneuver is expected and tracks a piecewise-constant input provided that the value of $Q_u$ is chosen to be sufficiently large.

The adaptive- $H_0$ GLR detector has the important ability to modify online the reference realization $a^{H_0}$ such that an adaptive parameter is matched to a reference realization $a^{H_0} \in H_0$ so that, prior to an abrupt change, the mean of the residuals is zero. Thus,

$$\sup_{\delta} L(H_t, H_0) = \frac{1}{2} \frac{d^2(k, i)}{J(k, i)} \quad (23a)$$

$$d(k, i) \triangleq d(k - 1, i) + \rho^T(k, i)V^{-1}(k)\gamma(k), \quad d(k^*, i) = 0 \quad (23b)$$

$$J(k, i) \triangleq J(k - 1, i) + \rho^T(k, i)V^{-1}(k)\rho(k, i), \quad J(k^*, i) = 0 \quad (23c)$$

$$\rho(k, i) \triangleq H \Gamma(k, i) \quad (23d)$$

$$\Gamma(k, i) \triangleq G_2 f_i(k) + F(k - 1)\Gamma(k - 1, i), \quad \Gamma(k^*, i) = 0 \quad (23e)$$

where $\Gamma(k, i)$ is the autocorrelation function of the Kalman gain, respectively, and $f_i$ is an a priori selected member (not identically zero) of the parametric family of input functions associated with hypothesis $H_t.$

After a change is detected, the maximum likelihood estimates of the onset time instant of the change and of the realization of the change are obtained from the maximizing hypothesis as follows:

$$\hat{k}^* = k_{\max}^*$$

$$z_{ML}^*(l) = a^{H_0}(l) + \theta_{ML} f_{ML}(l), \quad l = \hat{k}^*, \ldots, k$$

$$Q_a = 4\left[[z_{ML}^{(\max)}]^2/1_f \right]$$

$$Q_a = (4/25)[[z_{ML}^{(\max)}]^2/1_f]$$

The larger intensity, $Q_a$, is obtained following the formula recommended by Ref. 11 for homing guidance applications against a maneuvering target. The smaller intensity, $Q_{a1},$ is a heuristic trade-off between 1) optimal Gaussian noise rejection (for which $Q_{a1}$ should be set to zero) and 2) providing the filter with a sufficiently broad bandwidth to compensate for errors in the estimates $\hat{k}^*$ and $z_{ML}^*$ (for which $Q_{a1}$ must be sufficiently large). The estimator $E_0$ is designed to be employed when the uncertainty in the system is dominated by the unknown evasive maneuver. The estimator $E_1$ is adequate when the evasive maneuver has been detected and estimated (the uncertainty in the system is then dominated by the Gaussian noise processes).
C. The Bank of Guidance Laws

For simplicity, the bank of guidance laws is also limited to two members, referred to as DGL/C and DGL/1, respectively. First-order dynamics for both the pursuer and the evader and the availability of full state observation are assumed while deriving these laws.

The DGL/1 law was introduced in Ref. 9 as the solution to a linear, perfect-information pursuit–evasion game with bounded controls and with the miss distance \( y_e (t_f) \) as the cost function of the game. This guidance law (the optimal pursuer strategy) has the form

\[
(a^*_P)(t) = (a^*_P)^{\text{max}} \text{sgn}(ZEM_1(t))
\]

where \((a^*_P)\) is the pursuer’s commanded acceleration and \(ZEM_1\) is the zero error miss distance (the miss distance resulting when neither the pursuer nor the evader applies any control command until the end of the game), computed under the assumptions previously made. The explicit expression for \(ZEM_1\) is

\[
ZEM_1(t) = x_1(t) + x_2(t) t_0 - \Delta Z_P(t) + \Delta Z_E(t)
\]

where

\[
\Delta Z_P(t) = \dot{x}_1(t) \tau_P (e^{-\theta_P} + \theta_P - 1)
\]

\[
\Delta Z_E(t) = \dot{x}_1(t) \tau_E (e^{-\theta_E} + \theta_E - 1)
\]

and where \(\theta_P = \frac{t_0}{\tau_P}\) and \(\theta_E = \frac{t_0}{\tau_E}\).

The solution of this perfect information game yields a zero guaranteed miss distance (the value of the game) provided that the pursuer can use larger lateral accelerations than the evader and has noninferior agility (agility is defined as maximum lateral acceleration divided by the time constant, that is, \((a^*_P)^{\text{max}}/\tau_i, i \in \{E, P\}\)).

A. The Governor

The online governor employs the value of the indicator \(E\) to select a state estimator and a guidance law from the respective banks. This selection is based on the current level of uncertainty about the system, which is assessed on the basis of the output values of the maneuver detector and available prior information about the expected number of evasive maneuvers. For simplicity, only a single evasive maneuver (bang–bang with a single commanded change) is assumed to be expected.

B. The Reinitialization of the State Estimator

A reinitialization of the state estimator is employed to improve the accuracy of the state estimate and is achieved by exploiting the information contained in the detector’s estimates \(\hat{z}_{\text{ML}}\) and \(k^*\). Reinitialization is only necessary when the detector updates its current estimates, which takes place only in two situations: when an evasive maneuver is detected at time instant \(k\) (whenever \(E(k) = 0\)) and in the event of a false detection of a maneuver that is indicated by the sequence of events: \(\{E(k - 1) = 1, E(k) = 0\}\). (If more than one evasive maneuver is expected, the definition of a false detection is more complicated.)

The reinitialization of the state estimator at time instant \(k\) requires correcting the value of the state estimate \(\hat{x}(k - 1\hat{k} - 1)\) and of its covariance \(P(k - 1\hat{k} - 1)\).

The reinitialization of the state estimator can be carried out in various ways. The simplest approach to reinitialization is to reuse the previous state estimate without corrections. The drawback is that this simplest approach ignores any new information about the evasive maneuver delivered by the detector and contained in the estimates \(\hat{z}_{\text{ML}}\) and \(k^*\). Recall that the Kalman filters employed in the bank of state estimators provide an estimate of the process \(z\) in the form of the component \(\hat{z}_0\) of the state estimate vector. A second approach to reinitialization is then to reset the state estimate namely, the maneuver detector, the bank of state estimators, and the bank of guidance laws, as well as an online governor. The resulting integrated estimation and guidance approach is adaptive and hierarchical. The structure of the adaptive scheme is depicted in Fig. 3. The functions performed by this scheme are described as follows:

At each time instant, the online governor selects a pair consisting of a state estimator and a guidance law from the respective banks. This selection is based on the current level of uncertainty about the system, which is assessed on the basis of the output values of the maneuver detector and available prior information about the expected number of evasive maneuvers. For simplicity, only a single evasive maneuver (bang–bang with a single commanded change) is assumed to be expected.

\[
(E_i, DGL/j) = \begin{cases} 
(E_0, DGL/C) & \text{if } E(k) = 0 \\
(E_1, DGL/1) & \text{if } E(k) = 1
\end{cases}
\]
so that $\hat{x}_S(l)/l = \tilde{z}^{k^*}_{\text{ML}}(l), l = k^*, \ldots, k$. This particular approach can, however, degrade the state estimate due to the errors in $\tilde{z}^{k^*}_{\text{ML}}$ and $\tilde{k}^*$. A reasonable tradeoff between these two approaches is adopted here. The tradeoff is to constrain the state estimate only at the single time instant $k^*$ by imposing that $\hat{x}_S(k^*) = \tilde{z}^{k^*}_{\text{ML}}(k^*)$. Thus, the state estimate is corrected for the abrupt change occurring precisely at $k^*$.

Moreover, the subsequent state estimates in the interval $(k^*, k]$ will converge to $\tilde{z}^{k^*}_{\text{ML}}$ under the action of the Kalman filter whenever the value of $\tilde{z}^{k^*}_{\text{ML}}$ is correct, or else compensate for any error that might arise in $\tilde{z}^{k^*}_{\text{ML}}$.

To define this procedure more precisely, the reinitialized state estimate and covariance are calculated in two steps.

1) If a reinitialization of the state estimate was made at time $k - 1$, then this previous reinitialization is removed by restoring the original sequence of estimates as follows:

\[
\tilde{x}(k - 1|k - 1)_{\text{old}} = \hat{x}(k - 1|k - 1)_{\text{old}} - \Xi(k - 1)\delta\hat{x}(k - 2)_{\text{old}}
\]

(32a)

\[
\Xi(k - 1)\tilde{\hat{x}}(l - \tilde{k}(k - 1)\tilde{H})\tilde{F}
\]

(32b)

where $\delta\hat{x}(k - 2)_{\text{old}}$ is the correction employed by the previous reinitialization, the subscripts ($\text{old}$ and $\text{new}$) denote the state estimate without and with the previous reinitialization carried out at instant $k - 1$, respectively, $\tilde{k}(l)$ is the Kalman gain, and $\tilde{H}$ and $\tilde{F}$ are the measurement and state transition matrices employed by the state estimators.

If the state estimate was not previously reinitialized at time $k - 1$, then

\[
\hat{x}(k - 1|k - 1)_{\text{old}} = \hat{x}(k - 1|k - 1)_{\text{old}}
\]

(33)

2) The corrected state estimate, $\hat{x}(k - 1|k - 1)_{\text{new}}$, is calculated using the updated values of both $k^*$ and $\tilde{z}^{k^*}_{\text{ML}}(k^*)$. Let $\delta\hat{x}$ be the difference between the estimates of the process $\tilde{x}$ rendered by the detector and by the state estimator; that is, $\delta\hat{x}(l) = \tilde{z}^{k^*}_{\text{ML}}(l) - \hat{x}(l|l)$, and

\[
\hat{x}(k - 1|k - 1)_{\text{new}} = \hat{x}(k - 1|k - 1)_{\text{old}} + \delta\hat{x}(k - 1)_{\text{new}}
\]

(34)

where the correction term $\delta\hat{x}(k - 1)_{\text{new}}$ is obtained from

\[
\delta\hat{x}(l)_{\text{new}} = \Xi(l)\delta\hat{x}(l)_{\text{new}}, \quad l = k^*, k^* + 1, \ldots, k - 1
\]

(35a)

\[
\delta\hat{x}(k^*)_{\text{new}} = [0 \quad \ldots \quad 0 \delta\hat{x}(k^*)]^{T}
\]

(35b)

Eqs. (34) and (35) are derived in Appendix B.

For simplicity, the covariance of the reinitialized state estimate is updated in this work by

\[
P_{\text{new}}(k - 1|k - 1) = P_{\text{old}}(k - 1|k - 1)
\]

(36)

This simple approach neglects the uncertainty in the estimates delivered by the maneuver detector. However, because the estimates delivered by the adaptive-$\tau_0$ GLR detector are consistent, the approximation in Eq. (36) is valid over a sufficiently large time interval.

V. Application to Endgame Guidance

The state estimation and homing performance statistics of the decision-directed adaptive estimation and guidance scheme are obtained from a Monte Carlo simulation of a pursued-evasion engagement between an intercepter (the pursuer) and an incoming ballistic missile (the evader). In the simulated engagement, the dynamics of the system is represented by the nonlinear pursuit-evasion engagement described by the nonlinear set of equations (2). The simulation parameters are listed in Table 1. The measurement frequency, $f$, determines the sampling time interval: $\Delta = 1/f$. The strategy of the evader is a bang–bang maneuver command with a single switch over the time interval of the engagement. The initial heading angles are zero; that is, $\phi_0(0) = 0$ and $\phi_\theta(0) = 0$, and the initial evader’s command acceleration is $a_{\phi}^c = 15$ g. The DGL/C law employs an assumed information delay of $\Delta_{\text{day}} = 0.3$ s.

![Table 1 Simulation parameters](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pursuer velocity $V_p$</td>
<td>2300 m/s</td>
</tr>
<tr>
<td>Evader velocity $V_e$</td>
<td>2700 m/s</td>
</tr>
<tr>
<td>Pursuer maximal acceleration $\alpha_p^{\text{max}}$</td>
<td>30 g</td>
</tr>
<tr>
<td>Evader maximal acceleration $\alpha_e^{\text{max}}$</td>
<td>15 g</td>
</tr>
<tr>
<td>Pursuer dynamics time constant $\tau_p$</td>
<td>0.2 s</td>
</tr>
<tr>
<td>Evader dynamics time constant $\tau_e$</td>
<td>0.2 s</td>
</tr>
<tr>
<td>Initial distance $X_0$</td>
<td>20000 m</td>
</tr>
<tr>
<td>Measurement rate $f$</td>
<td>100 Hz</td>
</tr>
<tr>
<td>Measurement angular noise standard deviation $\sigma$</td>
<td>0.1 mrad</td>
</tr>
<tr>
<td>False alarm probability $\alpha$</td>
<td>0.001</td>
</tr>
</tbody>
</table>

The parameters of the GLR detector are as follows. The detector employs $70$ hypotheses, $w = 70$, describing a possible abrupt change in the evasive command. Each hypothesis, $H_i$, $i = 1, \ldots, w$, requires the selection of two parameters: 1) a time instant for the onset of the abrupt change, $k^*_i$, and 2) a normalized shape for the abrupt change, $f_i$. The hypotheses are selected to represent a bang–bang evasive maneuver by choosing the values of the parameters as follows. The value of $k^*_i$ is taken in the sliding interval $[k - w - 1, k - 1]$ such that each hypothesis $H_i$ has a different value $k^*_i$. Hence, each hypothesis assumes a different time instant for the onset of the change. The value of $f_i$ is chosen to be a nonzero constant value; that is, $f_i(t) = f_i \neq 0$. This choice of $f_i$ means that the evasive command acceleration has a constant value before and after the abrupt change. An abrupt change is then defined as a modification in the value of the evasive command. The value of the evasive command after an abrupt change is estimated by scaling of the normalized shape $f_i$ and by adding this rescaled $f_i$ to the value of the evasive command before the change (the reference realization); see Eq. (24b). The significance of this estimate is assessed by a test of the hypotheses that involves a threshold calculated using Eq. (22) with $\sigma = 0.001$ and computed to be $h = 10.83$. The initial reference realization (employed by the reference Kalman filter) is $a_{\phi}^{\text{ML}}(\tau) = 0$; that is, $a_{\phi}^{\text{ML}}(\tau)$ is initially mismatched with respect to the true evasive command, $\tilde{z}$. The reference Kalman filter also employs a nonzero process noise covariance matrix, $Q$, to provide some bandwidth to compensate for the uncertainties in the initiation of the abrupt change and for possible nonlinearities. This discrete-time process noise covariance matrix is computed as

\[
Q_k = \int_0^\Delta \Phi(\tau) Q \Phi^T(\tau) \, d\tau, \quad Q = \text{diag}(q_{11}, q_{22}, q_{33}, 0)
\]

(37)

where the transition matrix, $\Phi$, is provided in Eq. (12a), and where $q_{11} = 1 \text{ m}^2/\text{s}$, $q_{22} = 10 \text{ m}/\text{s}^3$, and $q_{33} = 1 \text{ m}^2/\text{s}^3$. Finally, the value of the bound for the estimate of $\tilde{z}$ is set to $z_{\text{max}}^{\text{ML}} = 100 \text{ g}$; this is much larger than $z_{\text{max}}^{\text{ML}}$ to allow for the presence of estimation errors in the outputs of the detector, $\tilde{z}$ and $\tilde{z}^{k^*}_{\text{ML}}$.

The following sections compare results obtained using the nonadaptive and the decision-directed adaptive procedures. The state estimation statistics and the homing performance results are presented.

A. The State Estimation Statistics

Let $t_{\text{gosw}}$ be the time-to-go at the onset time instant of the bang–bang evasive maneuver command; that is,

\[
t_{\text{gosw}} = \Delta_f - k^* \Delta
\]

(38)

The state estimation statistics of the decision-directed adaptive estimation scheme is compared to that of the nonadaptive $E_0$ and $E_1$ estimators for an example with $t_{\text{gosw}} = 2.0$ s. The comparison criterion is the average estimation error. The average estimation error was obtained by a Monte Carlo simulation that repeated the engagement 200 times; each repetition employed a different noise realization.

The absolute value of the average estimation error is depicted in Fig. 4 for the estimates of the lateral separation, the relative lateral velocities, and the evader’s acceleration. The decision maker switches between the estimators $E_0$ and $E_1$ after the detection of the change in the evader’s maneuver command.
The zero-effort miss distance during a single engagement example, assumption of no measurement noise, shows the time histories of ror and a faster convergence than the ones from the nonadaptive use the adaptive state estimator exhibit a smaller average estimation error law continues to keep it zero until the evader’s commanded DGL/1 law. However, once the estimator has converged, this guidance to drive the perceived zero-effort miss distance to zero using the DGL/1 law and of the DGL/C law, both using estimator adaptive estimation and guidance scheme are compared with that of B. The Homing Performance

In this section, the homing performance statistics of the integrated adaptive estimation and guidance scheme are compared with that of the DGL/1 law and of the DGL/C law, both using estimator $E_0$.

The first comparison, illustrated in Figs. 5a and 5b under the assumption of no measurement noise, shows the time histories of the zero-effort miss distance during a single engagement example, where $t_{go} = 0.8$ s. The time histories during the entire engagement are shown in Fig. 5a, whereas in Fig. 5b the time histories in the interval $t_{go} \in [0, 0.9]$ s are plotted against a different scale.

These figures provide a clear explanation of the basic characteristics of the three different estimation/guidance strategies. The time histories of the nonadaptive schemes show the basic difference between the DGL/1 and DGL/C laws, both with a wide band estimator ($E_0$).

By looking at the initial phase of the engagement, it is seen that due to the estimation delay it takes a considerable time (about 1 s) to drive the perceived zero-effort miss distance to zero using the DGL/1 law. However, once the estimator has converged, this guidance law continues to keep it zero until the evader’s commanded acceleration changes. This change creates a very large error (17 m) that is observed with a considerable time delay. Because of the short remaining time left in the engagement, this error cannot be fully corrected and the engagement terminates with a miss distance of more than 1 m.

When the DGL/C guidance law is employed, the time history of the zero effort miss distance is quite different. Because this guidance law takes the actual target acceleration into account only partially—see Eq. (30)—it cannot drive the actual zero-effort miss distance to zero at any time instant. Note that in the example scenario the normalized estimation delay (i.e., $\Delta_t/\tau_E$) is 1.5, so only 22% of the actual value of the target acceleration is taken into account by the DGL/C law. When the evader’s commanded acceleration changes, the error created by the abrupt change is smaller (only 10.5 m) due to the “biased” initial value. Therefore, the DGL/C law is able to achieve a miss distance of less than 1 m.

The zero-effort miss distance of the decision-directed adaptive scheme is identical with that of the nonadaptive DGL/C law until an abrupt change in the evader’s command is detected at $t_{go} \approx 0.55$ s. The decision-directed adaptive scheme then adapts the estimator and the guidance law. After detection, the adaptive scheme employs the DGL/1 law, which succeeds in driving the zero effort miss distance to zero shortly before the end of the engagement, thus achieving perfect interception.

In the second comparison, the homing performance statistics are obtained by a Monte Carlo simulation consisting of 200,000 repetitions of the engagement with noisy measurements. The main comparison criterion is $R_k(0.95)$, the required LR of the interceptor for a successful interception with SSKP = 0.95.

The required lethal radii are presented as a function of $t_{go}$, the time-to-go at the onset of the direction change in the evader’s maneuver command.

Each repetition employs a different noise realization and the onset time of the change in the evasive bang–bang maneuver command is randomly chosen from a set of 100 possible instants. The large number of repetitions provides an accurate tail distribution statistics for each of these onset times.

The value of $R_k(0.95)$ is shown in Fig. 6 as a function of the onset time of the evader’s maneuver command reversal. The results presented in the figure can be summarized as follows. The decision-directed adaptive scheme requires a LR that is always smaller than or equal to the LR required by the nonadaptive combination of $E_0$ with the DGL/C law. As compared with the nonadaptive combination of $E_0$ with the DGL/1 law, the decision-directed adaptive scheme requires a smaller LR except for two small regions at the time-to-go
obtained from the narrow-bandwidth filter by employing the DGL/1 law. The improved state estimate of the filter bandwidth as well as to correct the trajectory of the scheme has ample time to improve the state estimate by reducing the miss distance distribution than the combination of the DGL/1 or the DGL/C laws with a fixed estimator. In terms of computational complexity, the adaptive algorithm requires additional computations as claimed by the maneuver detector. The adaptive-H\(_0\) GLR detector with 70 hypotheses has moderate computational requirements, similar to those of an IMM estimator with nine models (see Ref. 28).

VI. Conclusions

The main contribution of this paper is the introduction of a new adaptive approach to improving the homing performance of an interceptor against a randomly maneuvering target. Such an improvement is achieved by exploiting the information generated by a detector of abrupt changes of the evasive maneuver. The resulting decision-directed adaptive estimation and guidance algorithm can be seen as an online optimization procedure permitting to advantageously modify the state estimator and the guidance law. A significant homing improvement is achieved as demonstrated by an example involving a simplified ballistic missile defense scenario. This improvement is expressed in terms of a more favorable cumulative distribution of the miss distance, which translates to a reduced value of the required lethal radius that guarantees a prescribed probability of target destruction.

The exposition in this paper is limited for the sake of simplicity to a given evasive maneuver structure (although the most efficient one) with a single change in the commanded acceleration. Nevertheless, the algorithms can easily be modified to accommodate for multiple maneuvers that could occur according to a prescribed probability distribution. Such an extension would require a careful modification of each of the component blocks: the detector, the bank of estimators, and the bank of guidance laws. Specifically, whereas the present GLR detector could efficiently deal with the detection of the onset time of a spiraling motion, the tracking of changes in the present GLR detector could efficiently deal with the detection of estimators, and the bank of guidance laws. The concept of the interactive hierarchical structure employed in the proposed approach is, however, expected to be helpful, even in this case.

Appendix A: Matrices of the Kalman Filter with a Wiener Process Acceleration Model

The following linear system with a Wiener process acceleration model (discretized over a sampling time interval \(\Delta\)) is employed by the Kalman filter in the bank of state estimators:

\[
\begin{align*}
\dot{x}(k+1) &= \tilde{F} \dot{x}(k) + \tilde{G}_1 \omega_0(k) + \tilde{w}(k), \\
\tilde{w}(k) &\sim \mathcal{N}(0, \tilde{Q}_x) \\
y_m(k) &= \tilde{H} \dot{x}(k) + \eta(k), \\
\eta &\sim \mathcal{N}(0, \tilde{Q}_y).
\end{align*}
\]

(A.1a)

The filter’s state vector is

\[
x = [x_1 \ x_2 \ x_3 \ x_4 \ x_5]^T
\]

(A.2)

where \(x_5\) is the discretized Wiener process approximating \(z\) (see Eq. (25)). The filter’s matrices are

\[
\begin{bmatrix}
1 & \Delta & \tau_{E} a_2 & -\tau_{E} a_2 + \Delta^2/2 \\
0 & 1 & \tau_{E} d_3 & -\tau_{E} d_3 + \Delta \\
0 & 0 & e^{-\Delta/\tau_{E}} & 0 \\
0 & 0 & 0 & e^{-\Delta/\tau_{E}} \\
0 & 0 & 0 & 0 \end{bmatrix}
\]

(A.3a)
\[ \hat{G}_1 = \begin{bmatrix} \tau_p a_2 - \Delta^2/2 \\ \tau_p a_2 - \Delta \\ 0 \\ a_4 \\ 0 \end{bmatrix} \]  
\[ \tilde{H} = [1 \ 0 \ 0 \ 0 \ 0] \]  
\[ \tilde{Q}_w = Q_u = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 & a_{15} \\ a_{12} & a_{22} & a_{23} & 0 & a_{25} \\ a_{13} & a_{23} & a_{33} & 0 & a_{35} \\ 0 & 0 & 0 & 0 & 0 \\ a_{15} & a_{25} & a_{35} & 0 & a_{55} \end{bmatrix} \]

where \( Q_u \in \mathbb{R}^4 \) is the jerk process intensity (see Eq. (25)), and

\[ a_1 = \tau_p e^{-\Delta/\tau_E} - \tau_p + \Delta \]
\[ a_2 = 1 - e^{-\Delta/\tau_E} \]
\[ a_3 = \Delta^3/6 \quad a_4 = 1 - e^{-\Delta/\tau_E} \]
\[ a_{11} = (1/60)[\tau^4(30 - 2\Delta^2/\tau_E + 120e^{-\Delta/\tau_E} - 90) + 60\tau_\Delta^2 \tau_\Delta^2(e^{-\Delta/\tau_E} - 1) + 40\Delta^3 \tau_\Delta^2 - 15\Delta^4 \tau_\Delta^2] \]
\[ a_{12} = \frac{1}{8}[4\tau_\Delta^2(e^{-\Delta/\tau_E} - 2e^{-\Delta/\tau_E} + 1) + 8\tau_\Delta^2(e^{-\Delta/\tau_E} - 1) + 4\Delta^3 \tau_\Delta^2 - 4\Delta^3 \tau_\Delta^2 + 8\Delta^4 \tau_\Delta^2 + 8\Delta^4 \tau_\Delta^2] \]
\[ a_{13} = \frac{1}{8}[6\tau_\Delta^2 + 3\Delta^3 \tau_\Delta^2(e^{-\Delta/\tau_E} - 1) - 1 + \Delta^3] \]
\[ a_{15} = \tau_p e^{-\Delta/\tau_E} - 1 + \Delta^2 \tau_p + \Delta^3/6 \]
\[ a_{22} = \frac{1}{8}[3\tau_\Delta^2(1 - e^{-2\Delta/\tau_E}) + 6\tau_\Delta^2(1 - e^{-2\Delta/\tau_E}) - 6\Delta^2 \tau_\Delta^2 + 2\Delta^3] \]
\[ a_{23} = \frac{1}{8}[\tau_\Delta^2(e^{-2\Delta/\tau_E} - 2e^{-\Delta/\tau_E} + 1) + 2\Delta^2 \tau_\Delta^2(e^{-\Delta/\tau_E} - 1) + \Delta^3] \]
\[ a_{25} = \tau_p e^{-\Delta/\tau_E} - \Delta \tau_p + \Delta^2/2 \]
\[ a_{33} = \tau_p e^{-\Delta/\tau_E} - 1 + \Delta^2 \tau_p + \Delta + 1 \]
\[ a_{35} = \tau_p e^{-\Delta/\tau_E} - 1 + \Delta^2 \tau_p + \Delta + 1 \]  
\[ a_{55} = \Delta \]  

where \( \hat{F}, \tilde{G}, \) and \( \tilde{H} \) and let \( \hat{x}(\cdot) \) be the state estimate. Suppose that at time instant \( k \), it is desired to force the past state estimate \( \hat{x}(k^*|k^*) \), \( k^* < k \), to adopt the value \( \hat{x}(k^*|k^*)_{\text{new}} \). The problem is to find a current state estimate consistent with the modified history of \( \hat{x}(\cdot) \).

Let the subscripts \((\cdot)_{\text{old}}\) and \((\cdot)_{\text{new}}\) denote the variables before and after modification of the estimate history, respectively, and let \( \delta \hat{x}(k^*|k^*) \) be the difference:

\[ \delta \hat{x}(k^*|k^*) = \hat{x}(k^*|k^*)_{\text{new}} - \hat{x}(k^*|k^*)_{\text{old}} \quad k^* < k \]  

By linearity, the difference \( \delta \hat{x}(k^*|k^*) \) is propagated forward in time using the Kalman filter:

\[ \hat{x}(l+1|l) = \hat{F}\hat{x}(l|l) + \tilde{G}u(l) \]  
\[ \delta \hat{x}(l|l) = \hat{x}(l|l-1) + \tilde{K}(l)(y_w(l) - \tilde{H}\hat{x}(l|l-1)) \]

Repetitive applications of the filter Eq. (B1) to Eqs. (B2) yield

\[ \delta \hat{x}(k|k) = \left( \sum_{i=0}^{k-k^*-1} \Xi(k-i) \right) \delta \hat{x}(k^*|k^*) \]

Equation (B.3) can be rewritten in recursive form as

\[ \delta \hat{x}(l|l) = \Xi(l) \delta \hat{x}(l-1|l-1) \quad l = k^* + 1, \ldots, k \]

Thus, by propagating and by reversing Eq. (B.1) and employing Eq. (B.5), the current state estimate consistent with the modification of \( \hat{x}(k^*|k^*) \) is

\[ \hat{x}(k|k)_{\text{new}} = \hat{x}(k|k)_{\text{old}} + \delta \hat{x}(k|k) \]

References


