Model-based Diagnostics for Propellant Loading Systems

Matthew Daigle
UC Santa Cruz
NASA Ames Research Center
Moffett Field, CA 94035
matthew.j.daigle@nasa.gov

Michael Foygel
Physics Department
South Dakota School of Mines and Technology
Rapid City, SD 57701
michael.foygel@sdsmt.edu

Vadim Smelyanskiy
Intelligent Systems Division
NASA Ames Research Center
Moffett Field, CA 94035
vadim.smelyanskiy@nasa.gov

Abstract—The loading of spacecraft propellants is a complex, risky operation. Therefore, diagnostic solutions are necessary to quickly identify when a fault occurs, so that recovery actions can be taken or an abort procedure can be initiated. Model-based diagnosis solutions, established using an in-depth analysis and understanding of the underlying physical processes, offer the advanced capability to quickly detect and isolate faults, identify their severity, and predict their effects on system performance. We develop a physics-based model of a cryogenic propellant loading system, which describes the complex dynamics of liquid hydrogen filling from a storage tank to an external vehicle tank, as well as the influence of different faults on this process. The model takes into account the main physical processes such as highly non-equilibrium condensation and evaporation of the hydrogen vapor, pressurization, and also the dynamics of liquid hydrogen and vapor flows inside the system in the presence of helium gas. Since the model incorporates multiple faults in the system, it provides a suitable framework for model-based diagnostics and prognostics algorithms. Using this model, we analyze the effects of faults on the system, derive symbolic fault signatures for the purposes of fault isolation, and perform fault identification using a particle filter approach. We demonstrate the detection, isolation, and identification of a number of faults using simulation-based experiments.

TABLE OF CONTENTS

1. INTRODUCTION

The loading of cryogenic spacecraft propellants is an inherently risky and unsafe operation, especially in the case of hydrogen [1–4]. Therefore, diagnostic solutions are necessary to quickly identify when a fault occurs, so that recovery actions can be taken or an abort procedure can be initiated before system safety is compromised. Model-based diagnosis approaches enable quick and robust detection, isolation, and identification of faults, because they rely on a detailed model of system behavior under nominal and faulty conditions.

Applying a model-based approach requires an in-depth analysis and understanding of the underlying physical processes in order to produce an accurate and reliable system model. However, cryogenic propellant loading involves complex physical processes that are difficult to capture. In this paper, we develop a medium-fidelity, lumped-parameter dynamical model of propellant loading that takes into consideration a variety of complex multi-phase phenomena that govern the storage and transfer of cryogenic propellants, yet is simple enough to allow for physics analysis and numerical simulations of real loading systems [5]. We concentrate on a system of liquid hydrogen (LH2) filling that is functionally representative of the Space Shuttle refueling system. In this system, LH2 is stored on the ground in a spherical, insulated, double-walled storage tank (ST), and is transfered to the external vehicle tank (ET) through a network of pipes and valves. The model takes into account the main physical processes such as highly non-equilibrium condensation and evaporation of the hydrogen vapor, pressurization, and also the dynamics of liquid hydrogen and vapor flows inside the system in the presence of helium gas. Since the model incorporates faults in the system, it provides a suitable framework for model-based diagnostics and prognostics algorithms.

We apply a model-based diagnostic approach to the system using a combined qualitative-quantitative diagnosis methodology based on the approach of [6]. Deviations in measured values from model-predicted values are compared to qualitative predictions made using the system model for quick fault isolation. Fault identification is performed using particle filters for joint state-parameter estimation. Using simulation-based experiments, we demonstrate the detection, isolation, and identification of a number of faults.

The paper is organized as follows. Section 2 describes the propellant loading system and its physics model. Section 3 overviews the diagnosis approach. Section 4 describes the fault detection methodology, Section 5 discusses fault isolation, and Section 6 develops the fault identification approach. Section 7 verifies the approach with a number of simulation-based experiments. Section 8 concludes the paper.
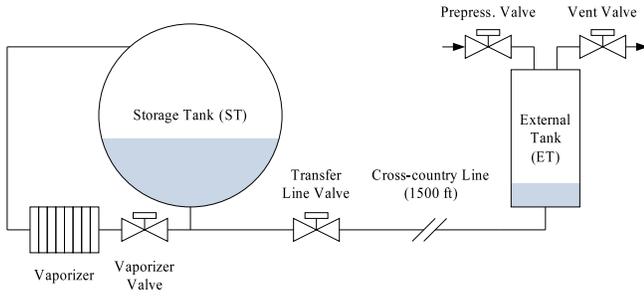
Figure 1. LH2 propellant loading schematic.

2. SYSTEM MODELING

Model-based diagnosis algorithms utilize a model of the system for the fault detection, isolation, and identification tasks. We advocate a physics-based modeling approach, because an understanding of the physical processes in both the nominal and faulty cases is necessary for successful fault identification. Information on fault severity is necessary in order to make appropriate recovery actions in response to a fault, especially in the propellant loading domain. In this section, we develop the physics model of the LH2 propellant loading system. Since the focus of the paper is on diagnostics, we review the main features of the model, and refer to [4, 5] for additional details. We first summarize the filling protocol, followed by mathematical descriptions of the tank histories.

Filling Protocol

The purpose of the LH2 propellant loading system is to move LH2 from the ST to the ET. Fig. 1 shows a simplified, but functionally equivalent schematic of the system. Initially, the ullages of the tanks are at atmospheric pressure due to the presence of gaseous hydrogen (GH2). Before filling, the tanks are first pressurized. The ST is pressurized to 54.7 psia, then 80.7 psia through the use of the vaporizer, which boils off LH2 from the ST and returns the GH2 to the ullage of the ST. The ET is filled with gaseous helium (GHe) through the prepressurization valve, until it reaches 38.7 psia. The purpose of pressurization is two-fold. First, it limits potential boiling of the propellant by keeping a high vapor pressure in the ullage of the tanks. Second, the pressure difference between the ST and ET drives propellant from the ST to the ET in the absence of a pump.

Filling progresses in stages with different filling rates, controlled by the position of the transfer line valve (in reality there are a number of valves between the tanks, but in this paper we consider a simplified representation consisting of a single valve). Slow fill begins first with a low flow rate and chilling of the ET. As the liquid drains out of the ST, its ullage pressure drops, so the vaporizer constantly maintains the ullage pressure to keep LH2 flowing to the ET. The flow through the vaporizer valve is modulated based on the error between the measured ST ullage pressure and the ST pressure set point. As the ET is filled, its ullage volume decreases, and, therefore, its ullage pressure increases. The ullage pressure
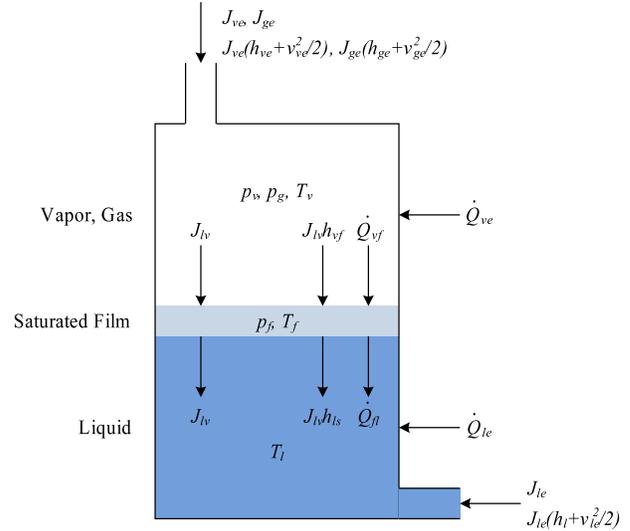
Figure 2. Control Volumes (CV), mass and energy flows in an LH2 tank.

in the ET is maintained using its vent valve, which opens and closes to maintain the pressure between 38.7 and 41.7 psia.

When the ET is 5% full, fast fill begins at a flow rate around 30 kg/s. When the ET is 72% full, the ullage pressure of the ST is reduced to 64.7 psia, reducing the flow rate. When the ET is 85% full, the fill rate is reduced further by partially closing the transfer line valve. When the ET is 98% full, topping begins at a lower flow rate. The ET vent valve is also opened, reducing the ET ullage pressure to 14.7 psia. Finally, at 100% full, topping ends and the tank is then continuously replenished until launch to replace the boil off. During replenish, the fill valve position is modulated to maintain the ET level at 100%.

Tank Modeling

For each tank, we consider three control volumes: the vapor, the liquid, and the vapor film, as shown in Fig. 2. By convention, positive mass/energy flows enter the CV, and negative flows exit the CV.

The vapor CV (v) consists of GH2 (subscript v) and GHe (subscript g), and is treated as a mixture of ideal gases with partial densities $\rho_{v(g)}$ and pressures $p_{v(g)}$, as well as a common temperature T_v, all related to each other by the following equations of state:

$$p_v = \rho_v R_v T_v \tag{1}$$

$$p_g = \rho_g R_g T_v. \tag{2}$$

The liquid CV (l) is where, far from the surface, the temperature is equal to T_l and the liquid is treated as incompressible. T_l is treated as a constant at 20 K.

The vapor film CV (f) separates the liquid and gas phases. It is treated as saturated hydrogen vapor, whose temperature is

equal to that of the liquid/vapor interface, where [1,4,7]

$$T_f = T_C \left(\frac{p_v}{p_C} \right)^n, \qquad (3)$$

with $p_C = 1.315$ MPa, $T_C = 33.2$ K, and $n = 5$ [7].

For the liquid CV, the net mass flow is defined by the external mass flow J_{le} (i.e., the flow across its boundaries other than through the interface), and the interphase (condensation-evaporation) flow J_{lv} (defined later):

$$\dot{m}_l = J_{le} + J_{lv} \qquad (4)$$

Similarly, the GH2 and GHe mass conservation for CV (v) is

$$\dot{m}_v = J_{ve} - J_{lv} \qquad (5)$$
$$\dot{m}_g = J_{ge}. \qquad (6)$$

where J_{ve} and J_{ge} are the external GH2 and GHe flows. For a given tank volume V, the vapor volume is fully defined by the mass of liquid, since it has a constant density:

$$V = V_v + V_l = V_v + m_l/\rho_l. \qquad (7)$$

The energy conservation law for the vapor CV yields:

$$\frac{d(m_v u_v + m_g u_g)}{dt} = \dot{U}_v = \dot{Q}_{ve} - \dot{Q}_{vf} - \dot{W} -$$
$$J_{lv}h_{vf} + J_{ve}(h_{ve} + v_{ve}^2/2) + J_{ge}(h_{ge} + v_{ge}^2/2), \quad (8)$$

where $u_{v(g)} = c_{V,v(g)}T_v$ are specific internal energies of the gases, $\dot{U}_v$ is the net rate of change in internal energy of the mixed gas, $\dot{Q}_{ve}$ is the net external heat flow into the CV (v) through the tank walls, $\dot{Q}_{vf}$ is the heat flow lost through the interface, $\dot{W} = -p_t dV_l/dt$ is related to the quasi-static power due to compression (expansion) of the CV (v), $h_{vf}J_{lv}$ is the interphase enthalpy flow (the specific enthalpy h_{vf} is equal to that of the saturated vapor), h_{ve} and h_{ge} are the specific enthalpies of the hydrogen and helium gas entering the CV, and v_{ve} and v_{ge} are the velocities of the incoming gases [8]. Here, the kinetic energies associated with both the GH2 and GHe mass flows entering the CV (v) are taken into consideration, because the corresponding velocities $v_{v(g)e}$ are much greater than the one related to interphase flow. The specific enthalpies $h_{v(g)} = u_{v(g)} + p_{v(g)}/\rho_{v(g)} = u_{v(g)} + R_{v(g)}T_v = c_{P,v(g)}T_v$. Here, $c_{P,v(g)} = c_{V,v(g)} + R_{v(g)}$ and is the specific heat at constant pressure. The temperature of the mixed gas is then described by

$$\dot{T}_v = \frac{1}{m_v c_V}(\dot{U}_v - \dot{m}_v c_{V,v} T_v - \dot{m}_g c_{V,g} T_v), \qquad (9)$$

where $c_{V,v}$ and $c_{V,g}$ are the specific heats at constant volume for the vapor and gas, and c_V is the specific heat at constant volume for the mixed gas.

If the film layer is considered negligibly thin so that one can ignore its mass [1], then the energy balance equation for the CV (f) can be written as

$$\dot{Q}_{vf} - \dot{Q}_{fl} + J_{lv}h_{lv} = 0, \qquad (10)$$

where $\dot{Q}_{vf}$ is the heat flow from the vapor CV, $\dot{Q}_{fl}$ is the heat flow to the liquid CV, and h_{lv} is the enthalpy (heat) of vaporization. Strictly speaking, h_{lv} depends on the saturated vapor temperature, such that it goes to zero when the surface temperature approaches the critical temperature T_C [9]. To take this effect into consideration, we use the following simple interpolation formula for $T_f \le T_C$:

$$h_{lv}(T_f) = h_{lv}^0 \left(\frac{T_C - T_f}{T_C - T_l} \right)^{1/2}, \qquad (11)$$

where for liquid hydrogen, $T_C = 33$ K and $h_{lv}^0 \cong u_{lv}^0 = 4.5 \times 10^5$ J/kg at $p = 1$ atm and $T_l = 20$ K [9]. The heat flow terms may be computed based on the liquid, vapor, and film temperatures, allowing for J_{lv} to be computed with (10).

The heat flows $\dot{Q}_{fl}$ and $\dot{Q}_{vf}$ are dominated either by conduction or convection, depending on the relative temperatures of the liquid, vapor, and film [5]. If $T_f > T_l$ then conduction heat transfer defines $\dot{Q}_{fl}$, else convection does. If $T_v > T_f$, then conduction heat transfer defines $\dot{Q}_{vf}$, else convection does. In the case of conduction, we use the following approximation:

$$\dot{Q}_{fl}^{cond} = A_f \alpha_{fl}^{cond}(T_f - T_l) \qquad (12)$$
$$\dot{Q}_{vf}^{cond} = A_f \alpha_{vf}^{cond}(T_v - T_f), \qquad (13)$$

where A_f is the surface area of the interface, and the α terms are heat transfer coefficients (see [5]). Typically, conduction expressions contain complex integral relations, but here, we use this algebraic approximation that has proven adequate for our system [5]. For convection, we use

$$\dot{Q}_{fl}^{conv} = A_f \alpha_{fl}^{conv}(T_f - T_l) \qquad (14)$$
$$\dot{Q}_{vf}^{conv} = A_f \alpha_{vf}^{conv}(T_v - T_f). \qquad (15)$$

Both the liquid and vapor CVs absorb external (e) heat from the tank walls. This heat is transferred by means of convection so that

$$\dot{Q}_{v(l)e} = A_{v(l)} \alpha_{v(l)e} \left(T_w - T_{v(l)} \right), \qquad (16)$$

where $A_{v(l)}$ are the internal tank surfaces in contact with vapor (liquid), and $\alpha_{v(l)e}$ are the convection heat transfer coefficients [5]. The wall temperature T_w is governed by the heat flow passing through the walls from the environment [1–3].

The temperature T_w of the tank wall, considered uniform, is defined by the heat exchange rate with the tank surroundings with the effective ambient temperature T_a:

$$\dot{Q}_w = A_w \alpha_w \left(T_a - T_w \right), \qquad (17)$$

where A_w is the external tank surface area. The wall temperature is governed by the tank energy conservation:

$$m_w c_w \dot{T}_w = \dot{Q}_w - \dot{Q}_{le} - \dot{Q}_{ve}. \qquad (18)$$

Here, the heat transfer coefficients describe natural convection inside and outside the tank walls [1, 10].

The Storage Tank

The above equations apply equally well to both the ST and the ET. We denote variables of the ST with a 1 subscript, and variables of the ET with a 2 subscript. For the ST, there is no GHe in the ullage volume, so $p_{g1} = \rho_{g1} = 0$. The external mass flow for the liquid consists of vaporizer flow J_{vap}, transfer line flow J_{tr}, and leak flow $J_{,leak,1}$:

$$J_{le1} = -J_{vap} - J_{tr} - J_{l,leak1}. \tag{19}$$

J_{tr} is described by

$$J_{tr} = \lambda_{tr}\sigma_{tr}\alpha_{tr}\sqrt{p_1 - p_2}, \tag{20}$$

where λ_{tr} is the valve position (an input), α_{tr} is the flow coefficient, $p_1 = p_{v1} + \rho_l g h_{l1}$ is the total ST pressure where h_{l1} is the height of the liquid, $p_2 = p_{v2} + p_{g2} + \rho_l g h_{l2}$ is the total ET pressure, g is the acceleration due to gravity and σ_{tr} is a (dimensionless) multiplicative factor describing a transfer line blockage fault.

In the vaporizer, a certain amount of LH2 is evaporated and returned to the ST, thus controlling its ullage pressure. We assume that all liquid flow through the vaporizer is converted instantaneously to GH2. The flow is given by

$$J_{vap} = \lambda_{vap}\sigma_{vap}\alpha_{vap}\sqrt{p_{v1} - p_{vap}} \tag{21}$$

where λ_{vap} is the vaporizer valve position, α_{vap} is the flow coefficient, p_{vap} is the pressure in the vaporizer (close to atmospheric pressure), and σ_{tr} is a multiplicative factor describing the vaporizer valve blockage fault. The vaporizer valve position is controlled by

$$\lambda_{vap} = \min\left(1, \max\left(0, 10\frac{p_{v1}^* - p_{v1}}{p_{v1}^*}\right)\right), \tag{22}$$

where p_{v1}^* is the desired ST ullage pressure.

The liquid leak flow is given by

$$J_{l,leak,1} = A_{l,leak,1}\alpha_{l,leak,1}\sqrt{p_1 - p_{atm}}, \tag{23}$$

where $A_{l,leak,1}$ is the the area of the leak hole, $\alpha_{l,leak,1}$ is the leak hole flow coefficient, and p_{atm} is atmospheric pressure.

In the vapor CV, $J_{ve1} = -J_{v,leak,1}$, where

$$J_{v,leak,1} = A_{v,leak,1}\alpha_{v,leak,1}\sqrt{p_{v1} - p_{atm}}. \tag{24}$$

Heat leaks constitute the external heat flow term, i.e., $\dot{Q}_{ve1} = \dot{Q}_{leak,1}$.

The External Tank

For the ET, the external liquid flow is given by

$$J_{le2} = J_{tr} - J_{l,leak,2} - J_{boil}, \tag{25}$$

where $J_{l,leak2}$ is described similar to that for the ST, and $J_{boil} = \dot{Q}_{le2}/h_{lv}(T_{w2} - T_{l2})$ is responsible for intense LH2 evaporation as the ET walls are being initially chilled down during the beginning of the slow fill stage.

In the vapor CV, the external flows are given by

$$J_{ve2} = -J_{v,vent,2} - J_{v,leak,2} \tag{26}$$

$$J_{ge2} = J_{g,pp} - J_{v,vent,2} - J_{g,leak,2}, \tag{27}$$

where $J_{g,pp}$ is the prepressurization flow of GHe, and

$$J_{v(g),vent,2} = \lambda_{vent,2}\sigma_{vent,2}\frac{A_{vent,2}\rho_{v(g)}\sqrt{\gamma(p_{v2}+p_{g2}-p_{atm})}}{\Gamma\sqrt{K_{vent,2}(\rho_{v2}+\rho_{g2})}}. \tag{28}$$

Here, the dimensionless flow coefficient K (the loss factor) can be found in Schmidt *et. al.* [10] (see Tables 7-2 and 7-3 therein); a dimensionless relative valve position assumes values between $\lambda_k = 1$ (fully open) and $\lambda_k = 0$ (fully closed) [10]. During filling, the valve opens when the pressure exceeds 41.7 psia and closes when it falls below 38.7 psia. The vapor/gas and heat leaks may be defined as with the ST.

Nominal Dynamics

Fig. 3 summarizes the major results of the simulation of a nominal loading regime, which are based on the parameters, initial conditions, and filling protocol that are typical for LH2 loading systems. These values are provided in [5].

It can be seen that the LH2 level in the ST drops monotonically (Fig. 3a) as the level in the ET rises (Fig. 3d). The pressure p_{v1} in the ST (Fig. 3b) is determined by the loading dynamics and controlled by the vaporizer. Once achieved during slow fill, the pressure in the ST is maintained at approximately 80.7 psia up to the start of the reduced-pressure fast fill (see Figs. 3b and 3j), at which point it is maintained at 64.7 psia. Meanwhile, the ET ullage pressure (Fig. 3e) is oscillating due to the cycling of the vent valve that maintains the pressure between the lower and upper thresholds of 38.7 and 41.7 psia. The fluctuations in the ET ullage temperature (Fig. 3f) as well as in the mass flow rates (Figs. 3j and 3h) are driven by the ET pressure oscillations.

The LH2 partial pressure in the ET rises due to the continuing hydrogen supply, while the GHe partial pressure drops because the helium is being permanently removed through the vent valve (Fig. 3i). In this case, due to the condensation blocking effect [4], the flow of the condensed vapor in the ET (Fig. 3h) is much smaller than that in the ST (Fig. 3g), because the vapor pressure is being maintained approximately equal to the equilibrium pressure of the condensed vapor at the temperature of LH2. The ullage temperature T_{v2} in the ET (Fig. 3f) increases initially due to the introduction of the GHe during the pressurization stage, then drops from the initial high value due to venting and near-wall boiling that generates relatively cold GH2 during filling. The liquid surface temperature T_{f1} in the ST increases (Fig. 3c) due to the vapor condensation at the interface. Simultaneously, the ST ullage temperature T_{v1} increases, mainly because the relatively hot

(a) ST LH2 Volume   (b) ST Ullage Pressure   (c) ST Temperatures

(d) ET LH2 Volume   (e) ET Ullage Pressure   (f) ET Temperatures

(g) ST Condensation Flow   (h) ET Condensation Flow   (i) ET Partial Pressures

(j) Transfer Line Flow
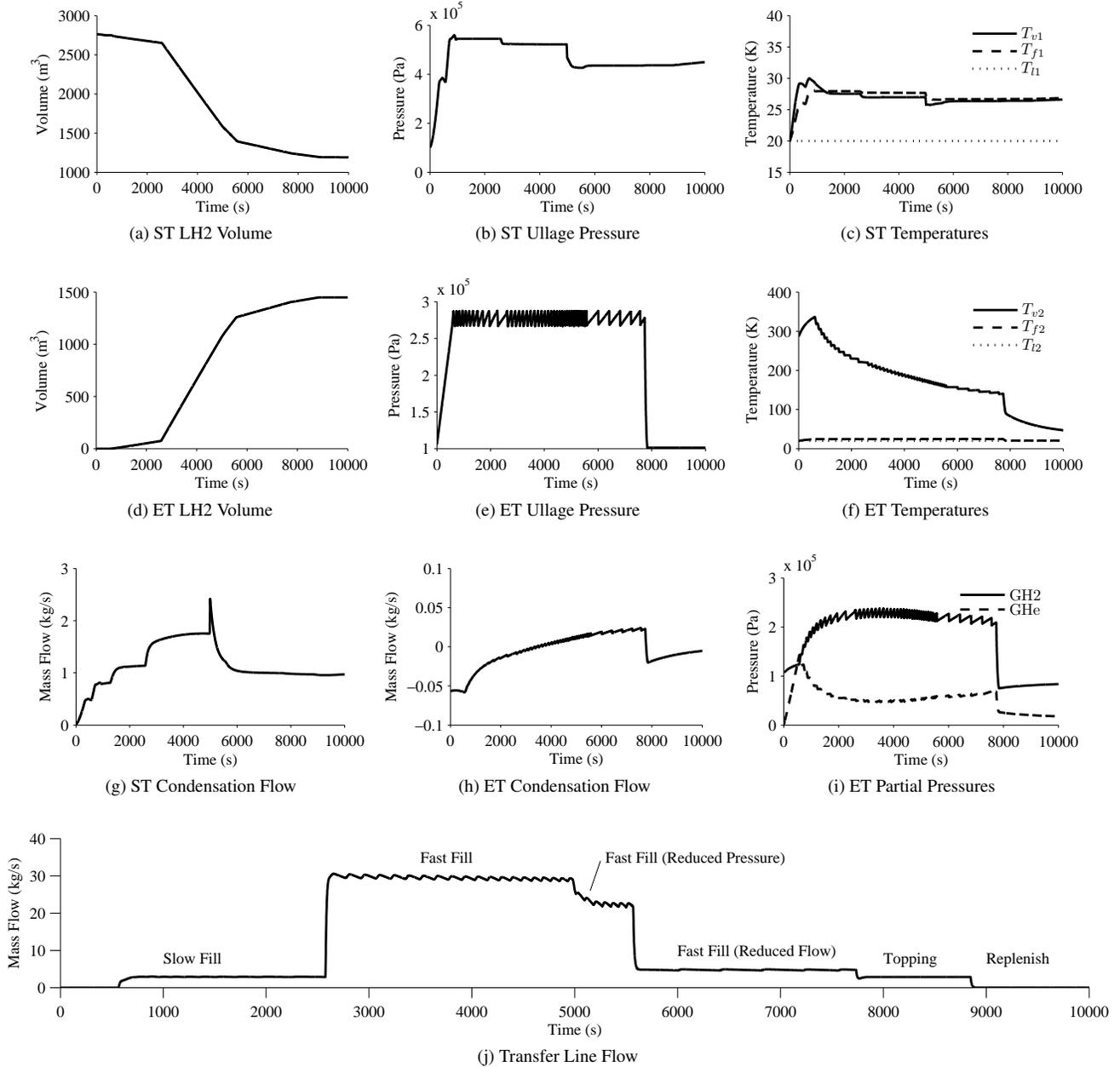
Figure 3. Nominal regime of LH2 propellant loading.

GH2 is supplied by the vaporizer as loading occurs. As a result, the ullage temperature approaches the temperature of LH2 saturated vapor at a pressure close to the final ST ullage pressure of approximately 5 atm (Fig. 3b).

3. DIAGNOSIS APPROACH

We apply a model-based diagnosis approach using the physics model of the LH2 system. In this paper, we consider the problem of single fault diagnosis. The system may be described in the following general form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t), \mathbf{v}(t))$$
$$\mathbf{y}(t) = \mathbf{h}(t, \mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{u}(t), \mathbf{n}(t)),$$

where $\mathbf{x}(t) \in \mathbb{R}^{n_x}$ is the state vector, $\boldsymbol{\theta}(t) \in \mathbb{R}^{n_\theta}$ is the parameter vector, $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ is the input vector, $\mathbf{v}(t) \in \mathbb{R}^{n_v}$ is the process noise vector, $\mathbf{f}$ is the state equation, $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ is the output vector, $\mathbf{n}(t) \in \mathbb{R}^{n_n}$ is the measurement noise vector, and $\mathbf{h}$ is the output equation.

Measurements are time-varying signals of $\mathbf{y(t)}$ obtained from the system sensors. In the LH2 system, we consider the following measurements for diagnosis: V_{l1}, p_1, T_{v1}, V_{l2}, p_2, T_{v2}, J_{tr}, λ_{vap}, and $\lambda_{vent,2}$.

We consider single, abrupt faults, modeled as unexpected step changes in system parameter values. We name faults by the
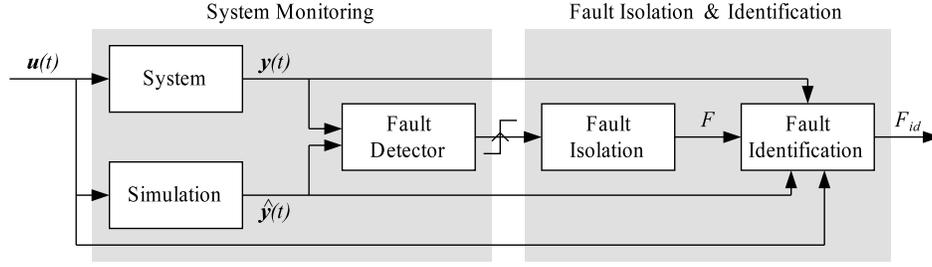
Figure 4. Diagnosis architecture.

associated parameter and the direction of change, i.e., θ^+ denotes a fault defined as an increase in the value of parameter θ, and θ^- denotes a fault defined as a decrease in the parameter value. For the LH2 system, we consider liquid and gas leaks, heat leaks, and valve clogging. The liquid and vapor leaks are defined by the equivalent leak areas, nominally zero, so faults are defined as increases in these areas, denoted by $A^+_{l,leak,1}$, $A^+_{v,leak,1}$, $A^+_{l,leak,2}$, and $A^+_{v,leak,2}$. The heat leaks are defined by the heat leak rate, nominally zero, so faults are increases in these values, denoted by $\dot{Q}^+_{leak,1}$ and $\dot{Q}^+_{leak,2}$. The valve clogging faults are described by the σ parameters, which are nominally 1, so faults are decreases in these values (0 at a minimum), denoted by σ^-_{tr}, σ^-_{vap}, and $\sigma^-_{vent,2}$.

The diagnosis architecture is shown in Fig. 4. The system receives inputs $\mathbf{u}(t)$ and produces outputs $\mathbf{y}(t)$. The physics simulation runs simultaneously, producing predicted outputs $\hat{\mathbf{y}}(t)$, given the inputs $\mathbf{u}(t)$. Using statistical methods, the fault detection module decides when a measurement has deviated from its nominal value, triggering fault isolation. Measurement deviations are then used to quickly isolate faults F. Fault identification computes, for each fault $f \in F$, the value of the fault parameter that best fits the outputs of the system, and the candidate with the lowest output error is selected as the best candidate.

4. FAULT DETECTION

In model-based fault detection, a model of the system provides reference outputs representing nominal system behavior. For each sensor output $y(t)$, we define the residual as $r(t) = y(t) - \hat{y}(t)$, where $\hat{y}(t)$ is the model-predicted output signal. Statistically significant deviations of the actual system outputs from the model-predicted outputs imply the presence of a fault. If the model is accurate, then fault detection thresholds can be small and faults detected quickly. The thresholds are dynamic in that they are defined with respect to nominal behaviors as a function of time. This is favorable to the current practice with such systems, where thresholds are static, preventing the detection of subtle deviations from nominal behavior that indicate faults.

We use the Z-test for robust fault detection using a set of sliding windows, as described in [11, 12]. The current mean of a

residual signal, $\mu_r(t)$, is estimated over a small window W_2:

$$\mu_r(t) = \frac{1}{W_2} \sum_{i=t-W_2+1}^{t} r(i).$$

The variance of the nominal residual signal, $\sigma_r^2(t)$, is computed using a large window W_1 preceding W_2, by a buffer W_{delay}, ensuring that W_1 does not contain any samples after fault occurrence. The variance is computed using

$$\sigma_r^2(t) = \frac{1}{W_1} \sum_{i=t-W_2-W_{delay}-W_1+1}^{t-W_2-W_{delay}} (r(i) - \mu_r'(t))^2,$$

where

$$\mu_r'(t) = \frac{1}{W_1} \sum_{i=t-W_2-W_{delay}-W_1+1}^{t-W_2-W_{delay}} r(i).$$

A given confidence level determines the bounds $z^- < 0$ and $z^+ > 0$ for a two-sided Z-test. The fault detection thresholds, $\varepsilon_r^-(t)$ and $\varepsilon_r^+(t)$, are dynamically computed using

$$\varepsilon_r^-(t) = z^- \frac{\sigma_r(t)}{\sqrt{W_2}} - E$$

$$\varepsilon_r^+(t) = z^+ \frac{\sigma_r(t)}{\sqrt{W_2}} + E,$$

where E is a modeling error term. A fault is detected if $\mu_r(t)$ lies outside of the thresholds at time t. If $\mu_r(t) < \varepsilon_r^-(t)$, a $-$ symbol for the measurement is used by the fault isolation module, and if $\mu_r(t) > \varepsilon_r^+(t)$, a + symbol is used. Generally, the parameters W_1, W_2, W_{delay}, the z bounds, and E must be tuned to optimize performance to minimize both false alarms and missed detections.

5. FAULT ISOLATION

We utilize a qualitative diagnosis methodology that isolates faults based on the transients they cause in system behavior, manifesting as deviations in observed measurement values from nominal measurement values [6]. The transients are abstracted using qualitative + (increase), $-$ (decrease), and 0 (no change) values to form *fault signatures*. Fault signatures represent these measurement deviations from nominal behavior as the immediate (discontinuous) change in magnitude, and the first nonzero derivative change.

6

The fault signatures can be derived automatically from a graph-based representation of the system model, known as a temporal causal graph (TCG) [6]. A TCG captures dynamic system variables as nodes in a graph, and the qualitative relations between them as edges. These edges are labeled with dt, representing integration, $+1$, denoting a proportionality, -1, denoting an inverse proportionality, and system parameters. Faults that are modeled as parameter changes appear on edges, allowing the qualitative effects of parameter changes to be propagated over the system variables. Propagation to measured variables reveals the qualitative effects of faults on measurements.

A partial TCG for the LH2 system, showing some key variables of the ST, is shown in Fig. 5. Variables within dashed boxes are measured. Here, we show only the σ_{vap} fault. A forward propagation algorithm may be used to derive the fault signatures. Details may be found in [6], so here, we use the vaporizer valve clogging fault as an example to illustrate the general procedure. According to the TCG, a decrease in σ_{vap} will lead to a decrease in J_{vap}, and, subsequently, a decrease in J_{v1}, and, due to the -1 label, an increase in J_{l1}. The decrease in J_{v1} will lead to a first-order decrease (due to the dt label) in m_{v1} and, subsequently, p_{v1}, which is a measured variable. This change then propagates as a decrease in J_{tr}, and then towards ET variables. The increase in J_{l1} will lead to a first-order increase in m_{l1} and V_{l1}, which is measured, leading to an increase in J_{tr}. This conflicts with the previously predicted decrease from the other path with the same derivative order, resulting in an ambiguity denoted with a $\star$. We can use the simulation and knowledge of the system dynamics to resolve these ambiguities. In this case, we know that the hydrostatic pressure due to the height of the liquid is negligible compared to p_{v1} due to the very low density of LH2. Therefore, we know that the decrease effect will dominate. The propagation continues to all system variables.

Fault signatures for the LH2 system are shown in Table 1. Due to the dynamics of the system, all faults appear as smooth changes in the measured values, so only the first change presents useful diagnostic information, hence, we show only the first symbol derived from forward propagation. Recall that the derived symbols represent deviations from nominal behavior, so, for example, the + symbol for V_{l1} caused by $A_{v,leak,1}^{+}$ does not necessarily mean that V_{l1} increases, rather, it increases with respect to its nominal value, i.e., it decreases at a slower rate. The remaining $\star$ symbols are for those in which the qualitative effect may be different depending on the state of the system. Note also that this is a hybrid system, i.e., it has continuous dynamics mixed with discrete dynamics due to the valves, and we show only the signatures for the mode where the vent valve is closed. When the vent valve opens, some of the signatures will flip. A systematic framework for dealing with this issue is discussed in detail in [13].

Using the signatures, we can perform diagnosability analysis to determine the effectiveness of the isolation step. We
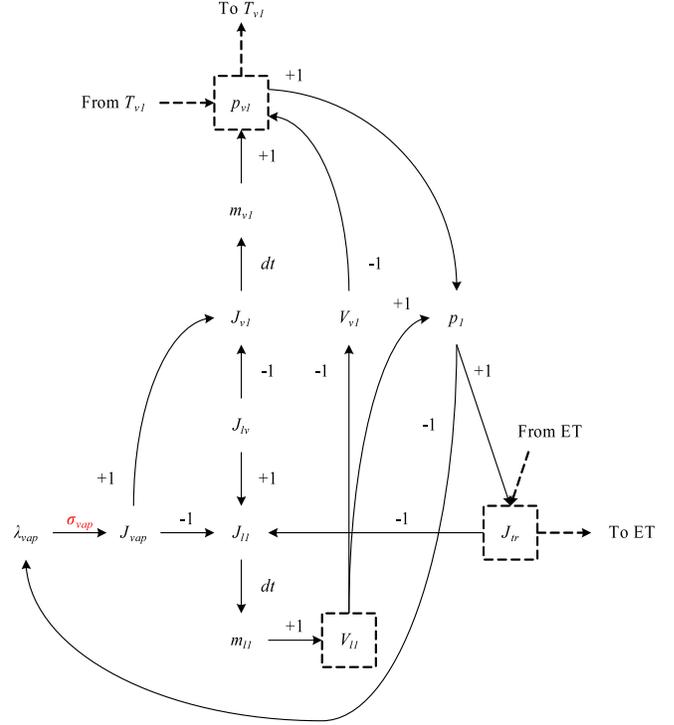
Figure 5. Partial TCG of the LH2 system.

Table 1. Fault Signatures for the LH2 System

Fault	V_{l1}	p_1	T_{v1}	V_{l2}	p_2	T_{v2}	J_{tr}	λ_{vap}	$\lambda_{vent,2}$
$A_{l,leak,1}^{+}$	−	−	⋆	−	−	⋆	−	+	−
$A_{v,leak,1}^{+}$	+	−	−	−	−	⋆	−	+	−
$\dot{Q}_{leak,1}^{+}$	−	+	+	+	+	⋆	+	−	+
$A_{l,leak,2}^{+}$	−	−	⋆	−	−	⋆	+	+	−
$A_{v,leak,2}^{+}$	−	−	⋆	+	−	−	+	+	−
$\dot{Q}_{leak,2}^{+}$	+	+	⋆	−	+	+	−	−	+
σ_{tr}^{-}	+	+	⋆	−	−	−	−	−	−
σ_{vap}^{-}	+	−	−	−	−	⋆	−	+	−
$\sigma_{vent,2}^{-}$	+	+	⋆	−	+	+	−	−	+

can see that in most cases, the set of signatures produced distinguishes most faults. One exception is the pair $A_{v,leak,1}^{+}$ and σ_{vap}^{-}, which have all the same signatures, therefore, fault identification will have to distinguish between them, since quantitatively, they should have different effects on the measured variables. If we measure J_{vap}, then the faults can be distinguished, because σ_{vap}^{-} would produce a decrease in J_{vap}, whereas $A_{v,leak,1}^{+}$ would produce an increase. The faults $\dot{Q}_{leak,2}^{+}$ and $\sigma_{vent,2}^{-}$ are similarly indistinguishable.

We can also use the signatures to perform measurement selection. The measurement of λ_{vap} always has the same signature as p_1, therefore one of these can be eliminated without loss of diagnosability. However, both quantities are relatively easy to measure, so there is value in keeping them both. λ_{vap} is typically more sensitive to faults than p_1, since p_1 is controlled by the vaporizer, which may mask faults. Simi-

larly, both p_2 and $\lambda_{vent,2}$ are not needed. Changes in p_2 are reflected in changes in the switching frequency of the vent valve, and therefore provide the same information, although changes in the vent valve position can be detected much more easily, again because the control of p_2 leads to fault masking in that measurement. It is also beneficial to keep as many measurements as possible, because this gives more information for fault identification. Further, the qualitative fault signatures say nothing about fault magnitude. Some faults may not produce large enough changes in some variables for the changes to be detected, therefore, having more measurements helps to alleviate this problem.

6. FAULT IDENTIFICATION

Fault identification is initiated immediately after the initial set of fault candidates is produced after fault detection. Each candidate has its own identification module that updates its estimate at every time step. Identification is performed using the particle filtering algorithm for joint state-parameter estimation [14, 15], which has seen previous application in model-based diagnosis and prognosis algorithms [16–19]. The magnitude of the fault parameter θ is estimated along with the system state.

The identification module must compute $p(\mathbf{x}_k, \boldsymbol{\theta}_k | \mathbf{y}_{0:k})$. A general solution to this problem is the *particle filter*, which may be directly applied to nonlinear systems with non-Gaussian noise terms. Particle filters offer approximate (suboptimal) solutions to the state estimation problem for systems where optimal solutions are unavailable or intractable [14, 15]. In particle filters, the state distribution is approximated by a set of discrete weighted samples, called *particles*. As the number of particles is increased, performance increases and the optimal solution is approached. Due to the highly nonlinear dynamics of the LH2 loading system, particle filters are favored over other estimation approaches such as the extended Kalman filter.

The particle approximation to the state distribution is given by

$$\{(\mathbf{x}_k^i, \boldsymbol{\theta}_k^i), w_k^i\}_{i=1}^N,$$

where N denotes the number of particles, and for particle i, $\mathbf{x}_k^i$ denotes the state vector estimate, $\boldsymbol{\theta}_k^i$ denotes the parameter vector estimate, and w_k^i denotes the weight. The posterior density is approximated by

$$p(\mathbf{x}_k, \boldsymbol{\theta}_k | \mathbf{y}_{0:k}) \approx \sum_{i=1}^N w_k^i \delta_{(\mathbf{x}_k^i, \boldsymbol{\theta}_k^i)}(d\mathbf{x}_k d\boldsymbol{\theta}_k),$$

where $\delta_{(\mathbf{x}_k^i, \boldsymbol{\theta}_k^i)}(d\mathbf{x}_k d\boldsymbol{\theta}_k)$ denotes the Dirac delta function located at $(\mathbf{x}_k^i, \boldsymbol{\theta}_k^i)$.

We use the sampling importance resampling (SIR) particle filter with systematic resampling [14, 20]. The pseudocode for a step of the filter is shown as Algorithm 1. Each particle

---

Algorithm 1 SIR Filter

Inputs: $\{(\mathbf{x}_{k-1}^i, \boldsymbol{\theta}_{k-1}^i), w_{k-1}^i\}_{i=1}^N, \mathbf{u}_{k-1:k}, \mathbf{y}_k$
Outputs: $\{(\mathbf{x}_k^i, \boldsymbol{\theta}_k^i), w_k^i\}_{i=1}^N$
for $i = 1$ **to** N **do**
  $\boldsymbol{\theta}_k^i \sim p(\boldsymbol{\theta}_k | \boldsymbol{\theta}_{k-1}^i)$
  $\mathbf{x}_k^i \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}^i, \boldsymbol{\theta}_{k-1}^i, \mathbf{u}_{k-1})$
  $w_k^i \leftarrow p(\mathbf{y}_k | \mathbf{x}_k^i, \boldsymbol{\theta}_k^i, \mathbf{u}_k)$
end for
$W \leftarrow \sum_{i=1}^N w_k^i$
for $i = 1$ **to** N **do**
  $w_k^i \leftarrow w_k^i / W$
end for
$\{(\mathbf{x}_k^i, \boldsymbol{\theta}_k^i), w_k^i\}_{i=1}^N \leftarrow \texttt{Resample}(\{(\mathbf{x}_k^i, \boldsymbol{\theta}_k^i), w_k^i\}_{i=1}^N)$

---

is propagated forward to time k by sampling new parameter and state values. The particle weight is assigned using $\mathbf{y}_k$. The weights are then normalized, and then the particles are resampled (see [14]).

Note that the parameters $\boldsymbol{\theta}_k$ evolve by some unknown process that is independent of the state $\mathbf{x}_k$. The particle filter algorithm requires some type of evolution to the parameters. We use a random walk, i.e., for parameter θ, $\theta_k = \theta_{k-1} + \xi_{k-1}$, where ξ_{k-1} is Gaussian noise. The particles generated with parameter values closest to the true values should match the outputs better, and, therefore, be assigned higher weight, thus allowing the particle filter to converge to the true values. The selected variance of the random walk noise affects both the rate of this convergence and the estimation performance after convergence. Using the simulation, we can determine appropriate values for the random walk variances.

Under the single fault assumption, we can run a set of parallel particle filters, one for each consistent fault candidate. This reduces the dimensionality of the estimation task over combined estimation of all fault parameters, and allows the particle filters to be much more efficient. The particle filters are initiated at the point of fault detection, using the model-estimated state at that time point for initialization of the states. The fault parameter estimate starts at its value during nominal operation (e.g., 0 for leak areas). When the fault isolation module reduces the set of candidates F, the identification task continues only for those faults remaining in F.

Because diagnosability may be limited, fault identification must also be used to help refine fault candidates. The particle filter for the true fault candidate should estimate the correct fault parameter and track the faulty outputs with low error, whereas the particle filters for the incorrect faults will not track and result in large error. We compute the mean squared output error for each candidate from the point of fault detection to the present time. The candidate with the lowest output error e is considered to be the true candidate.

7. RESULTS

We illustrate the diagnosis process with the $A^+_{v,leak,1}$ fault injected at 2600 s with magnitude 3×10^{-3} m^2. Note that from diagnosability analysis, the observed signatures will be consistent with both the ST vapor leak fault $A^+_{v,leak,1}$ and the vaporizer valve clogging fault σ^-_{vap}, so the fault identification stage will have to resolve the ambiguity. Relevant measurements are shown in Fig. 6. The diagnoser, running at a sample time of once per second, detects the fault at 2605 s due to an observed decrease in p_1. The initial list of consistent faults is $\{A^+_{v,leak,1}, A^+_{v,leak,2}, A^+_{l,leak,1}, A^+_{l,leak,2}, \sigma^-_{vap}\}$. At 2608 s, a decrease in J_{tr} is detected, followed by an increase in λ_{vap} at 2612 s and a decrease in T_{v1} at 2690 s. The candidate set remains the same. At 2703 s, a decrease in V_{l2} is detected, which rules out the ET vapor/gas leak fault $A^+_{v,leak,2}$. At this point, the output error of the $A^+_{v,leak,1}$ fault is an order of magnitude less than for all other remaining candidates. At 2806 s, a detected increase in V_{l1} rules out $A^+_{l,leak,1}$ and $A^+_{l,leak,2}$, resulting in the candidate list $\{A^+_{v,leak,1}, \sigma^-_{vap}\}$. By this time, the particle filter for $A^+_{v,leak,1}$ has also converged on the correct value of the fault magnitude, as shown in Fig. 7. The fault was detected within 5 seconds of its occurrence, and identification confirmed the true fault with correct magnitude within 200 s. It is important to quickly discriminate between a vapor leak in the ST and a clogging of the vaporizer, because the former requires an abort, whereas the system can continue fueling in a safe manner with the latter. As discussed in Section 5, measuring J_{vap} would allow a faster discrimination of these two faults.

Diagnosis results over the complete set of faults are shown in Table 2. Here, Δt_d denotes the time to detect the fault, Δt_i denotes the time to isolate, taken as the last time at which the candidate set is reduced, F_{id} denotes the output of the fault identification module at the end of the scenario, and f^* denotes the final output of the diagnoser, i.e., the identified fault with the lowest error. For the fault identification stage, 50 particles were used per particle filter, which seemed to offer a reasonable trade off between computation time and identification accuracy.

Overall, the results are fairly good. In all cases, the correct fault was identified with sufficient accuracy. In some cases, the effects of the faults are very subtle at first, and it takes some time before the changes they produce can be distinguished from the sensor noise. For the vent valve clogging fault, the effects are visible only when the valve is open, so if the fault first appears when the valve is closed, it will take time for the fault to be detected. The σ^-_{tr} fault was detected immediately because the change in J_{tr} was significant. Without this measurement, it would take some time for the fault to be visible in the liquid volumes of the tanks.

For some faults, the operation of the vaporizer masks the change in ST pressure, therefore, it is important to monitor the vaporizer valve position in order to obtain a more sensitive
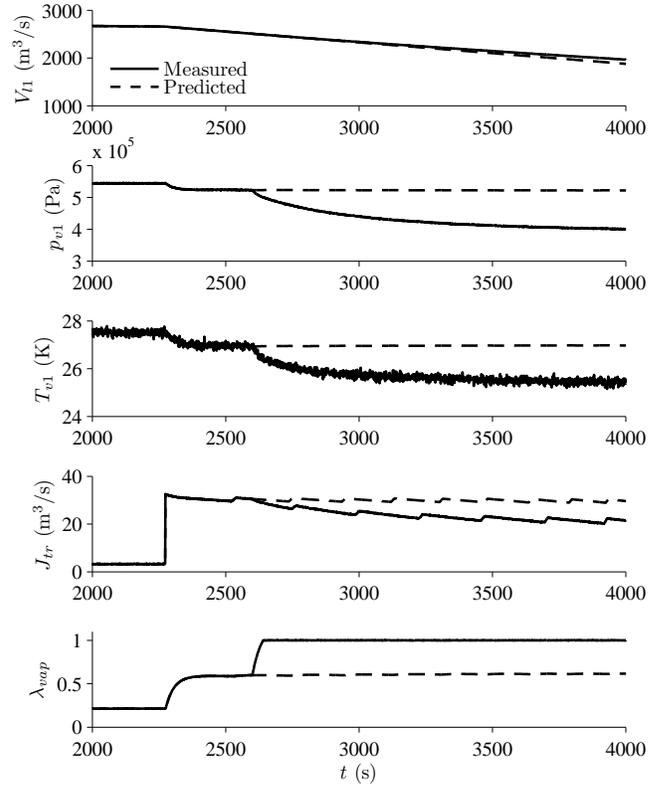
Figure 6. Predicted and observed outputs for the $A^+_{v,leak,1}$ fault injected at 2600 s with magnitude 3×10^{-3} m^2.
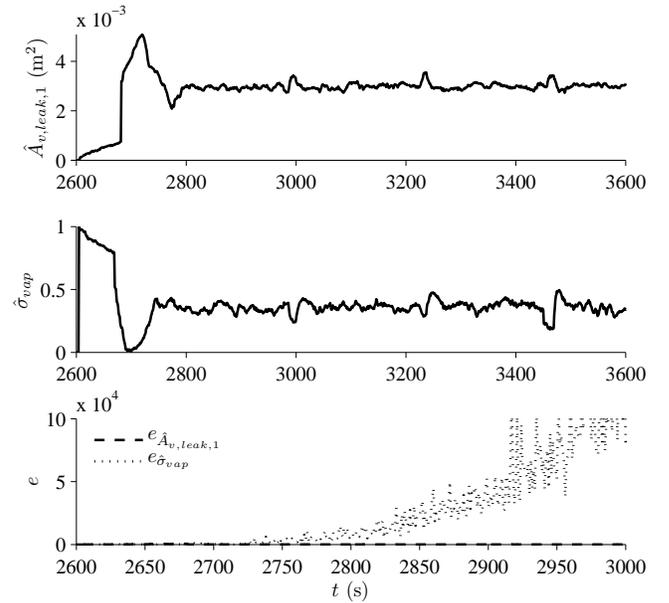
Figure 7. Estimated fault magnitudes and output error for the $A^+_{v,leak,1}$ fault injected at 2600 s with magnitude 3×10^{-3} m^2. The $A^+_{v,leak,1}$ estimator converges to the true value by 2800 s. The σ_{vap} estimator converges by 2800 s also, but it cannot track the outputs, as shown by the corresponding error $e_{\hat{\sigma}_{vap}}$, which is greater than $e_{\hat{A}_{v,leak,1}}$ at 2800 s.

Table 2. Diagnosis Results

Fault	Magnitude	Δt_d	Δt_i	F_{id}	f^*
$A_{l,leak,1}^+$	1×10^{-3} m^2	17	113	$A_{l,leak,1}^+ = 9.98 \times 10^{-4}, e = 97.1$	$A_{l,leak,1}^+ = 9.98 \times 10^{-4}$
$A_{v,leak,1}^+$	3×10^{-3} m^2	5	206	$A_{v,leak,1}^+ = 3.04 \times 10^{-4}, e = 27.1$	$A_{v,leak,1}^+ = 3.04 \times 10^{-4}$
				$\sigma_{vap}^- = 0.35, e = 7.89 \times 10^5$	
$\dot{Q}_{leak,1}^+$	3×10^4 W	111	151	$\dot{Q}_{leak,1}^+ = 2.95 \times 10^4, e = 9.44$	$\dot{Q}_{leak,1}^+ = 2.95 \times 10^4$
$A_{l,leak,2}^+$	1×10^{-3} m^2	28	131	$A_{l,leak,2}^+ = 1.10 \times 10^{-3}, e = 121$	$A_{l,leak,2}^+ = 1.10 \times 10^{-3}$
$A_{v,leak,2}^+$	1×10^{-4} m^2	39	138	$A_{v,leak,1}^+ = 1.30 \times 10^{-4}, e = 409$	$A_{v,leak,2}^+ = 1.07 \times 10^{-4}$
				$A_{v,leak,2}^+ = 1.07 \times 10^{-4}, e = 7.50$	
				$A_{l,leak,1}^+ = 6.78 \times 10^{-5}, e = 463$	
				$A_{l,leak,2}^+ = 5.16 \times 10^{-4}, e = 2.86 \times 10^5$	
				$\sigma_{vap}^- = 0.95, e = 410$	
$\dot{Q}_{leak,2}^+$	3×10^4 W	18	131	$\dot{Q}_{leak,1}^+ = 4.15 \times 10^3, e = 1930$	$\dot{Q}_{leak,2}^+ = 3.19 \times 10^4$
				$\dot{Q}_{leak,2}^+ = 3.19 \times 10^4, e = 8.97$	
σ_{tr}^-	0.5	0	13	$\sigma_{tr}^- = 0.50, e = 56.8$	$\sigma_{tr}^- = 0.50$
σ_{vap}^-	0.5	7	71	$A_{v,leak,1}^+ = 2.14 \times 10^{-3}, e = 3.47 \times 10^5$	$\sigma_{vap}^- = 0.50$
				$\sigma_{vap}^- = 0.50, e = 8.06$	
$\sigma_{vent,1}^-$	0.5	141	152	$\dot{Q}_{leak,2}^+ = 3.21 \times 10^3, e = 105$	$\sigma_{vent,1}^- = 0.50$
				$\sigma_{vent,1}^- = 0.50, e = 9.55$	

detection. Changes in the vent valve switching frequency are also easy to observe, but one must wait for the upper pressure threshold to be reached before it can be determined whether it closes early or late. We also observed that faults in the ET have very small effects on measured variables in the ST, resulting in the large size of F_{id} for $A_{v,leak,2}^+$. In such cases, fault identification becomes even more crucial to resolving ambiguities.

8. CONCLUSIONS

In this paper, by applying our previously developed physics-based model of a propellant loading system [5], we analyzed the effects of faults, and applied a model-based diagnosis approach to fault detection, isolation, and identification. The detection stage compares observed and model-predicted outputs to detect faults. The isolation stage compares model-predicted fault transients to observed measurement deviations to quickly isolate faults and reduce the complexity of the identification stage. The fault identification stage uses particle filters to estimate the values of fault parameters and resolve isolation ambiguities.

The simulation results showed that with such a model-based diagnosis approach, faults can be quickly detected, isolated, and identified, and this knowledge can be used to determine if the system can continue loading safely or if an abort is required. We considered only single faults here, and in the future we will extend the framework to multiple faults. Sensor faults were not considered here, but can be easily incorporated into a model-based framework by including models of the sensors [12]. Using the physics model for prognostics and loading optimization is also of interest.

We would like to validate the approach by using available historical data of faulty situations. Such data can be difficult to find, and, further, there is only a limited capability to inject faults into the system for the purpose of diagnosis algorithm validation, due to the high cost of a fueling operation and the explicit danger that fault injection presents in such a system. Therefore, an accurate simulation with which to validate diagnosis algorithms, such as that used here, is very valuable.

REFERENCES

[1] P. N. Estey, D. H. Lewis Jr., and M. Connor, "Prediction of a propellant tank pressure history using state space methods," *Journal of Spacecraft and Rockets*, vol. 20, no. 1, pp. 49–54, 1983.

[2] C. N. Torre, J. A. Witham, E. A. Dennison, R. C. McCool, and M. W. Rinker, "Analysis of a low-vapor-pressure cryogenic propellant tankage system," *J. Spacecraft*, vol. 26, pp. 368–378, 1987.

[3] D. M. DeFelice and J. C. Aydelott, "Thermodynamic analysis and subscale modeling of space-based cryogenic propellant systems," in *Proceedings of the 23rd Joint Propulsion Conference*, Jun. 1987.

[4] V. V. Osipov and C. B. Muratov, "Dynamic condensation blocking in cryogenic refueling," *Applied Physics Letters*, vol. 93, no. 22, pp. 224 105–1–4, 2008.

[5] V. Osipov, C. Muratov, M. Daigle, M. Foygel, V. Smelyanskiy, and A. Patterson-Hine, "A dynamical physics model of nominal and faulty operational modes of propellant loading (liquid hydrogen): From space shuttle to future missions," NASA Ames Research Center, Tech. Rep. NASA/TM-2010-216394, Jul. 2010.

[6] P. J. Mosterman and G. Biswas, "Diagnosis of continuous valued systems in transient operating regions," *IEEE Trans. on Systems, Man and Cybernetics, Part A*, vol. 29, no. 6, pp. 554–565, 1999.

[7] J. A. Clark, "Universal equations for saturation vapor pressure," in *Proceedings of the 40th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, Jul. 2004.

[8] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed. Pergamon Press, 1987.

[9] J. A. Dean, *Lange's Handbook of Chemistry*, 12th ed. McGraw-Hill, New York, 1979.

[10] F. W. Schmidt, R. E. Henderson, and C. H. Wolgemuth, *Introduction to Thermal Sciences*, 2nd ed. J. Wiley & Sons, 1993.

[11] G. Biswas, G. Simon, N. Mahadevan, S. Narasimhan, J. Ramirez, and G. Karsai, "A robust method for hybrid diagnosis of complex systems," in *Proc. of the 5th Symposium on Fault Detection, Supervision and Safety for Technical Processes*, Jun. 2003, pp. 1125–1131.

[12] M. Daigle, I. Roychoudhury, G. Biswas, X. Koutsoukos, A. Patterson-Hine, , and S. Poll, "A comprehensive diagnosis methodology for complex hybrid systems: A case study on spacecraft power distribution systems," *IEEE Transactions of Systems, Man, and Cybernetics, Part A*, vol. 4, no. 5, pp. 917–931, Sep. 2010.

[13] S. Narasimhan and G. Biswas, "Model-based diagnosis of hybrid systems," *IEEE Trans. on Systems, Man and Cybernetics, Part A*, vol. 37, no. 3, pp. 348–361, May 2007.

[14] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[15] O. Cappe, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proceedings of the IEEE*, vol. 95, no. 5, p. 899, 2007.

[16] I. Roychoudhury, G. Biswas, and X. Koutsoukos, "Distributed diagnosis of dynamic systems using dynamic bayesian networks," in *Proceedings of the 20th International Workshop on Principles of Diagnosis*, Jun. 2009, pp. 329–336.

[17] M. E. Orchard, "A particle filtering-based framework for on-line fault diagnosis and failure prognosis," Ph.D. dissertation, Georgia Institute of Technology, 2007.

[18] B. Saha and K. Goebel, "Modeling Li-ion battery capacity depletion in a particle filtering framework," in *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2009*, Sep. 2009.

[19] M. Daigle and K. Goebel, "Model-based prognostics under limited sensing," in *Proceedings of the 2010 IEEE Aerospace Conference*, Mar. 2010.

[20] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.