Abstract—Fault-tolerant control and operation of complex unmanned and aircraft systems is an emerging technology intended to provide the designer and operator with flexibility, interoperability, sustainment and reliability under changing operational requirements or mission profiles. Moreover, it is intended to reconfigure online hardware and software to maintain the operational integrity of the system in the event of contingencies (fault/failure modes). This paper presents an hierarchical architecture that uses available sensor information, fault isolation, failure prognosis, system restructuring and controller reconfiguration. The fault tolerant control framework relies on prognostic information to reconfigure system components and preserve the operational integrity of the aircraft. The hierarchical structure starts at the lowest component level and migrates to the middle system/subsystem level ending with the final mission level. We illustrate the methodology using an electro-mechanical actuator (EMA).

I. INTRODUCTION-BACKGROUND

The emergence of complex and autonomous systems, such as Unmanned Aerial Vehicles (UAVs), have ignited the development of new control paradigms that try to accommodate incipient failures and maintain stability during the emergency situation. As reported in [1] (i) controlled flight into terrain, (ii) loss-of-control in flight and (iii) system/component failure or malfunction, are the primary causes of fatal accidents in the commercial fleet worldwide for the period of 1987-2005. As a result of a joint effort between industry and the government to improve safety, the number of fatal commercial aircraft accidents, dropped by 65% during the period of 1996-2007 [2]. In fact the first cause of accidents (controlled flight into terrain) has been virtually eliminated. On the other hand the same cannot be said for the other two factors. Moreover, it System/component failure and malfunctions are recognized as contributing factors to aircraft loss-of-control in flight.

The NASA Aviation Safety Program (ASP) founded in 1997 [3] is responsible for the application of reconfigurable strategies to general aviation, which till then were meant for military fixed wing aircraft programs. NASA ASP research focuses on vehicle design, manufacturing, operation, and maintenance. Presently, two major NASA ASP initiatives, Integrated Vehicle Health Management (IVHM) and the Integrated Resilient Aircraft Control (IRAC) project, are addressing these needs by funding the private sector, academia and government sponsored laboratories to develop tools to protect against hardware system/component failure or malfunctions.

Joint Strike Fighter (JSF) program is an example of a state of the art aircraft IVHM system, which provides logistic support to the end-user and also provides off-board trending across the entire fleet [4]. Nevertheless, more work is needed in order to develop reliable, effective health management systems exploiting detection, diagnostics, and prognostics for more efficient implementation of mitigation strategies [5]. Moreover, little work has been published regarding the use of prognosis in control system. In 2006, Bogdanov et al. [6], [7] presented a framework to take into account long-term lifetime prediction as a constraint used in an optimal control cost function. This paper presents an integrated framework that explicitly includes fault detection, isolation and failure prognosis to system reconfiguration.

The rest of this paper is organized as follows. Section II presents an overview of fault diagnosis, failure prognosis and fault tolerant control architectures; Sections III and IV present the proposed reconfiguration strategies; Section V evaluates proposed approach using an Electromechanical Actuator (EMA); and Section V concludes the paper highlighting major accomplishments and providing directions for future work.

II. AN INTEGRATING CBM+/PHM END-TO-END ARCHITECTURE FOR FAULT DIAGNOSIS AND FAILURE PROGNOSIS

A rigorous and verifiable framework for diagnosis and prognosis, has been developed, tested and applied to various military and commercial systems at Georgia Tech over the past years. A schematic of the framework, which involves both online as well as offline modules [8], is depicted in Fig. 1 and the fundamental enabling technologies are briefly summarized in the sequel.

Physics of Failure Mechanisms - The foundation for the development and application of Prognostics and Health Management (PHM) technologies is a thorough understanding of the physics of failure mechanisms as critical systems are subjected to stress conditions. From the physical components/systems themselves to a good understanding of how such systems fail and under what conditions leads to optimum Condition Indicator (CI) extraction and selection and, eventually, to accurate diagnostics and prognostics.
Failure Modes and Effects Criticality Analysis (FMECA) - The starting point for “good” diagnostics / prognostics is a thorough FMECA. It describes the failure modes, sensor suite, CIs, possible diagnostics and prognostic algorithms. It forms the first essential step in the systems engineering process for health management of critical aircraft components/systems. FMECA typically proceeds from the bottom-up, i.e. considers the effects of individual component or part failures (faults). It migrates next to the subsystem and system levels. Failure analysis, on the other hand, is a top-down approach that takes advantage of Reliability Centered Maintenance (RCM) methods to trace system faults to failing components.

Sensors and Sensing Strategies - Sensors and sensing strategies constitute the essential requirements for fault diagnosis and failure prognosis of failing components/systems. The type, location and characteristic properties of PHM sensors, i.e. sensors that are specifically designed to monitor fault signatures, present major challenges to the system designer.

Data Pre-Processing - Raw sensor data (current, voltage, vibration, temperature, etc.) are usually noisy and high dimensional. Therefore a preprocessing stage is usually involved which a) reduces the dimensionality of the data and b) removes unwanted noise, thus improving the (fault) Signal to Noise Ratio (SNR). Based on the type of data and the specific application a number of pre-processing can be used including, filtering, Time Synchronous Averaging (TSA) of vibration data etc...

Condition Indicator Extraction and Selection - CI selection and extraction is of paramount importance for fault diagnosis. The objective is to transform high dimensional raw data into tractable low dimensional form (information) without loss of useful information. CI extraction, is an algorithmic process where features/CIs are extracted in a computationally efficient manner from sensor data. The CI extraction usually involves general purpose algorithms. On the other hand CI selection is application dependent and aims at selecting CIs that possesses properties of fault distinguishability and detectability [8]. We will emphasize this module of the architecture since it is the foundational element for good diagnostics and prognostics.

Fault Diagnosis - A fault diagnosis procedure involves the tasks of fault detection, fault isolation and identification (assessment of the severity of the fault) and it is usually either model based or model free (data driven). However intermediate approaches also exist, for example using a particle filter-based module, which is based on a nonlinear dynamic state model,

$$\begin{align*}
    x_d(t+1) &= f_d(x_d(t), u(t)) \\
    x_r(t+1) &= f_r(x_r(t), x_d(t), w(t)) \\
    f_r(t) &= h_r(x_r(t), x_d(t), v(t))
\end{align*}$$

where \(f_d, f_r\) and \(h_r\) are non-linear mappings, \(x_d\) is a collection of Boolean states associated with the presence of a particular operating condition in the system (normal operation, fault type #1, #2, etc.), \(x_r\) is a set of continuous-valued states that describe the evolution of the system given those operating conditions, \(f_p\) is a feature measurement, \(w\) and \(v\) are non-Gaussian distributions that characterize the process and feature noise signals, respectively. The function \(h_r\) is a mapping between the feature value, \(f_p\), and the fault state \(x_r\). This particle filter-based approach gives estimates of fault conditions in a probabilistic manner at any given instant of time, and can be combined with specific confidence and false alarm metrics provided by the user.

Failure Prognosis - The prognostic framework takes advantage of a nonlinear process (fault / degradation) model, a Bayesian estimation method using particle filtering and real-time measurements (Fig. 2). Prognosis is achieved by prediction and filtering. Prediction uses both the knowledge of the previous state estimate and the process model to generate the a priori state pdf estimate for the next time instant,

$$p(x_{t|t-1} | y_{1:t-1}) = \int p(x_{t} | x_{t-1})p(x_{t-1} | y_{1:t-1})$$

Usually, Sequential Monte Carlo (SMC) algorithms, or particle filters, are used to numerically solve this equation in real-time, due to the lack, in most cases of an analytic solution. Particle filtering approximates the state pdf using samples or “particles” having associated discrete probability masses (“weights”) as,

$$p(x_t | y_{1:t}) \approx \frac{1}{w_t} \delta(x_t - x^i_{t|t})$$

where \(x^i_{t|t}\) is the state trajectory and \(y_{1:t}\) are the measurements up to time \(t\). The simplest implementation of this algorithm, the Sequential Importance Re-sampling (SIR) particle filter, updates the weights using the likelihood of \(y_t\) as,

$$w_t \leftarrow w_{t-1} \cdot p(x_t | y_t)$$

Long-term predictions are used to estimate the probability of failure in a system given a hazard zone that is defined via a probability density function with lower and upper bounds for the domain of the random variable, denoted as \(H_{lb}\) and \(H_{up}\), respectively. The probability of failure at any future time is estimated by combining both the weights \(w_{lb,k}^{t+1}\) of predicted trajectories and specifications for the hazard zone (Fig. 3). The resulting RUL pdf, where \(RUL_{lb}\) refers to RUL, provides the basis for the generation of confidence intervals and expectations for prognosis,
\[ \hat{p}_{\text{RUL}} = \sum_{i=1}^{n} p \left( \text{Failure} \mid X = \hat{x}_{\text{exp}}, H_{ib} \cdot H_{ab} \right) \cdot w_{\text{RUL}} \] (5)

Figure 2. Particle filtering-based failure prognosis framework

Performance and Effectiveness Metrics

Performance metrics are used for all major modules of the integrity management architecture. Correlation metrics are defined for the optimum selection and extraction of CIs; confidence and false alarm rates are exploited to ascertain that fault detection results meet customer specifications; several metrics are defined for prognostic routines placing emphasis on different aspects of the prognostic process [9].

Case Study: An Electro-Mechanical Actuator (EMA)

- A case study involving a critical system typically found in many applications-the EMA- is used to demonstrate generic aspects of the framework. EMAs are finding extensive utility as drives for modern aircraft systems, in addition to classical hydraulic devices. An EMA is configured as a closed-loop system consisting of a controller, motor(s), and sensing apparatus like a resolver (top of Fig. 4) [10]. A simulation model of a motor with a turn-to-turn insulation winding fault (bottom of Fig. 4) is the test case for the proposed framework.

III. FAULT TOLERANT CONTROL / ADVERSE EVENT MITIGATION

Definition (Fault Tolerant Control [11]): “Control systems that possess the ability to accommodate system component failures automatically (while) maintaining overall system stability and acceptable performance.”

Traditionally, FTC systems are classified either as passive or active [12], with the former being designed to make the closed loop system robust against system uncertainties and anticipated faults [13].

Model Predictive Control (MPC) MPC, or receding horizon optimal control (RHOC), is a form of control in which the current control action is obtained by solving online, at each sampling instant, a finite horizon open-loop optimal control problem; the first control action in this sequence, of the optimal control sequence is applied to the plant [14], [15].

Therefore, MPC generates a discrete-time controller which takes action at regularly spaced, discrete time instances, where the latest measured output, \( y_k \), and previous measurements, \( y_{k-1}, y_{k-2}, \ldots \), are known.

To calculate the next control input the controller: a) estimates and b) optimizes,

1. Estimation. The controller updates the true value of the controlled variable, \( y_k \), and any internal variables that influence the future trend, (i.e. \( y_{k+1}, y_{k+2}, \ldots \)).

2. Optimization. Values of set points, measured disturbances, and constraints are specified over a finite horizon of future sampling instants, \( k+1, k+2, ..., k+P \) where \( P \in \mathbb{Z}^+ \). The controller computes \( M \) modes \( u_1, u_{k+1}, \ldots, u_{k+M} \), where \( 1 \leq M \leq P \) is referred to as the control horizon.

The MPC is obtained by solving the optimization problem,

\[ J = \int_{0}^{\infty} \left[ (r-y)^T Q (r-y) + \Delta u^T R \Delta u \right] dt + \rho \varepsilon^2 \] (6)

where the variables \( r, y \) and \( \Delta u \) correspond to the input reference, plant output and control correction. \( Q \) and \( R \) are, problem dependent, weight matrices, which are defined a-priori as the inverse of the maximum allowable tracking error and control correction, respectively. An illustration of the non-linear system with MPC is provided in Fig. 5 [10].
Figure 5. Block diagram of MPC with plant and signals

The presented fault-tolerant methodology, builds upon a central theme starting with low-level reconfiguration but also promoting “intelligent” concepts, such as game theory, “smart” search engines, etc., as we migrate to the higher echelons. This way an adverse event mitigation strategy that is mathematically rigorous and generic, while incorporating prognostics, is created.

IV. RECONFIGURATION STRATEGY

The approach to fault-tolerance utilizes a three-tiered architecture as shown in Fig. 6. The highest tier passes down trajectories and performance requirements for all of the subsystems in accordance with the current mission/flight. The middle level manages the subsystems, i.e. each subsystem must satisfy the requirements passed from the high level controller by distributing command signals and performance requirements to all of the components within that subsystem. The lowest level manages the individual components; it predicts the components’ RUL and modifies the control of components in order to extend their RUL. The fault-tolerant control scheme performs control reconfiguration, redistribution and mission/flight adaptation as necessary to meet specified objectives. It should be noted that the adverse event mitigation architecture utilizes real-time prognostic information in the design of the control algorithms. Given accurate on-line prognostic information in terms of estimates of the RUL of a failing component/subsystem, the proactive fault accommodation system manages the accumulation of further damage through control actions until major flight/mission objectives are achieved, despite the fact that the system is in an impaired state [16]. The implications to system survivability, safety and availability to complete a critical flight / mission are significant [17].

The method for estimating fault growth for the system will utilize the fault diagnosis and prognosis routines developed at the component level [18], [19].

Low-Level Reconfiguration

The low-level reconfiguration approach is depicted in Fig. 7.

Mid-Level Redistribution

The middle level of the fault mitigation architecture enables the transition from the component-level reconfiguration to the subsystem and system fault tolerance, thus expanding significantly the practical utility of these emerging technologies.

Figure 6. Block diagram of proposed 3-tier fault-tolerant control

Figure 7. Flowchart of low-level reconfigurable control architecture.

The Redistribution Controller at the middle level is tasked with the rerouting of the remaining available control authority between the subsystems when one or more are experiencing a fault mode. Game theory is a generic means of finding solutions to this type of problem. The problem is set as a two-level hierarchy, where the upper tier is occupied by a supervisor (or manager) who controls a set of variables that affect the behavior of the interacting, goal-oriented players or agents [20]. The players manipulate the control of individual components according to some bounded rational strategy for minimizing their own cost functions, based on limited information about system states and the strategies of the other players. The middle level supervisor will modify the strategies of the players in such a way that the collection converges to a Nash Equilibrium that satisfies both RUL and performance objectives. The utility functions of the players may be maximized via a reinforcement learning scheme or an outer correction loop feedback arrangement.

High-Level Flight / Mission Adaptation

The final and highest level of the Adverse Event Mitigation hierarchy is intended to safeguard strict flight/mission objectives by deploying flight adaptation mechanisms when the middle and low-level of the fault-tolerant control scheme fail to achieve such objectives due to the severity of the contingency. Flight adaptation allows the control architecture to pursue relaxed flight objectives that do not belong to the strict or hard class, to achieve greater
vehicle/system usefulness and absolutely necessary flight goals. The assigned flight objectives are expressed as system performance variables or a sequence of waypoints in the vehicle case. Mission adaptation alters parameters of the system states or of the individual waypoints, such as velocities and accelerations used by the planners to generate flight paths. The mission adaptation component via state or waypoint parameter adaptation enables the system/aircraft to accomplish an altered yet admissible mission. Changes to the parameters can be implemented within the receding window employed by the middle level control redistribution or path re-planning stage [21]-[25].

V. EVALUATION

Thermal Model – For this study an EMA having a brushless DC motor (BLDC) is used, and the investigated fault concerns winding degradation. Since winding degradation is directly related to operating temperature a thermal mode is needed. In the case of a BLDC motor, the winding temperature is related to the power loss in the copper windings, assuming the copper losses are the primary source of power loss. A first order thermo-electrical model is used to describe the relationship between power loss in in the copper windings with respect to the winding-to-ambient temperature.

Prognosis Model - The electrical endurance qualities of insulation materials are affected by temperature and time. In [26] the concept of the ten-degree rule was introduced, stating that the thermal life of insulation is halved for each increase of 10(K) in the exposure temperature. Later, Dakin [27] postulated that the rate of thermal aging of insulation was another way of stating that the rate of temperature-included changes (deterioration) obeyed the Arrhenius chemical rate equation. This is translated to the following equation, which gives the life of insulation aged at elevated temperatures [26]:

\[ L = L_0 \exp \left( \frac{E_a}{k_B T_w} \right) \]  

where \( L \) is the life in units of time (hr), \( L_0 \) a constant of proportionality, \( E_a \) the activation energy (eV), \( T_w \) the winding temperature (K), and \( k_B \) the Boltzmann constant. Accumulating the ratio of RUL during each operating point with the operating temperature, the percentage of life remaining, \( L_{plb}(t) \), can be estimated by:

\[ L_{plb}(t) = 1 - \frac{1}{L_0} \int_0^t \exp \left( -\frac{E_a}{k_B T_w} \right) d\tau \]  

This expression can be used with to project the RUL of the motor windings for a specific commanded input.

Actuator Model - A high-fidelity 5th order state-space model was developed [28]. The model is employed to relate the control inputs and measured outputs of the actuator to the internal system states of the BLDC motor.

Simulation Results

Different operating conditions were simulated to estimate the time evolution of turn-to-turn winding faults of the BLDC. The RUL estimates were generated for different motor currents.

As could someone expect, since an increased temperature, which is directly related to the magnitude of the current, degrades the winding condition, in order to extent the RUL the magnitude of the operating current should be reduced. This is exactly what is performed by the MPC which reduces the operating current magnitude based on the RUL requirement (Fig. 8). Fig. 9 depicts the results of applying MPC for different operating scenarios. Again it is evident that as the RUL reduces (left-to-right), the MPC places more emphasis on reducing the magnitude of the motor current, thus also increasing the rise time of the actuator position increases and at the same time causing a decrease at the magnitude of the winding temperature [10], [28].

VI. CONCLUSIONS AND FUTURE WORK

Fault/failure conditions can have critical implications to life critical applications. Fault-tolerant and reconfigurable control strategies can offer a remedy in such situations improving reliability and survivability. A hierarchical framework to fault tolerant control was presented in this work. The framework uses continues monitoring to detect and diagnose incipient faults and then utilizes prognostic information as part of the reconfiguration strategy. A simulation example was presented to explicitly demonstrate how this approach can be used under a fault scenario occurring in an EMA. The scenario was successfully tackled at the component level. In future work other modules of the integrated fault-tolerant control hierarchy, such as the control re- distribution, mission adaptation, will be addressed.
REFERENCES