



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

MANEJO DE INCERTIDUMBRE ANTE PÉRDIDA PARCIAL O TOTAL DE DATOS
EN ALGORITMOS BASADOS EN MÉTODOS SECUENCIALES DE MONTE CARLO
Y NUEVA DEFINICIÓN DE PROBABILIDAD DE FALLA EN EL CONTEXTO DE
MONITOREO EN LÍNEA

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA
INGENIERÍA, MENCIÓN ELÉCTRICA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

DAVID ESTEBAN ACUÑA URETA

PROFESOR GUÍA:
MARCOS ORCHARD CONCHA

MIEMBROS DE LA COMISIÓN:
JORGE SILVA SÁNCHEZ
JUAN YUZ EISSMANN

SANTIAGO DE CHILE
2016

RESUMEN DE LA TESIS PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELÉCTRICO Y AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA
POR: DAVID ESTEBAN ACUÑA URETA
FECHA: 2016
PROF. GUÍA: SR. MARCOS ORCHARD CONCHA

MANEJO DE INCERTIDUMBRE ANTE PÉRDIDA PARCIAL O TOTAL DE DATOS EN ALGORITMOS BASADOS EN MÉTODOS SECUENCIALES DE MONTE CARLO Y NUEVA DEFINICIÓN DE PROBABILIDAD DE FALLA EN EL CONTEXTO DE MONITOREO EN LÍNEA

Desde hace algunos años que la disciplina especializada en el diagnóstico de fallas en sistemas y el pronóstico de eventos que pudieran afectar la salud de éstos, denominada *Prognostics and Health Management* (PHM), ha ido tomando cada vez más fuerza y ha empezado a consolidarse como una nueva área de la ingeniería. Esta disciplina tiene como eje principal la caracterización del estado de salud de los sistemas dinámicos, ya sea a través de enfoques probabilísticos, posibilísticos, o epistémicos, por nombrar algunos; siendo el enfoque probabilístico el adoptado en la presente Tesis.

Dado que la mayoría de los sistemas se conciben mediante complejas relaciones que suelen establecerse en base a no linealidades e incertidumbre no necesariamente Gaussiana, el empleo de métodos secuenciales de Monte Carlo -también llamados *filtros de partículas*- ha tomado gran relevancia para el diagnóstico de la salud de los sistemas. Los filtros de partículas proveen una distribución empírica de probabilidad para los estados del sistema siguiendo un enfoque Bayesiano.

Es muy frecuente que en rutinas de estimación, y en especial al usar filtro de partículas, algunas de las mediciones que debiesen proporcionar los sensores se pierdan de manera parcial o total. En el caso de una pérdida parcial, se propone una estrategia basada en la Teoría de Imputaciones Múltiples desarrollada en los años 60' que permite incorporar la incertidumbre que hay al perder solo parte de la información. Por otra parte, cuando se pierde acceso a la totalidad de las mediciones se puede hablar de un problema análogo al de pronóstico. Se propone un algoritmo que recibe como entrada una distribución empírica producto del final de una etapa de estimación basada en filtro de partículas, en la cual se propaga incertidumbre haciendo uso de sigma-points y del ajuste paramétrico de mezclas de Gaussianas para la aproximación de distribuciones de probabilidad.

Además de los algoritmos anteriormente mencionados, en esta Tesis se presenta la deducción y demostración de una nueva medida de probabilidad de falla de sistemas en el contexto de monitoreo en línea, la cual modifica el estado-del-arte y abre nuevas ramas de investigación que pueden influir fuertemente en el desarrollo de nuevos sistemas de toma de decisiones.

Los desarrollos presentados en esta Tesis abordan principalmente el caso del problema de estimación y pronóstico del Estado-de-Carga de baterías de ion-litio para el análisis y validación de los algoritmos que se proponen dada su relevancia en la autonomía de los dispositivos electrónicos que requieren procesamiento de información en tiempo real; tales como tablets, celulares, vehículos o micro-redes, por nombrar algunos ejemplos.

Para mi familia, expresión maravillosa del amor de Dios en mi vida

Agradecimientos

Primeramente quisiera agradecer a Dios por todos los dones, todos los medios, y todas las personas que ha puesto en mi camino hasta ahora y que han contribuido en especial durante mis estudios universitarios. Mi vida no tiene sentido sin Él.

Quisiera agradecer profundamente a mi familia, por su amor incondicional. A mi padre, por su incansable cariño y mansedumbre, además de instruirme en ser un hombre de una sola línea; a mi madre, por toda la paciencia, ternura y delicadeza con la que suele tratarme; mi hermano Ángel, que siempre sabe escucharme y orientarme con maestría con una sabiduría que le ha sido otorgada desde lo alto; a mi hermano Daniel, que con gran madurez me enseña innumerables cosas, desde la profundidad del discernimiento de espíritu hasta la dulzura del canto; y finalmente pero no menos importante, a mi queridísima abuelita Carmen, que vela por mí en la oración todos los días de su vida.

Quisiera agradecer también muy especialmente al profesor Marcos Orchard, cuya vocación académica fue la fuente inspiración para profundizar en las ciencias de la ingeniería, además por la confianza que puso en mí desde un principio y a lo largo de aproximadamente tres años donde trabajamos juntos, planteándome desafíos, orientándome, y creyendo firmemente en mis aptitudes. Así también, agradecer al profesor Jorge Silva, quien fue fuente de inspiración para atreverme a escudriñar en aspectos más teóricos. También quiero agradecer a todos mis amigos y compañeros que me han acompañado durante éstos años de estudio; a Michel, Tomislav, Antonio, Leonel, Cristián, Esteban, Felipe. También a quienes son o han sido parte del grupo de investigación donde participo; a Benjamín, Matías, Daniel, Ricardo, Hugo, Pablo, Francisco, Aramis, Juan Pablo, Jorge, Ismael, Rubén, Vanessa, César y Gina.

Aprovecho el espacio para agradecer al profesor Jonathan Jiménez, por su amistad y porque tuve el privilegio de participar de sus clases de TaeKwonDo durante aproximadamente siete años y medio en pleno periodo universitario.

Muchas gracias a todos, tanto a los nombrados como a tantos otros que han contribuido no sólo en lo pragmático, sino simplemente con ser como son y por haber compartido conmigo.

Tabla de Contenido

Introducción	1
1. Fundamentos Generales	2
1.1. Baterías de Ion-Litio	3
1.1.1. Representación en Espacio-Estados	4
1.2. Filtro de Partículas	5
1.2.1. Sequential Importance Sampling	6
1.2.2. Resampling	7
1.3. Imputaciones Múltiples	9
1.3.1. Filtro de Partículas con Imputaciones Múltiples	9
1.4. Outer Feedback Correction Loops	10
1.5. Unscented Transform y Sigma-Points	11
1.6. Weighted Expectation Maximization para GMMs	12
2. Filtro de Partículas con Imputaciones Múltiples	14
2.1. Implementación de una Estrategia de Imputaciones Múltiples	15
2.2. Resultados Experimentales	17
2.2.1. Impacto en la Etapa de Estimación	18
2.2.2. Impacto en la Etapa de Pronóstico	22
3. Nueva Definición de Probabilidad de Falla en el Contexto de Monitoreo en Línea	25
3.1. Formalización Matemática	26
4. Pronóstico de Fallas Basado en Filtro de Partículas y el Uso de Sigma-Points	31
4.1. Propagación de Incertidumbre Mediante Sigma-Points	33
4.2. Comentarios Sobre el Algoritmo Propuesto	36
4.2.1. En Cuanto a los Requerimientos	36
4.2.2. Convergencia del Algoritmo WEM	37
4.3. Análisis del Algoritmo: Estimación de EoD	38
4.3.1. Métricas de Evaluación de Desempeño	42
4.3.2. Resultados Experimentales	42
Conclusión	48
Bibliografía	51

Índice de Ilustraciones

2.1. El modelo implementado no corresponde a un HMM, pues la transición de los estados depende de la salida pasada del sistema ($v(k - 1)$ en éste caso). . . .	16
2.2. Módulo de la impedancia interna como función del SoC [%] para una desconexión de 1000 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 1000-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$	19
2.3. Estimación de SoC como función del tiempo[s] para una desconexión de 1000 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 1000-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$	20
2.4. Estimación de voltaje como función del SoC [%] para una desconexión de 1000 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 1000-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$	20

2.5.	Estimación y pronóstico del módulo de la impedancia interna como función del SoC[%] para una desconexión de 500 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 500-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$	22
2.6.	Estimación y pronóstico del SoC como función del tiempo[s] para una desconexión de 500 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 500-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$	23
2.7.	Estimación y pronóstico de voltaje como función del SoC[%] para una desconexión de 500 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 500-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$	24
4.1.	Ejemplo de rotación de sigma-points para el caso de dos dimensiones, [14].	38
4.2.	Curva de OCV de una celda de ion-litio (línea negra) y proyección de zona lineal (línea segmentada transversal a las zonas) como función del SoC, [41, 42].	39
4.3.	El modelo corresponde a un HMM, pues la transición de los estados sólo depende del instante anterior y deben ser inferidos indirectamente a través de mediciones.	40
4.4.	Curvas promedio de PMFs para el EoD producto de 30 realizaciones considerando una población de 45 partículas y el ajuste de GMMs de 5 componentes generadas a partir de la solución analítica aproximada (izquierda) y del algoritmo basado en sigma-points (derecha).	45
4.5.	Curvas promedio de PMFs para el EoD producto de 30 realizaciones considerando una población de 90 partículas y el ajuste de GMMs de 10 componentes generadas a partir de la solución analítica aproximada (izquierda) y del algoritmo basado en sigma-points (derecha).	46
4.6.	Curvas promedio de PMFs para el EoD producto de 30 realizaciones considerando una población de 135 partículas y el ajuste de GMMs de 15 componentes generadas a partir de la solución analítica aproximada (izquierda) y del algoritmo basado en sigma-points (derecha).	47

Introducción

En la implementación de algoritmos basados en filtro de partículas, para propósitos de estimación de estados de un sistema, a menudo se debe lidiar con el problema de mediciones perdidas. Un diseño eficaz requiere una metodología apropiada para caracterizar la incertidumbre inherente a ésta problemática a lo largo del proceso de estimación, incorporando conocimiento desde fuentes alternativas de información. El primero de los trabajos presentados en esta Tesis analiza éste problema y presenta resultados de una estrategia de imputaciones múltiples que mejora el desempeño de estimadores de Estado-de-Carga (SoC, del inglés *State-of-Charge*) basados en filtro de partículas para el caso baterías de ion-litio. El esquema de caracterización de incertidumbre es probado y validado en un caso de estudio donde se modela el SoC de una batería en una representación en el espacio-estados que requiere mediciones tanto de voltaje como de corriente de descarga para la estimación. Se asume una desconexión súbita del sensor de voltaje de la batería que causa una pérdida de datos significativa. Los resultados muestran que el empleo de imputaciones múltiples junto con filtro de partículas permite obtener una caracterización de los intervalos de confianza razonable para los estimadores de los estados, incluso cuando la desconexión del sensor de voltaje es persistente. Más aún, si las mediciones de voltaje vuelven a estar disponibles, los intervalos de confianza se ajustan a niveles que son comparables al caso donde no hay pérdida de datos. Debido a que los estimadores de los estados son usados como condición inicial para módulos de pronóstico del tiempo de descarga (EoD, del inglés *End-of-Discharge*) de baterías, también se estudia cómo impacta la estrategia de imputaciones múltiples en la calidad de los estimadores del EoD.

Cuando se tiene una pérdida total de las mediciones, ésta condición se convierte en un problema análogo al problema de pronóstico. Además del enfoque de imputaciones múltiples para el caso de pérdida parcial de datos, la Tesis incluye otras dos temáticas que se enmarcan en el concepto de pronóstico, todo bajo una perspectiva probabilista. Se propone un algoritmo para la realización de pronóstico para sistemas en general que considera una representación en el espacio-estados y dinámica no lineal con ruidos no necesariamente Gaussianos propagando incertidumbre mediante el uso de sigma-points. Se establecen ventajas y desventajas del algoritmo de pronóstico propuesto. Por otra parte, se incluye en la Tesis una nueva definición de probabilidad de falla que en el contexto de monitoreo en línea. La nueva definición se muestra como la consecuencia de una formalización matemática probabilista, en donde se exponen tanto la deducción de ésta como también la demostración del cumplimiento de las propiedades de una medida de probabilidad.

Capítulo 1

Fundamentos Generales

Los sistemas del mundo real son comúnmente dinámicos, no lineales, y pueden involucrar relaciones de alta dimensionalidad entre variables. Los modelos planteados en espacio-estados ofrecen por su parte un buen fundamento teórico sobre el cual trabajar en módulos de diagnóstico o pronóstico dedicados a monitorear componentes críticos de los sistemas, cuyas fenomenologías se rigen por principios de la física. Los módulos de diagnóstico, por un lado, están orientados a dar una noción de la salud del sistema en tiempo-real, es decir, se encargan de estimar la severidad de las posibles fallas del sistema instante a instante. Por otra parte, los módulos de pronóstico están orientados a proveer información sobre la salud futura del sistema bajo cierto perfil de utilización, el cual determina el tiempo de vida remanente de éste. Ambos módulos permiten tener la noción de riesgo y confiabilidad en procesos de toma de decisiones.

La incertidumbre debido a la falta de conocimiento sobre la verdadera dinámica de los sistemas o al registro de mediciones ruidosas (o no linealidades de los sensores) pueden ser incorporadas en la representación en espacio-estados de manera directa. Esto permite adoptar enfoques de filtrado estadístico Bayesiano, donde el objetivo principal es estimar la distribución o función de densidad de probabilidad (PDF - del inglés *Probability Density Function*) subyacente de los estados condicional a las mediciones registradas hasta el momento para realizar inferencia estadística. Dado que soluciones analíticas pueden ser encontradas bajo ciertas condiciones (sistemas lineales y Gaussianos, por ejemplo), el real problema a enfrentar es la evaluación de integrales de alta complejidad donde los métodos numéricos tienden a fracasar (por problemas de convergencia o estabilidad numérica), inclusive más aún en el caso de sistemas que involucran alta dimensionalidad. Una forma alternativa de abordar esta problemática es usar filtros de partículas, los cuales definen la base sobre la que se desarrollan los algoritmos propuestos en esta Tesis.

Los objetivos generales de la presente Tesis apuntan a mejorar aspectos de representación de incertidumbre no Gaussiana durante procesos de estimación de estados en sistemas no lineales, así como la caracterización de la propagación de la misma en algoritmos de pronóstico de eventos. Por otra parte, los objetivos específicos apuntan a la implementación del concepto de imputaciones múltiples para garantizar la continuidad de la ejecución de procesos de estimación basados en filtros de partículas ante pérdida parcial de mediciones, redefinir y

formalizar matemáticamente la medida de probabilidad para eventos futuros (catastróficos en el marco de esta Tesis), y finalmente proponer un nuevo método basado también en filtro de partículas para propagar incertidumbre que permita caracterizar de buena forma colas de PDFs en instantes futuros.

En éste capítulo se pretende sentar las bases teóricas para generar intuición y facilitar la comprensión de los desarrollos que se muestran en los siguientes capítulos para la presentación de los nuevos algoritmos propuestos en la presente Tesis. Primeramente se muestra información sobre las baterías de ion-litio y la estimación SoC, puesto que dentro de dicha problemática se prueban y validan los algoritmos que se proponen. Luego se introducen brevemente diversos métodos matemáticos que constituyen la base sobre la cual se desarrollan algoritmos que se proponen en los siguientes capítulos de esta Tesis.

1.1. Baterías de Ion-Litio

A lo largo de las últimas décadas se ha experimentado un incremento significativo en el desarrollo y la producción de vehículos eléctricos. La industria automotriz ha estado poniendo sus esfuerzos en investigación en el desarrollo de dispositivos de almacenamiento de energía (ESDs - del inglés *Energy Storage Devices*) para la producción de vehículos eléctricos híbridos (HEV - del inglés *Hybrid Electric Vehicle*) o completamente eléctricos (EV - del inglés *Electric Vehicle*). Como resultado, los ESDs juegan un rol crucial actualmente en cuanto a la autonomía de los sistemas que energizan. Éste hecho ha motivado la investigación de las celdas de ion-litio debido a sus ventajas por sobre otros tipos de ESDs, siendo su mayor densidad de carga (por unidad de masa o volumen) una de las características más importantes a considerar. Desde la industria automotriz, el concepto de sistemas de gestión de baterías (BMS - del inglés *Battery Management Systems*) [40] surge naturalmente en la búsqueda de sistemas capaces de proveer protección y condiciones óptimas de operación para baterías, mientras se tiene en cuenta de forma simultánea las predicciones de vida útil a través de la supervisión de datos adquiridos en tiempo-real. En este sentido, el Estado-de-Carga (SoC) [40] (una medida de la energía almacenada disponible), el Estado-de-Salud (SoH - del inglés *State-of-Health*) [40] (una medida de la degradación en cuanto a capacidad de almacenamiento de energía), y la vida útil remanente (RUL - del inglés *Remaining Useful Life*) [39] son parámetros que proveen información importante sobre la condición actual de una batería. Desafortunadamente, debido a la incapacidad de medirlas directamente de una manera en línea, los sistemas BMS deben incorporar rutinas de estimación y pronóstico en tiempo-real para poder llevar a cabo sus objetivos.

El SoC es concebido como una cuantificación de la energía almacenada disponible con respecto a la verdadera capacidad de almacenamiento, expresado como un porcentaje de ésta última. Conformar una característica importante a considerar para la autonomía de los sistemas cuando son energizados por ESDs, ya sea como fuente principal o como fuente de respaldo. Debido a que no es posible medir el SoC directamente, los algoritmos de estimación y pronóstico están encargados de la obtención de estimaciones y predicciones válidas a partir de mediciones provistas por sensores que usualmente son ruidosas como la corriente de descarga, el voltaje y la temperatura, llevándolo a su vez a cabo un manejo apropiado del sistema. A

decir verdad, el conocimiento sobre éste es esencial para el control de sistemas autónomos donde el tiempo de EoD juega un rol clave.

De acuerdo a [42], una amplia variedad de métodos han sido propuestos en la literatura para modelar baterías en aplicaciones fuera de línea; como modelos electroquímicos, por ejemplo. Otros métodos más apropiados para implementaciones en línea están basados en representaciones de circuito abierto (OCV - del inglés *Open Circuit Voltage*). Éstos métodos relacionan directamente el SoC y las medidas de voltaje pero requieren largos periodos de descanso para las baterías, siendo ineficientes para estimación en tiempo-real. El método de *Electrochemical Impedance Spectroscopy* (EIS) requiere equipamiento costoso, siendo infactible para aplicaciones prácticas. En este sentido, los esfuerzos en cuanto a la investigación se enfocan en desarrollar algoritmos de estimación y pronóstico basados en relaciones fenomenológicas a través de lógica difusa, redes neuronales y enfoques Bayesianos [36], entre otros.

1.1.1. Representación en Espacio-Estados

Una de las ventajas principales de adoptar un enfoque Bayesiano para estimación cuando se tienen mediciones ruidosas es que es posible incorporar conocimiento a priori sobre la dinámica de los sistemas de forma directa en el modelo así como también las incertidumbres asociadas. Adicionalmente, es posible capturar directamente fenomenología física crítica bajo una representación en espacio-estados, relacionándola a un modelo de observación el cual permite la convergencia de estimadores mediante el cómputo de la verosimilitud de datos medidos secuencialmente.

Propuesto por [41], el modelo en espacio-estados para celdas de ion-litio usado es el siguiente:

Modelo de Transición de Estados

$$x_1(k+1) = x_1(k) + \omega_1(k) \quad (1.1)$$

$$x_2(k+1) = x_2(k) - \frac{\hat{v}(x_1(k), x_2(k), u(k)) \cdot u(k) \cdot T_s}{E_{crit}} + \omega_2(k) \quad (1.2)$$

Modelo de Observación

$$y(k) = \hat{v}(x_1(k), x_2(k), u(k)) + \eta(k) \quad (1.3)$$

con

$$\begin{aligned} \hat{v}(x_1(k), x_2(k), u(k)) = & v_L + (v_0 - v_L) \cdot e^{\gamma \cdot (x_2(k)-1)} + \alpha \cdot v_L \cdot (x_2(k) - 1) \dots \\ & \dots + (1 - \alpha) \cdot v_L \cdot (e^{-\beta} - e^{-\beta \cdot \sqrt{x_2(k)}}) - u(k) \cdot x_1(k) \end{aligned} \quad (1.4)$$

donde $\omega_1(k) \sim \mathcal{N}(0, \sigma_1)$ y $\omega_2(k) \sim \mathcal{N}(0, \sigma_2)$ corresponden a ruido aditivo Gaussiano y $\eta(k) \sim \mathcal{N}(0, \sigma_{obs})$ es también una variable aleatoria distribuida como una Gaussiana que da cuenta de la incertidumbre en las mediciones. El tiempo de muestreo es $T_s[\text{sec}]$ y $u(k) = i(k)[A]$ es

la corriente de descarga considerada como la variable de entrada del sistema, mientras que $y(k) = v(k)[V]$ es el voltaje de la batería considerado a la vez salida de éste. Las variables de estado $x_1(k)$ y $x_2(k)$ son escogidas estratégicamente bajo un significado físico como el módulo de la impedancia interna y el SOC, respectivamente. Finalmente, como el SOC es expresado como un porcentaje de energía, E_{crit} representa una constante de normalización cuya unidad de medida es $[VA \text{ sec}]$. Todos los otros parámetros del modelo son constantes conocidas dentro de cada ciclo de descarga de la batería. Sus valores pueden ser obtenidos siguiendo el procedimiento descrito en [41, 42], mediante una prueba de descarga completa a corriente nominal (constante).

Una alternativa que permite disminuir el tiempo de cómputo para implementaciones de filtrado estadístico en tiempo-real cuando se cuenta con sensores de voltaje de excelente precisión y exactitud, es reemplazar $\hat{v}(x_1(k), x_2(k), u(k))$ por $y(k)$ en la Ec. (1.2) del modelo de transición de estados.

1.2. Filtro de Partículas

Los métodos secuenciales de Monte Carlo (SMC - del inglés *Sequential Monte Carlo*), también conocidos como filtros de partículas (PFs - del inglés *Particle Filters*) son métodos computacionales diseñados para estimar recursivamente la evolución a lo largo del tiempo de la distribución de probabilidad a posteriori de sistemas dinámicos no lineales y no necesariamente Gaussianos, dado un conjunto de mediciones adquiridas secuencialmente a través de sensores. Éste problema es conocido por el nombre de *problema de filtrado óptimo*. A continuación se muestra el marco matemático sobre el cual se sustenta la solución a éste problema mediante el uso de PFs.

Sea $X = \{X_k, k \in \mathbb{N}\}$ un proceso estocástico Markoviano de primer orden denotando un sistema cuyo vector de estados es de dimensión n_x , que tiene una distribución inicial $p(x_0)$ y una probabilidad de transición denotada por $p(x_k|x_{k-1})$. Además, denotando por $Y = \{Y_k, k \in \mathbb{N} \setminus \{0\}\}$ a las observaciones de dimensión n_y ruidosas y condicionalmente independientes. El sistema completo se representa en la forma de espacio-estados como

$$\begin{aligned} x_k &= f(x_{k-1}, \omega_{k-1}) \\ y_k &= g(x_k, v_k) \end{aligned} \tag{1.5}$$

donde ω_k y v_k son variables aleatorias independientes y no Gaussianas necesariamente. Debido a que es difícil computar la distribución a posteriori de filtrado $p(x_k|y_{1:k})$, los estimadores Bayesianos se construyen a partir de la *regla de Bayes*.

Bajo supuestos Markovianos, la distribución a posteriori de filtrado puede ser descompuesta en

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \tag{1.6}$$

En este contexto, los métodos SMC ofrecen una alternativa a técnicas de integración numérica que fallan al ser empleados en sistemas embebidos de tiempo-real debido a los altos

requerimientos computacionales. Los métodos SMC, o PFs, son técnicas de computación estocástica diseñados para simular sistemas altamente complejos de una manera eficiente. En estimación Bayesiana, éstas técnicas simulan distribuciones de probabilidad mediante el uso de una colección de N_p muestras asociadas a pesos llamadas *partículas* $\{x_k^{(i)}, W_k^{(i)}\}_{i=1}^{N_p}$, con $\sum_{i=1}^{N_p} W_k^{(i)} = 1$, tales que generan una distribución de probabilidad constituida por una suma de funciones delta-Dirac.

$$\hat{p}(x_k|y_{1:k}) \approx \sum_{i=1}^{N_p} W_k^{(i)} \delta_{x_k^{(i)}}(x_k) \quad (1.7)$$

El proceso de pesaje de cada muestra es hecho aplicando el algoritmo de remuestreo de importancia secuencial (SIR - del inglés *Sequential Importance Resampling*), el cual es explicado en las siguientes subsecciones.

1.2.1. Sequential Importance Sampling

El concepto de muestreo de importancia (IS - del inglés *Importance Sampling*) es usado al simular muestras de una distribución de probabilidad alternativa para estimar una distribución de probabilidad a posteriori. El punto clave para muestrear de manera exitosa es escoger apropiadamente una distribución de importancia o *importance distribution*, que es la distribución alternativa anteriormente mencionada. Muestrear directamente de distribuciones a posteriori es una tarea común cuando se desean obtener estimaciones de Monte Carlo (MC). Sin embargo, la mayor parte de las veces no es posible debido a que es muy costoso computacionalmente. Por ejemplo, la Ec. (1.8) muestra el cálculo de esperanzas

$$\hat{f}(x_k) = E_{X|Y}\{f(x_k)\} = \int_X f(x_k) p(x_k|y_{1:k}) dx_k \quad (1.8)$$

Al obtener N_p muestras aleatorias independientes idénticamente distribuidas desde $p(x_k|y_{1:k})$, la integral puede ser aproximada por una suma de funciones delta-Dirac.

$$\hat{f}(x_k) \approx \int_X f(x_k) \left(\frac{1}{N_p} \sum_{i=1}^{N_p} \delta_{x_k^{(i)}}(x_k) \right) dx_k \quad (1.9)$$

$$= \frac{1}{N_p} \sum_{i=1}^{N_p} f(x_k) \delta_{x_k^{(i)}}(x_k) \quad (1.10)$$

$$= \frac{1}{N_p} \sum_{i=1}^{N_p} f(x_k^{(i)}) \quad (1.11)$$

Éstas aproximaciones pueden no ser factibles cuando no sea posible muestrear directamente desde $p(x_k|y_{1:k})$, por lo tanto el algoritmo de muestreo de importancia secuencial (SIS - del inglés *Sequential Importance Sampling*) evita éstas dificultades al obtener muestras desde una *importance distribution*, $\pi(\tilde{x}_k|\tilde{x}_{0:k-1}^{(i)}, y_{1:k})$, aproximando la distribución de probabilidad a

posteriori objetivo mediante un pesaje apropiado. Los pesos son actualizados recursivamente de la siguiente forma:

$$w_k^{(i)} = W_{k-1}^{(i)} \cdot \frac{p(y_k | \tilde{x}_k^{(i)}) p(\tilde{x}_k^{(i)} | x_{k-1}^{(i)})}{\pi(\tilde{x}_k^{(i)} | \tilde{x}_{0:k-1}^{(i)}, y_{1:k})} \quad (1.12)$$

donde $\{\tilde{x}_k^{(i)}\}_{i=1}^{N_p}$ es un conjunto de N_p muestras aleatorias obtenidas desde $\pi(\cdot | \tilde{x}_{0:k-1}^{(i)}, y_{1:k})$. Más aún, si los pesos son normalizados

$$W_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N_p} w_k^{(i)}} \quad (1.13)$$

entonces la distribución a posteriori puede ser aproximada por la Ec. (1.7).

1.2.2. Resampling

A medida que se van actualizando los pesos de las partículas mediante SIS, se puede ir apreciando una tendencia de crecimiento en la varianza de los pesos de éstas, lo cual conduce a una degeneración del algoritmo al asignar valores despreciables muy cercanos a cero a los pesos de algunas de ellas. Por ende, aquellas partículas se vuelven inservibles pues corresponden a trayectorias del vector de estados muy poco probables. Para resolver este problema, se incluye un paso de remuestreo o *resampling* en el algoritmo SIS, lo cual conduce a un nuevo algoritmo llamado SIR.

Una expresión analítica para medir el grado de degeneración de la población de partículas está dado por el *effective particle sample size* mostrado en la Ec. (1.14).

$$N_{eff}(k) = \frac{N_p}{1 + Var_{p(\cdot | y_{1:k})}(w(x_k))} \quad (1.14)$$

Dado que no es posible calcular N_{eff} directamente, una estimación de su valor puede ser obtenida a través de

$$\hat{N}_{eff}(k) = \frac{1}{\sum_{i=1}^{N_p} (W_k^{(i)})^2} \quad (1.15)$$

En síntesis, la etapa de *resampling* consiste en remover partículas cuyo peso es muy pequeño mientras que se retienen y replican aquellas con peso considerable [12, 27]. La regla de decisión que permite conocer cuándo es apropiado aplicar esta etapa está dada por lo siguiente. Siempre y cuando $\hat{N}_{eff} \leq N_{thresh}$, con N_{thresh} un umbral fijo, la degeneración de la población de partículas es inminente y debe aplicarse remuestreo.

En general, el PF o algoritmo SIR se encuentra dividido en dos etapas. Primeramente, una *predicción* es realizada utilizando el modelo de transición de estados para generar una distribución de probabilidad a priori $p(x_k | x_{k-1})$. En seguida se realiza una etapa de *actualización* donde se modifican los pesos asociados a las partículas de acuerdo a su verosimilitud $p(y_k | x_k)$. Si la población de partículas se degenera en algún momento, se adiciona un paso de remuestreo, tal como es descrito con anterioridad.

Algorithm 1 SIR Particle Filter

1. Importance Sampling

for $i = 1, \dots, N_p$ **do**

- Muestrear $\tilde{x}_k^{(i)} \sim \pi(x_k | x_{0:k-1}^{(i)}, y_{1:k})$ y

definir $\tilde{x}_{0:k}^{(i)} \triangleq (x_{0:k}^{(i)}, \tilde{x}_k^{(i)})$

- Computar los pesos de importancia

$$w_k^{(i)} = W_{k-1}^{(i)} \cdot \frac{p(y_k | \tilde{x}_k^{(i)}) p(\tilde{x}_k^{(i)} | x_{k-1}^{(i)})}{\pi(\tilde{x}_k^{(i)} | \tilde{x}_{0:k-1}^{(i)}, y_{1:k})}$$

end for**for** $i = 1, \dots, N_p$ **do**

- Normalizar

$$W_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N_p} w_k^{(i)}}$$

end for

2. Resampling

if $\hat{N}_{eff} \geq N_{thres}$ **then****for** $i = 1, \dots, N_p$ **do**

- $x_{0:k}^{(i)} = \tilde{x}_{0:k}^{(i)}$

end for**else****for** $i = 1, \dots, N_p$ **do**

- Muestrear un índice $j(i)$ distribuido de acuerdo a la distribución discreta que satisface

$$P(j(i) = l) = W_k^{(l)} \text{ para } l = 1, \dots, N_p$$

- $x_{0:k}^{(i)} = \tilde{x}_{0:k}^{j(i)}$ y $W_k^{(i)} = \frac{1}{N_p}$

end for**end if**

1.3. Imputaciones Múltiples

La pérdida de datos es un problema que puede ser abordado principalmente desde dos perspectivas. Por una parte, las técnicas de imputación singular o *single imputation* se encargan básicamente de completar los conjuntos de datos mediante la asignación de cierto valor a cada dato faltante. La ventaja de esta perspectiva es que facilita el empleo de métodos diseñados para conjuntos de datos completos. Sin embargo, éstas técnicas son bastante simplistas en cuanto a que carecen de caracterización de incertidumbre tanto para la variabilidad en cuanto a valores que puede adoptar cada dato perdido, como también en cuanto a la incertidumbre asociada al modelo del cual se infieren dichos valores. Por otra parte, la idea de imputaciones múltiples (MI - del inglés *Multiple Imputation*) retiene las ventajas de emplear técnicas de imputación singular y además da cuenta de la incertidumbre asociada al mecanismo de pérdida de datos. Las imputaciones múltiples [44] consisten en crear varios conjuntos de datos completos imputando m valores por cada dato perdido de modo que la variabilidad de posibles valores en torno a los valores verdaderos sea incorporada a la hora de realizar inferencia estadística. No obstante, las imputaciones múltiples tienen desventajas tales como la necesidad de generar diversos posibles valores y la exigencia de una gran capacidad tanto en cuanto a espacio de almacenamiento en memoria como en capacidad de procesamiento de cada uno de los conjuntos completos de datos generados.

Un asunto importantísimo es la tarea de escoger un número correcto de imputaciones [19]. Obviamente, el costo computacional es mayor mientras se incrementa el número de imputaciones. En este sentido, en [44] (p. 114) se muestra que una aproximación de la eficiencia para un estimador -bajo un enfoque *batch*, como es concebido MI originalmente- está dada por

$$\left(1 + \frac{\gamma}{m}\right)^{-1/2} \quad (1.16)$$

en unidades de errores estándar, donde m es el número de imputaciones y γ es la fracción de información faltante en la estimación. Consecuentemente, se pueden obtener excelentes resultados al utilizar un bajo número de imputaciones ($m = 3, 4, 5$).

1.3.1. Filtro de Partículas con Imputaciones Múltiples

Originalmente introducido en [21], el filtro de partículas con imputaciones múltiples (MI PF - del inglés *Multiple Imputation Particle Filter*) extiende al algoritmo PF mediante la incorporación de un procedimiento de imputaciones múltiples en casos donde datos de medición no se encuentran disponibles. de modo que el algoritmo pueda incluir la incertidumbre correspondiente en el proceso de estimación. El principal supuesto estadístico bajo este enfoque es que el mecanismo de pérdida de datos es *Missing at Random* (MAR), por lo tanto, éste no depende de los datos que se pierden dado los que se observan o, en otras palabras, la pérdida de datos en una variable no se relaciona con el valor de dicha variable. De existir tal relación, ésta puede explicarse a partir de las relaciones mutuas con otras variables completamente observadas.

Por legibilidad, es necesario un cambio de notación (válida sólo para ésta sección). Tal

como es establecido en [21], se denotan ahora las mediciones como un vector particionado $U_k = (Z_k, Y_k)$, donde Z_k corresponde a la parte de datos perdidos e Y_k es la parte observada. Entonces, el algoritmo MI PF cumple las mismas funciones y es idéntico al algoritmo SIR PF, excepto en cuanto a que el primero implementa medidas frente a las pérdidas de datos.

Un modelo de imputación expresado como una distribución de probabilidad ϕ se requiere para poder obtener m muestras (imputaciones), que es

$$z_k^j \sim \phi(z_k | y_{1:k}) \quad (1.17)$$

donde $j = \{1, \dots, m\}$ es un supraíndice que denota la imputación. Similarmente a lo que es IS (ver Sección 1.2.1), cada imputación está asociada a un peso p_k^j que cumple la condición $\sum_{j=1}^m p_k^j = 1$. De acuerdo a lo establecido en [30], la distribución de probabilidad filtrada puede ser expresada como

$$p(x_k | y_{1:k}) = \int p(x_k | u_{1:k-1}, y_k) p(z_k | y_{1:k}) dz_k \quad (1.18)$$

Al realizar aproximaciones de MC resulta

$$p(x_k | y_{1:k}) \simeq \sum_{j=1}^m p_k^j p(x_k | u_{1:k-1}, u_k^j) \quad (1.19)$$

donde $u_k^j = (z_k^j, y_k)$ son conjuntos de datos completados a partir de valores imputados. Adicionalmente, al aplicar PF en cada uno éstos conjuntos de datos resulta

$$p(x_k | u_{1:k-1}, u_k^j) \approx \sum_{i=1}^{N_p} W_k^{(i,j)} \delta_{x_k^{(i,j)}}(x_k) \quad (1.20)$$

donde los índices i y j indican a la partícula y a la imputación, respectivamente. Por ende, una aproximación de la distribución de probabilidad a posteriori deseada es

$$p(x_k | y_{1:k}) \approx \sum_{j=1}^m \sum_{i=1}^{N_p} p_k^j W_k^{(i,j)} \delta_{x_k^{(i,j)}}(x_k) \quad (1.21)$$

1.4. Outer Feedback Correction Loops

La *evolución artificial* [39, 31] es un método usado para estimación en modelos con parámetros desconocidos que propone la extensión del vector de estados del sistema para incorporar dichos parámetros en la estimación. Las ecuaciones de estado relacionadas a parámetros desconocidos del modelo asumen modelos de caminatas aleatorias o *random walks* para implementar procedimientos de aprendizaje que permitan la estimación simultánea de parámetros y de las variables de estado originales del sistema. Los lazos de corrección externa (OFCL - del inglés *Outer Feedback Correction Loops*) [38] son una clase de algoritmos que proponen

manipular los hiperparámetros que caracterizan los antes mencionados modelos de caminatas aleatorias. Por ejemplo, si se considera el siguiente sistema

$$x_{k+1} = g(x_k, \theta, \omega_k)$$

donde x_k es el vector de estados, θ es el vector de parámetros del modelo y ω_k es un proceso de ruido estocástico. Si $\theta = [\theta^c \ \theta^{uc}]^T$, donde θ^{uc} es un vector de parámetros desconocidos, la evolución artificial propondría considerar el sistema alternativo

$$\begin{aligned} x_{k+1} &= g(x_k, \theta, \omega_k) \\ \theta_{k+1}^{uc} &= \theta_k^{uc} + \omega_k^\theta \end{aligned}$$

donde ω_k^θ es un proceso de ruido estocástico cuyos hiperparámetros son modificados gradualmente por el OFCL para reducir la incertidumbre asociada a θ^{uc} .

1.5. Unscented Transform y Sigma-Points

La *Unscented Transform* (UT) [23, 24] es un método diseñado para aproximar transformaciones no lineales de variables aleatorias. A pesar de que algunas funciones no lineales puedan ser aproximadas por una serie de Taylor truncada, la UT se enfoca en preservar algunos estadísticos característicos de la transformación no lineal usando para esto conjuntos de muestras seleccionadas determinísticamente, llamadas *sigma-points*.

Sea $x \in \mathfrak{R}^{n_x}$ una variable aleatoria con media \bar{x} y matriz de covarianza P_x , y $g(\cdot)$ una función no lineal de variable real. El objetivo es aproximar

$$y = g(x) \tag{1.22}$$

El método de la UT establece lo siguiente. Primeramente, se computa un conjunto de $2n_x + 1$ muestras escogidas determinísticamente asociadas a pesos, llamadas sigma-points $\mathcal{S}_i = \{\mathcal{X}^{(i)}, \mathcal{W}^{(i)}\}$, $i = 1, \dots, 2n_x + 1$, tales que

$$\mathcal{X}^{(0)} = \bar{x} \quad i = 0 \tag{1.23}$$

$$\mathcal{X}^{(i)} = \bar{x} + \left(\sqrt{(n_x + \kappa) \cdot P_x} \right)_i \quad i = 1, \dots, n_x \tag{1.24}$$

$$\mathcal{X}^{(i)} = \bar{x} - \left(\sqrt{(n_x + \kappa) \cdot P_x} \right)_i \quad i = n_x + 1, \dots, 2n_x \tag{1.25}$$

donde

$$\mathcal{W}^{(0)} = \frac{\kappa}{n_x + \kappa} \quad i = 0 \tag{1.26}$$

$$\mathcal{W}^{(i)} = \frac{1}{2(n_x + \kappa)} \quad i = 1, \dots, 2n_x \tag{1.27}$$

Los pesos satisfacen la condición $\sum_{i=0}^{2n_x} \mathcal{W}^{(i)} = 1$. Por otro lado, $\kappa > 0$ es un parámetro de escala que determina qué tan lejos resultan los sigma-points del valor esperado.

Aplicando la transformación no lineal a cada sigma-point

$$\mathcal{Y}^{(i)} = g(\mathcal{X}^{(i)}) \quad i = 0, \dots, 2n_x \quad (1.28)$$

entonces los estadísticos de y pueden ser aproximados como

$$\bar{y} \approx \sum_{i=0}^{2n_x} \mathcal{W}^{(i)} \mathcal{Y}^{(i)} \quad (1.29)$$

$$P_y \approx \sum_{i=0}^{2n_x} \mathcal{W}^{(i)} (\mathcal{Y}^{(i)} - \bar{y})(\mathcal{Y}^{(i)} - \bar{y})^T \quad (1.30)$$

$$P_{xy} \approx \sum_{i=0}^{2n_x} \mathcal{W}^{(i)} (\mathcal{X}^{(i)} - \bar{x})(\mathcal{Y}^{(i)} - \bar{y})^T \quad (1.31)$$

los cuales corresponden a la media, la matriz de covarianza y la matriz de covarianza cruzada, respectivamente.

La importancia de la UT yace en que provee una metodología para aproximar la PDF de la transformación no lineal de una variable aleatoria preservando con precisión al menos los dos primeros momentos más que intentar aproximar la transformación no lineal en sí misma.

1.6. Weighted Expectation Maximization para GMMs

El algoritmo *Expectation Maximization* (EM) [33] es un método iterativo para encontrar estimadores de máxima verosimilitud de los parámetros de un modelo estadístico asociado a datos incompletos. Cada iteración realiza dos etapas. En la primera etapa, también llamada **E**-step, la esperanza de la verosimilitud denotada por Q es evaluada usando los estimadores de los parámetros del modelo actuales. Luego en la segunda etapa, también llamada **M**-step, se modifican los estimadores de los parámetros del modelo maximizando la esperanza previamente computada. Este proceso es repetido hasta que se alcance alguna condición de convergencia.

En términos matemáticos, asumiendo que $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$ donde \mathcal{Z} denota un set de datos completo y \mathcal{X} e \mathcal{Y} son los datos observados y no observados (perdidos), respectivamente. Por lo tanto, si los datos son obtenidos desde una PDF cuyos parámetros son denotados por Θ , entonces

$$p(z|\Theta) = p(x, y|\Theta) = p(y|x, \Theta)p(x|\Theta) \quad (1.32)$$

La etapa **E**-step consiste en evaluar la siguiente función determinística

$$Q(\Theta, \Theta^{(i-1)}) = \mathbb{E}\{\log(p(\mathcal{Y}|\mathcal{X}, \Theta)) | \mathcal{X}, \Theta^{(i-1)}\} \quad (1.33)$$

donde $\mathcal{X} = x$ es conocido y $\Theta^{(i-1)}$ es la estimación del parámetro de máxima verosimilitud de la iteración previa. Notar que como \mathcal{Y} es inobservable, se asume como una variable aleatoria

y, tomando la esperanza, Q se vuelve una función determinística de Θ . A continuación, $Q(\Theta, \Theta^{(i-1)})$ es maximizado con respecto a Θ en la etapa **M**-step, resultando

$$\Theta^{(i)} = \underset{\Theta}{\operatorname{argmax}} Q(\Theta, \Theta^{(i-1)}) \quad (1.34)$$

En esta Tesis el algoritmo EM es empleado en el contexto de estimación de densidad de kernel (KDE - del inglés *Kernel Density Estimation*) para ajustar los parámetros de modelos de mezcla de Gaussianas (GMMs - del inglés *Gaussian Mixture Models*) debido a la capacidad que tienen de aproximar cualquier PDF con una precisión relacionada directamente al número de componentes considerada. En este sentido, $p(x|\Theta)$ puede ser expresada como

$$p(x|\Theta) = \sum_{j=1}^M \alpha_j p_j(x|\mu_j, \Sigma_j) \quad (1.35)$$

con $\alpha_j > 0$ los pesos de la mezcla tales que cumplen $\sum_{j=1}^M \alpha_j = 1$, y

$$p_j(x|\mu_j, \Sigma_j) = \frac{1}{(2\pi)^{d/2} |\Sigma_j|^{1/2}} e^{-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)} \quad (1.36)$$

donde μ_j y Σ_j son el vector media y la matriz de covarianza de la componente j -ésima de la mezcla. Finalmente, los parámetros de la PDF descrita por el GMM son definidos como $\Theta = (\alpha_1, \dots, \alpha_M, \mu_1, \dots, \mu_M, \Sigma_1, \dots, \Sigma_M)$.

Considerando el conjunto de observaciones $x = (x_1, \dots, x_N)$, y

$$h_j^{(i)}(x_k) = \frac{\alpha_j^{(i)} p_j(x_k|\mu_j^{(i)}, \Sigma_j^{(i)})}{p(x_k|\Theta^{(i)})} \quad (1.37)$$

con $p(x_k|\Theta^{(i-1)}) = \sum_{j=1}^M \alpha_j^{(i-1)} p_j(x_k|\mu_j^{(i-1)}, \Sigma_j^{(i-1)})$, la aplicación del algoritmo EM en este caso resulta en la siguiente actualización de los parámetros para la i -ésima iteración:

$$\alpha_j^{(i)} = \frac{1}{N} \sum_{k=1}^N h_j^{(i-1)}(x_k) \quad (1.38)$$

$$\mu_j^{(i)} = \frac{\sum_{k=1}^N x_k h_j^{(i-1)}(x_k)}{\sum_{k=1}^N h_j^{(i-1)}(x_k)} \quad (1.39)$$

$$\Sigma_j^{(i)} = \frac{\sum_{k=1}^N (x_k - \mu_j^{(i)})(x_k - \mu_j^{(i)})^T h_j^{(i-1)}(x_k)}{\sum_{k=1}^N h_j^{(i-1)}(x_k)} \quad (1.40)$$

¿Qué sucede si el conjunto de datos observados \mathcal{X} posee un peso positivo asociado a cada uno de sus elementos? Tal como en el caso de PFs, ajustar PDFs continuas desde un conjunto finito de muestras asociadas a pesos con el algoritmo EM requiere la incorporación de dichos pesos de alguna manera. En esta Tesis se aplica la misma versión del algoritmo EM empleada en [55] llamada *Weighted Expectation Maximization* (WEM) [33, 55] para los fines antes mencionados. Las modificaciones en las Ecs. (1.38)-(1.40) recaen en el modo en que se computa $p_j(x_k|\mu_j^{(i-1)}, \Sigma_j^{(i-1)})$. Esto es, el lado derecho de la Ec. (1.36) debe incluir la multiplicación del peso correspondiente a x_k .

Capítulo 2

Filtro de Partículas con Imputaciones Múltiples

Las rutinas de estimación y pronóstico tienen una fuerte dependencia de las mediciones adquiridas en tiempo-real para su implementación y, por lo tanto, éstos esquemas están sujetos a pérdidas de información debido a, por ejemplo, problemas de transmisión o desconexión de sensores. Completar la información de los datos perdidos dentro del conjunto de datos adquiridos no es tan simple como rellenarlos con valores promedio. En este sentido, varias estrategias pueden ser adoptadas para resolver el problema de estimación de estados secuencial con conjuntos de datos incompletos. Entre ellos, los métodos de imputaciones singulares o *single imputation* (asumir un cierto valor para cada dato perdido) fallan debido a que carecen de caracterización de incertidumbre. La idea de imputaciones múltiples fue propuesta en [44] para resolver éste problema. Ésta idea propone considerar diversos valores por cada dato perdido y combina la distribución de probabilidad inducida por cada uno de ellos en una sola solución para estimación paramétrica. Ésta técnica dio lugar al MI PF [21], donde se usan métodos de filtro de partículas [11], tomándo en cuenta la incertidumbre asociada a los datos perdidos a través de una estrategia de MI.

En ésta parte de la Tesis se presenta una mejora al enfoque Bayesiano basado en filtro de partículas adoptado en [42] mediante una caracterización de incertidumbre en tiempo-real para un caso de estudio particular, que es la estimación del SoC de baterías de ion-litio, basándose en una estrategia de imputaciones múltiples [2]. Se consideran dos casos de validación para este algoritmo. El primer caso considera la situación donde se pierde un total de 1000 mediciones de voltaje consecutivas, emulando la desconexión del sensor correspondiente durante la ejecución de un ciclo de descarga específico. El segundo caso explora el impacto de perder 500 mediciones de voltaje consecutivas justo antes de la ejecución de un módulo de pronóstico. Los resultados muestran que la incertidumbre asociada a la pérdida de información es acotada con el método propuesto. Más aún, la incertidumbre obtenida con el método se muestra más pequeña que la de simplemente descartar mediciones incompletas y computar la predicción a n -pasos de la PDF a priori.

2.1. Implementación de una Estrategia de Imputaciones Múltiples

En [42] se propone un procedimiento detallado para la estimación y pronóstico de SoC. Sin embargo, ¿qué pasa cuando desconexiones repentinas (o pérdida de datos) afectan el desempeño del sensor? Puede aparecer sesgo posiblemente en las estimaciones de SoC, afectando profundamente la etapa completa de estimación y proveyendo de esta forma información inválida. En consecuencia la autonomía del sistema no podría ser garantizada de ahí en más. En este sentido, se propone un nuevo enfoque desde la Teoría de Imputaciones Múltiples para caracterización de incertidumbre en estimadores de SoC basados en PFs donde se pierden mediciones de voltaje durante periodos extensos de tiempo (teniendo en cuenta que las mediciones de corriente de descarga se encuentran siempre disponibles).

Para la implementación de un MI PF en la estimación de SoC, se dispone de sensores de voltaje de alta precisión y exactitud, con lo que el modelo presentado en la Sección 1.1.1 puede simplificarse de la siguiente manera:

Modelo de Transición de Estados

$$x_1(k+1) = x_1(k) + \omega_1(k) \quad (2.1)$$

$$x_2(k+1) = x_2(k) - \frac{y(k) \cdot u(k) \cdot T_s}{E_{crit}} + \omega_2(k) \quad (2.2)$$

Modelo de Observación

$$y(k) = \hat{v}(x_1(k), x_2(k), u(k)) + \eta(k) \quad (2.3)$$

con

$$\begin{aligned} \hat{v}(x_1(k), x_2(k), u(k)) = & v_L + (v_0 - v_L) \cdot e^{\gamma \cdot (x_2(k) - 1)} + \alpha \cdot v_L \cdot (x_2(k) - 1) \dots \\ & \dots + (1 - \alpha) \cdot v_L \cdot (e^{-\beta} - e^{-\beta \cdot \sqrt{x_2(k)}}) - u(k) \cdot x_1(k) \end{aligned} \quad (2.4)$$

donde la única diferencia con el modelo presentado en la Sección 1.1.1 es que en la Ec. (2.2) se reemplaza $\hat{v}(x_1(k), x_2(k), u(k))$ por $y(k)$ ya que éste último equivale a la salida del sistema, que es adquirida por un sensor de voltaje de buena calidad en cuanto a precisión y exactitud.

El filtro de partículas con imputaciones múltiples -o MI PF- en éste caso se implementa particularmente usando imputaciones de voltaje de una manera diferente dependiendo de en qué etapa del procedimiento de filtrado se encuentre actualmente en aplicación, ya sea el paso de predicción o el de actualización. Ésto se debe precisamente a que la Ec. (2.2) exhibe una dependencia explícita de mediciones pasadas en la transición de estados, con lo que puede concluirse que el modelo empleado no es un modelo oculto de Markov (HMM - del inglés *Hidden Markov Model*) [43].

Durante el paso de predicción, y si las mediciones anteriores de voltaje fueron perdidas, el algoritmo de imputaciones múltiples sugiere obtener muestras (denominadas imputaciones) de voltaje desde una PDF propuesta denotada por ϕ . Cada una de las imputaciones definen

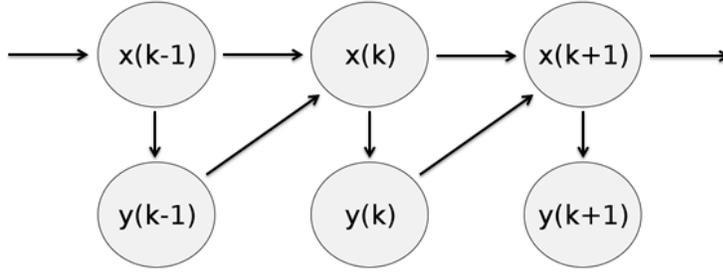


Figura 2.1: El modelo implementado no corresponde a un HMM, pues la transición de los estados depende de la salida pasada del sistema ($v(k - 1)$ en éste caso).

una PDF a priori diferente para el siguiente instante de tiempo. Esto pues $x_2(k + 1)$ depende de $v(k)$ de acuerdo a la Ec. (2.2) y a lo ilustrado en la Fig. 2.1. Sin embargo, dado que las ecuaciones de transición de estados ubican las partículas en diferentes lugares del espacio-estados, al aplicar la regla de Rubin [44] de teoría de imputación múltiple resulta como consecuencia un aumento de la población de partículas. Ésta consecuencia no es directamente una implicación de la regla de Rubin, sino que es más bien a causa de la dependencia de distribución a priori en la medición de voltaje $v(k)$.

La regla de Rubin sugiere que si la variable condicionante de una distribución se pierde (su valor es desconocido), entonces dicha distribución puede ser aproximada por una apropiada suma ponderada de distribuciones condicionadas en varios posibles valores. Éstas últimas son las denominadas *imputaciones* [44].

Asumiendo que la distribución a priori es conocida y que el verdadero valor de voltaje es conocido, entonces las imputaciones de voltaje pueden ser también consideradas en el paso de actualización del algoritmo PF. Más aún, en ése caso las partículas resultantes (las cuales representan la distribución a posteriori) se mantendrán en los mismos lugares dentro del espacio-estados. Por lo tanto, el número de partículas no se ve incrementado a propósito de aplicar la regla de Rubin [44].

Dado que las imputaciones generan un incremento en el número de partículas si es que se ocupan durante el paso de predicción, se requiere entonces la incorporación de un paso de reducción en el algoritmo de modo que se mantenga un número acotado de partículas a través del tiempo; evadiendo un incremento progresivo y exponencial en la población de partículas. De este modo, el SIR PF funciona como lo hace originalmente, especialmente después de que las mediciones de voltaje vuelvan a estar disponibles otra vez.

En particular, el problema de reducir el número de partículas desde $N_p \cdot m$ a N_p -donde N_p es el tamaño original de la población de partículas y m es el número de imputaciones- podría ser resuelto con remuestreo. Sin embargo, éste tipo de técnica falla porque el algoritmo de remuestreo tiende a retener solamente a las partículas de mayor probabilidad, con lo que se descarta la caracterización provista por la estrategia MI que yace principalmente en las colas de la distribución. Por lo tanto, se propone una solución subóptima. El foco principal consiste en preservar la distribución de probabilidad descrita por $N_p \cdot m$ partículas usando solamente N_p de ellas. Entonces, como un intento de resolver este problema las partículas son ordenadas

como función del SoC ($\{x_k^{(l)}, \mathcal{W}_k^{(l)}\}_{l=1}^{N_p \cdot m}$) y luego divididas en grupos de m partículas. De cada grupo se obtiene una sola partícula que resulta del promedio ponderado de las partículas del grupo y su probabilidad se define como la suma de las probabilidades de éstas. Por lo que se generan N_p nuevas partículas como

$$\bar{\mathcal{W}}_k^{(i)} = \sum_{l=m \cdot (i-1)+1}^{m \cdot i} \mathcal{W}_k^{(l)} \quad (2.5)$$

$$\bar{x}_k^{(i)} = \frac{1}{\bar{\mathcal{W}}_k^{(i)}} \cdot \sum_{l=m \cdot (i-1)+1}^{m \cdot i} \mathcal{W}_k^{(l)} x_k^{(l)} \quad (2.6)$$

$\forall i \in \{1, \dots, N_p\}$. El gran supuesto adoptado para éste paso de reducción es que la impedancia interna permanece constante al menos mientras el SoC de la batería se encuentra por sobre un 20 %, lo cual en la práctica lo vuelve casi independiente del valor del SoC. Por supuesto, otros factores también pueden afectar el valor de la impedancia interna como la temperatura de la batería, por ejemplo. De hecho, ésa es la principal razón de que éste parámetro deba ser estimado a partir de mediciones de voltaje y de corriente de descarga. El impacto de estos factores no es considerado en ésta versión en particular del algoritmo, pero pueden formar parte de trabajo de investigación futuro. A pesar de que la dinámica de la impedancia interna es afectada por diversas variables (incluyendo el SoC), todo los resultados presentados se basan en el modelo de transición de estados descrito por las Ecs. (2.1)-(2.2).

La implementación del MI PF propuesto solamente resuelve el problema de pérdida de observaciones de voltaje, mientras que la corriente de descarga se asume en el modelo como una variable de entrada que es conocida en cada instante de tiempo. El modelo de imputación adoptado es definido a través de la PDF inducida por la Ec. (2.3), proveyendo conocimiento a priori de la variabilidad del voltaje que es condicional a la estimación de los estados actualmente disponible. Por lo tanto, en el caso de estimadores basados en PFs, se generan diferentes modelos de imputación para el voltaje de la batería cuando se substituye cada partícula de la distribución a priori en la Ec. (2.3).

Por construcción, el algoritmo propuesto muestra una fuerte dependencia -en el sentido de su desempeño- en la calidad del modelo de observación. Sin embargo, ésta característica incluye al mismo tiempo una fuente de información importante la cual no es tomada en cuenta por los algoritmos de predicción a n -pasos.

Denotando los distintos conjuntos de mediciones con imputaciones como $\tilde{y}_{1:k}^j = \{\tilde{y}_{1:k-1}, y_k^j\}$ donde $\tilde{y}_{1:k} = \{\tilde{y}_{1:k-1}, y_k^1, \dots, y_k^m\}$, con $j \in \{1, \dots, m\}$, la implementación del MI PF se resume en el Algoritmo 2.

2.2. Resultados Experimentales

En ésta parte se aplica el algoritmo de imputaciones múltiples al caso de estimación de SoC de baterías de ion-litio. Particularmente, se pretende que éste método mejore el modo que el SoC es monitoreado en un BMS. Se analiza un conjunto de datos de un ciclo completo

Algorithm 2 Multiple Imputation Particle Filter

1. MI Importance Sampling

if y_k está disponible **then**

- SIR PF

else**for** $j' = 1, \dots, m$ **do****for** $i = 1, \dots, N_p$ **do**

- Muestrear $\tilde{x}_k^{(i,j')} \sim \pi(x_k | x_{0:k-1}^{(i)}, \tilde{y}_{k-1}^{j'})$ y definir $\tilde{x}_{0:k}^{(i,j')} \triangleq (x_{0:k-1}^{(i)}, \tilde{x}_k^{(i,j')})$

end for**end for**

- Computar m imputaciones $y_k^j \sim \phi(\{\tilde{x}_k^{(i,j')}, w_k^{(i,j')}\}, \eta_k)$ y sus pesos asociados p_k^j .

- Reducir la población de partículas de $N_p \cdot m$ a N_p .

$$\tilde{x}_k^{(i,j)} \rightarrow \tilde{x}_k^{(i)}$$

- Definir los pesos de importancia

$$w_k^{(i,j)} = w_{k-1}^{(i)} \cdot \frac{p(y_k^j | \tilde{x}_k^{(i)}) \cdot p(\tilde{x}_k^{(i)} | x_{k-1}^{(i)})}{\pi(\tilde{x}_k^{(i)} | \tilde{x}_{0:k-1}^{(i)}, \tilde{y}_{1:k}^j)}$$

for $i = 1, \dots, N_p$ **do**

- Aplicar regla de Rubin

$$w_k^{(i)} = \sum_{j=1}^m w_k^{(i,j)}$$

end for**for** $i = 1, \dots, N_p$ **do**

- Normalizar

$$\mathcal{W}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^N w_k^{(i)}}$$

end for**end if**

de descarga que contiene un total de 2920 muestras que fueron obtenidas con un equipo experimental localizado en el Laboratorio de Control Avanzado I de la Universidad de Chile. Los valores para los parámetros del modelo descrito por las Ecs. (2.1)-(2.3) son los siguientes: $\alpha=0.0776$, $\beta=16$, $\gamma=19.6445$, $v_0=4.12$, $v_L=3.9773$, y $E_{crit}=19865$. Las mediciones en línea son provistas por datos experimentales de voltaje y corriente de descarga de una celda de ion-litio 18650 con un tiempo de muestreo de $T_s = 1[s]$. El experimento de adquisición de datos se diseña siguiente un perfil de corriente de descarga equivalente al del consumo de un robot terrestre de cuatro ruedas; ver *Set #1* de la Sección 3.2 en [34].

2.2.1. Impacto en la Etapa de Estimación

Para probar el algoritmo durante la etapa de estimación, se considera el caso donde se pierde un total de 1000 mediciones de voltaje consecutivas simulando la desconexión de un sensor. Las estimaciones son obtenidas usando 60 partículas y 10 imputaciones. El desempeño es analizado y comparado considerando un promedio de 30 realizaciones para tres casos diferentes: i) SIR PF con conjunto de datos completo, ii) cómputo de una predicción a 1000-

pasos en ausencia de mediciones, y iii) MI PF con un conjunto de datos incompleto.

La PDF que es usada para muestrear imputaciones de voltaje corresponde a la distribución inducida por la Ec. (1.3), donde las partículas son obtenidas usando la distribución a priori mostrada en las Ecs. (1.1)-(1.2). Éstas partículas, junto con 33 realizaciones del ruido de observación $\eta(k)$, son substituidas en la Ec. (1.3) para generar una distribución empírica condicional a la pérdida de mediciones de voltaje. Las imputaciones son muestreadas aleatoriamente desde la distribución antes mencionada y por lo tanto se les asigna un peso equivalente.

Para éste caso de estudio, el SIR PF convencional se aplica en todos los casos hasta que se empiece a producir pérdida de mediciones. El foco yace en comparar la estrategia MI con una simple predicción a n -pasos [36] que podría ser aplicada cuando se pierden mediciones de voltaje. Además, la estrategia MI es comparada con las estimaciones basadas en PF que son obtenidas sin pérdida de datos. Ambas comparaciones llevan a resultados de impedancia interna, SoC y voltaje, las cuales son expuestas en las Figs. 2.2, 2.3 y 2.4, respectivamente. Para un mejor análisis, las mismas condiciones son adoptadas en todos los casos hasta el instante en que empiezan a perderse datos.

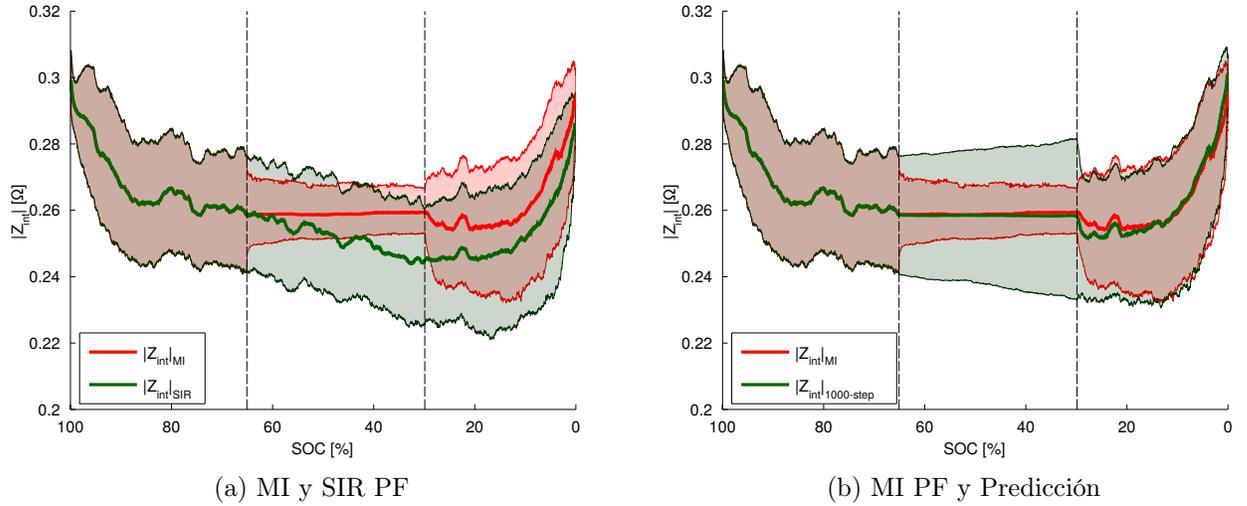


Figura 2.2: Módulo de la impedancia interna como función del SoC [%] para una desconexión de 1000 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 1000-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$.

Tal como se muestra en la Fig. 2.2a, la hipótesis de un valor constante para el módulo de la impedancia interna de la batería se vuelve inválido a medida que se incrementa el número de mediciones perdidas. De hecho, las estimaciones del MI PF difieren significativamente de las estimaciones computadas en base al conjunto de datos completo principalmente

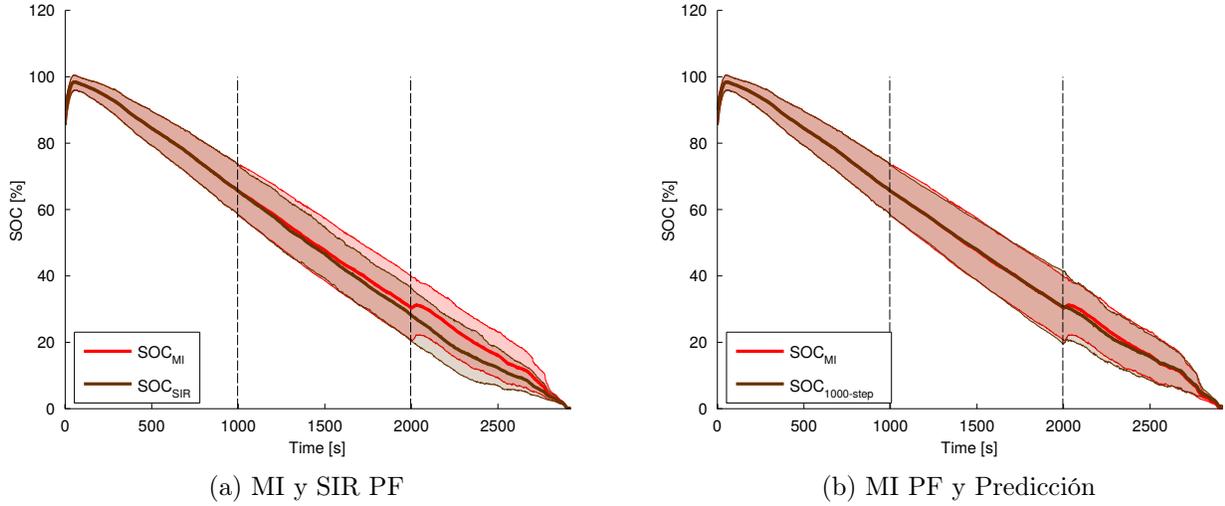


Figura 2.3: Estimación de SoC como función del tiempo [s] para una desconexión de 1000 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 1000-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$.

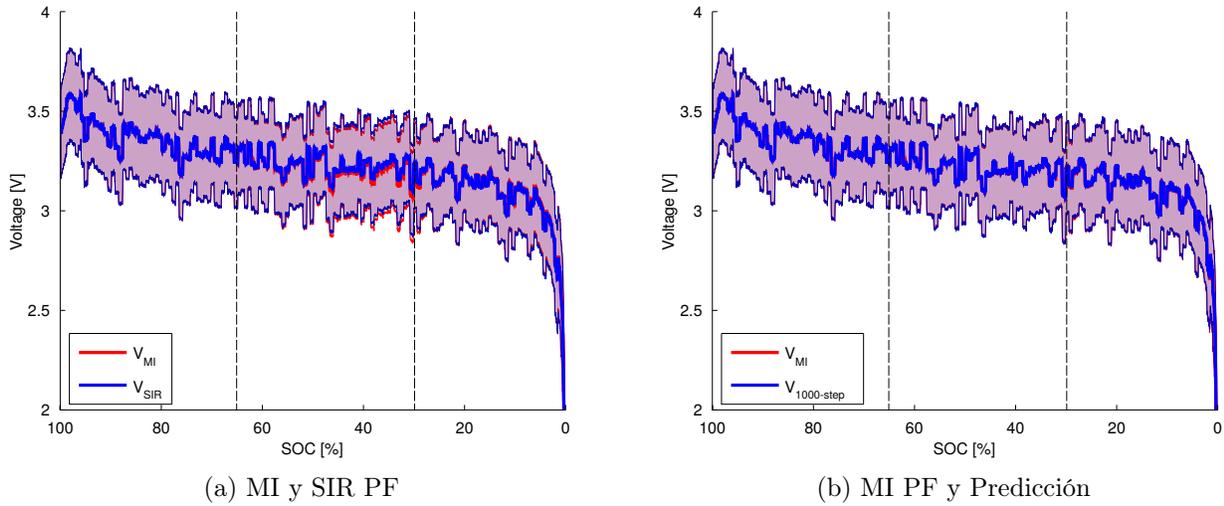


Figura 2.4: Estimación de voltaje como función del SoC [%] para una desconexión de 1000 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 1000-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$.

debido a dicha hipótesis. Adicionalmente, la Fig. 2.2b muestra que las estimaciones del MI PF son similares, en términos de valores esperados, a los generados usando una predicción a 1000-pasos. Por ende, la principal diferencia entre éstos dos métodos, en términos de caracterización de incertidumbre, se deben a las hipótesis asociadas con la evolución del primer estado. La estrategia enfocada en el cómputo de la predicción a 1000-pasos usa la Ec. (1.1) para la transición del primer estado, la cual no representa la dinámica verdadera del proceso sino más bien corresponde a un proceso de aprendizaje llamado evolución artificial [42]; ver Sección 1.4. Si la PDF a priori es caracterizada usando solo el concepto de evolución artificial, la incertidumbre se incrementa durante el transcurso del tiempo incluso recolectando información parcial a través de sensores activos. A pesar de que en éste caso se asume un solo sensor de voltaje para el pack de celdas completo, el potencial de una estrategia MI PF que no depende solamente del modelo de evolución artificial se vuelve más evidente en el caso en que se adquiere información parcial desde otros sensores; los cuales podrían proveer, por ejemplo, información a nivel de celdas de una batería. A pesar de éste hecho, la principal característica del MI PF propuesto es asegurar una caracterización robusta y acotada de la incertidumbre asociada con el SoC, la cual puede ser visualizada en la Fig. 2.3. La Fig. 2.3a muestra cómo la incertidumbre del MI PF se sobrelapa a la provista por el SIR PF, mientras que en la Fig. 2.3b la incertidumbre asociada al MI PF sobrelapa a penas a la caracterización provista por la predicción a 1000-pasos. Es interesante notar que la estrategia MI evade el uso de una etapa de remuestreo, generando similares resultados al de una predicción a varios pasos. No obstante, cuando las mediciones de voltaje dejan de perderse, se incorpora un sesgo en ambos casos (MI PF y la predicción a 1000-pasos). Este problema es generado por la hipótesis de que la impedancia interna de la batería permanece constante, lo cual introduce un sesgo que afecta a la estimación del SoC como un intento de corregir el primer estado.

Las diferencias que priman entre el MI PF y la predicción a n -pasos son en términos de la manera en la cual se toma en cuenta la incertidumbre. Por una parte, el método de la predicción a n -pasos [36] usa el modelo de transición de estados para reubicar las partículas, reemplazando los datos perdidos con la esperanza del modelo de observación, y despreciando el ruido de medición. Se realiza un paso de *regularización* de modo que las partículas tienen pesos equivalentes. Éstos procedimientos son repetidos a lo largo del horizonte de predicción completo.

Por otro lado, la implementación propuesta del MI PF incorpora la incertidumbre que se relaciona con el sistema de medición mismo, simulando la convolución con ruido de medición típico mediante muestreo. La importancia subyacente de generar una caracterización de incertidumbre realista durante el proceso de estimación es que el algoritmo provee condiciones apropiadas para una posible etapa de pronóstico.

En el caso de estimaciones de voltaje, los resultados se muestran en la Fig. 2.4. Puede notarse en la Fig. 2.4a que se adhiere un pequeño sesgo a la distribución del voltaje cuando se usa MI PF, principalmente por el sesgo asociado a la estimación de la impedancia interna. El uso de unas pocas imputaciones (10 en este caso de estudio) provee una caracterización razonable de la variabilidad de la salida al generar una aproximación robusta de las estadísticas verdaderas incluso cuando los datos son parcialmente perdidos. El sesgo permanece despreciable considerando que el total de datos perdidos alcanza los 1000. Sin embargo, la Fig. 2.4b muestra que la predicción a 1000-pasos comparte dicho comportamiento al describir

curvas prácticamente idénticas.

2.2.2. Impacto en la Etapa de Pronóstico

Anteriormente se analiza el efecto de perder datos consecutivos en una ventana de 1000 muestras durante la etapa de estimación. También se analiza el problema que surge de ésta situación, dando lugar a estimaciones sesgadas y, por lo tanto, que carecen de caracterización de incertidumbre en algunos casos. Pero, ¿qué tal si en vez de continuar en estimación luego de la pérdida de datos, se ejecuta una rutina de pronóstico?

Cuando se involucran sistemas de información o de toma de decisiones, el conocimiento sobre el comportamiento futuro del sistema juega un rol fundamental para planear estrategias de control que aseguren un óptimo uso de los recursos disponibles. En ésta parte se expone el desempeño de una rutina de pronóstico descrita en [42] para estimación de SoC de baterías de ion-litio teniendo una pérdida de 500 mediciones de voltaje justo antes de que sea ejecutada una rutina de pronóstico.

Para probar el algoritmo durante la etapa de pronóstico se considera una pérdida de 500 mediciones de voltaje consecutivas justo antes de que la etapa de estimación finalice. Los resultados que se muestran en las Figs. 2.5, 2.6 y 2.7.

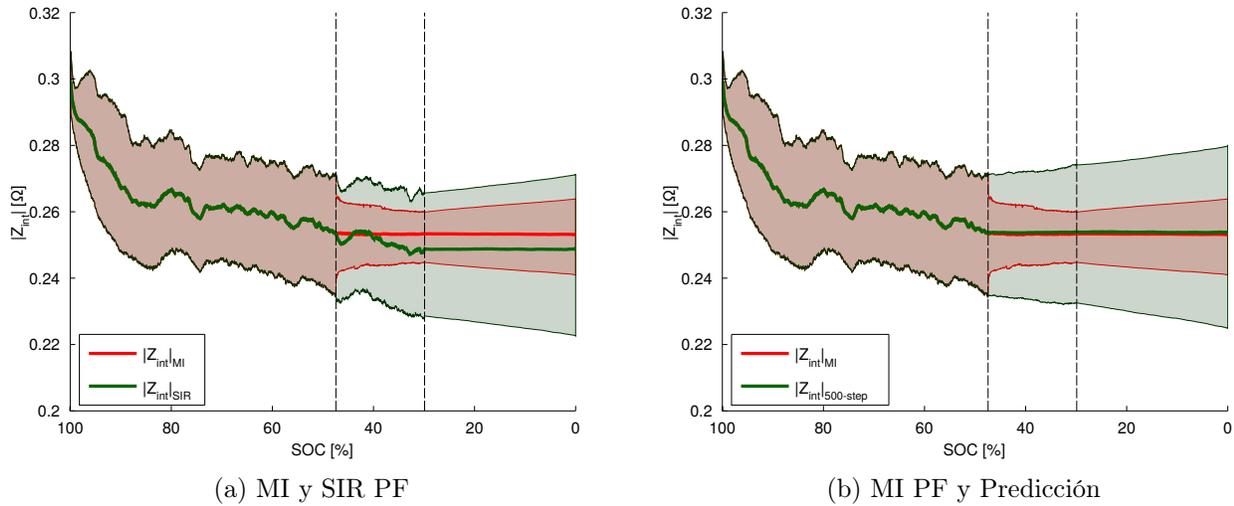


Figura 2.5: Estimación y pronóstico del módulo de la impedancia interna como función del SoC [%] para una desconexión de 500 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 500-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$.

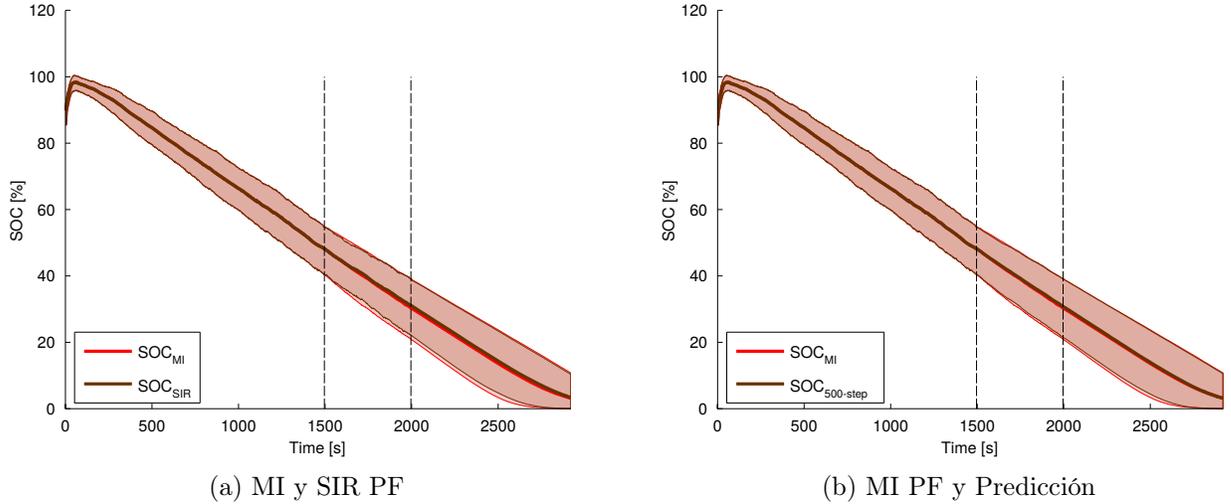


Figura 2.6: Estimación y pronóstico del SoC como función del tiempo [s] para una desconexión de 500 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 500-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$.

Los resultados se obtienen con el uso de 60 partículas y 10 imputaciones, al igual que el estudio del impacto en la etapa de estimación mostrado con anterioridad. El desempeño es analizado considerando un promedio de 30 realizaciones en tres casos diferentes: i) SIR PF con conjunto de datos completo, ii) cómputo de una predicción a 1000-pasos en ausencia de mediciones, y iii) MI PF con un conjunto de datos incompleto. Cada realización incluye a su vez el desempeño promedio de 30 realizaciones de pronóstico con el objetivo de tener resultados lo más cercanos a los esperados.

Cuando se ejecuta el módulo de pronóstico, la corriente de descarga verdadera es desconocida y, por ende, se simula usando cadenas de Markov que ayudan a caracterizar el perfil de uso pasado de la batería para simular el futuro. Por esta razón, se requieren muchas realizaciones de la cadena de Markov antes de pronosticar adecuadamente el tiempo de EoD. Más información sobre éste tópico en particular se puede hallar en [42, 34].

La Figura 2.5 muestra el efecto de la pérdida de mediciones en los resultados de pronóstico para el caso del estado relacionado a la impedancia interna de la batería. Como es mencionado antes, el uso de evolución artificial en el modelo de transición de estados fuerza a la esperanza de la distribución predicha a permanecer constante, aunque su incertidumbre crece continuamente a lo largo del tiempo. El caso del segundo estado es diferente, el cual está asociado con el SoC de la batería; ver Fig. 2.6. Aunque cuando se comparan los resultados del SIR PF con los del MI PF (ver Fig. 2.6a) en términos de la esperanza de la PDF predicha, éstos resultados solo ofrecen diferencias menores, hay muy ligeras disimilitudes en términos

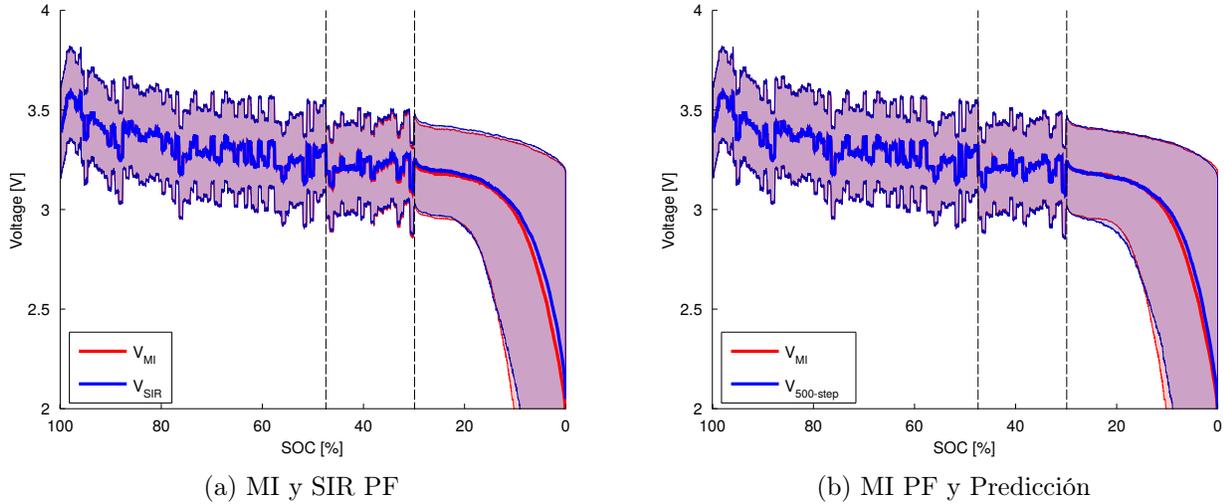


Figura 2.7: Estimación y pronóstico de voltaje como función del SoC [%] para una desconexión de 500 mediciones de voltaje consecutivas denotadas en la región definida entre ambas líneas punteadas verticales. a) Comparación entre MI PF (línea roja) y SIR PF (línea verde) con intervalos de confianza de un 95 %. b) Comparación entre MI PF (línea roja) y predicción a 500-pasos (línea verde) con intervalos de confianza de un 95 %. Sólo el SIR PF cuenta con mediciones de voltaje disponibles todo el tiempo mientras que se considera que hay pérdidas entre las líneas punteadas verticales en el resto de los casos. El tiempo de muestreo es $T_s = 1[s]$.

del tamaño de los intervalos de confianza que son construidas desde las distribuciones predichas para el SoC. Éstas ligeras diferencias son las consecuencias de haber perdido datos antes de ejecutar la rutina de pronóstico. Se pueden obtener conclusiones similares analizando la predicción del voltaje de la batería; ver Fig. 2.7. En este sentido es importante notar que el método propuesto no afecta negativamente la precisión de las predicciones. Mas aún, en cosas donde la evolución de uno de los estados se basa en el concepto de evolución artificial (el cual es muy útil para aprender los verdaderos valores de un modelo paramétrico, pero impreciso para describir su evolución futura), ayuda a acotar la incertidumbre asociada a predicciones a largo plazo, mejorando la precisión de resultados de pronóstico. f

Capítulo 3

Nueva Definición de Probabilidad de Falla en el Contexto de Monitoreo en Línea

A pesar de que en general los esquemas de monitoreo en tiempo real pueden proveer información sobre la condición del sistema a través de la estimación de variables críticas para éste, es natural que se busque encontrar una caracterización de la evolución futura del sistema una vez se detecte una falla. Este conocimiento facilita los procesos de toma de decisiones, generando resultados más confiables y permitiendo tomar medidas que apunten a prolongar lo más posible la vida útil del sistema, encapsulada bajo el concepto de vida útil remanente (RUL - del inglés *Remaining Useful Life*) o, alternativamente, el término de vida (EoL - del inglés *End-of-Life*) [53]; los que se relacionan de la siguiente manera:

$$RUL = EoL - k_p \quad (3.1)$$

donde se asume que el sistema en cuestión se encuentra saludable al menos hasta el instante k_p -ésimo, instante en el cual comienzan a ejecutarse rutinas de pronóstico. Sin embargo, en vez de hablar sobre RUL o EoL los cuales son conceptos relacionados más bien a problemas de degradación, se usa el concepto de tiempo de falla (ToF - del inglés *Time-of-Failure*) que representa el instante de tiempo donde toma lugar una falla catastrófica. El concepto de ToF es más general y apropiado para un espectro más amplio de aplicaciones.

La estimación del ToF es sumamente relevante debido a que provee información sobre la continuidad de la operabilidad de un sistema, es decir, permite saber por cuánto tiempo más puede operar el sistema antes de incurrir en algún tipo de falla que dependiendo de la aplicación puede resultar ser catastrófica. Dada la complejidad de los sistemas industriales, en general no es posible calcular el ToF con precisión infinita, y por lo tanto es relevante caracterizar la incertidumbre asociada a las estimaciones del ToF. De acuerdo a [47], los métodos para predicción del ToF se pueden clasificar en general en dos: los que son en línea y los fuera de línea. Los métodos fuera de línea típicamente hacen uso de algoritmos computacionalmente caros que ofrecen resultados exactos (o precisos). En contraste, los métodos en línea deberían ofrecer un delicado balance entre eficiencia y eficacia. En [47] se mencionan algunos ejemplos para metodologías fuera de línea para caracterizar la incertidumbre en

análisis de quebraduras [48, 49], pronóstico de daño de estructuras [17, 5], electrónica [20], y rodamientos mecánicos [28]. Ejemplos de implementación de pronóstico en línea y esquemas de estimación de RUL pueden encontrarse en [15]. También varios esfuerzos han apuntado a cuantificar incertidumbre asociada con estimación de RUL (o ToF) para los problemas de autonomía energética y degradación de capacidad de baterías de ion-litio [39, 45, 36, 35, 42] y válvulas neumáticas [8]. Metodologías inspiradas en técnicas eficientes de muestreo [9] y métodos analíticos [46] también han sido explorados recientemente.

Los algoritmos de pronóstico basados en PFs usan aproximaciones de la verdadera PDF del vector de estados basándose en distribuciones empíricas; es decir, aproximaciones para la PDF de los estados de un sistema mediante una colección de muestras asociadas a un peso cada una. Específicamente, en [39, 35, 42] el ToF se computa usando un conjunto de N_p partículas (cada partícula es una dupla estado-peso $\{x_k^{(i)}, W_k^{(i)}\}$) tales que

$$\mathcal{P}(ToF = k) = \sum_{i=1}^{N_p} W_k^{(i)} p(failure|X = x_k^{(i)}) \quad (3.2)$$

donde $p(failure|X)$ corresponde a la probabilidad de falla catastrófica del sistema, condicional al valor del vector de estados $x \in \mathfrak{R}^{n_x}$ (valor de “dimensión de la falla” bajo el concepto de *hazard zone* [51, 53]). De hecho, la Ec. (3.2) es un caso particular de

$$\mathcal{P}(ToF = k) = \int_{\mathfrak{R}^{n_x}} p(failure|x_k) p(x_k|y_{1:k_p}) dx_k \quad (3.3)$$

donde la distribución de probabilidad del estado, condicional a mediciones $y_{1:k_p}$, es aproximada por: $p(x_k|y_{1:k_p}) \approx \sum_{i=1}^{N_p} W_k^{(i)} \delta_{x_k^{(i)}}(x_k)$.

Es importante mencionar que para el caso de sistemas dotados de variables críticas que caracterizan su salud (típicamente funciones monótonas de la severidad de la falla), la medida de probabilidad mostrada con anterioridad en las Ecs. (3.2) y (3.3) ha sido mal interpretada como una *Cumulative Mass Function (CMF)*. Lo último debido muy posiblemente a que cuando las variables críticas son monótonas, la medida de probabilidad se muestra quasi-monótona; creciente partiendo en cero y alcanzando su máximo valor en uno. Ésta interpretación por ende se limita sólo al caso de sistemas estrictamente degenerativos; es decir, sistemas cuyo indicador de salud evoluciona de forma monótona (creciente o decreciente). Sin embargo en el caso general, como la medida de probabilidad para el ToF está definida en instantes discretos, ésta debe ser reinterpretada como una *Probability Mass Function (PMF)*.

3.1. Formalización Matemática

Para poder interpretar correctamente la forma en que se mide la incertidumbre en cuanto a las probabilidades de fallar de los sistemas, en el contexto de monitoreo en línea, es primero necesario identificar correctamente los elementos clave que permiten la apropiada caracterización de la medida de probabilidad.

Un evento catastrófico, o falla, puede ser entendido como un proceso estocástico de Bernoulli no estacionario; es decir, cuyas probabilidades varían conforme evoluciona el sistema.

Denotando por \mathcal{F}_k y \mathcal{H}_k la falla y la buena salud del sistema (en el k -ésimo instante de tiempo) respectivamente, es posible definir un espacio de probabilidad $(\Omega_{fail}, \mathcal{B}_{fail}, \mathcal{P}_{fail})$ donde

- $\Omega_{fail} = \{(\mathcal{H}_{k_p}, \mathcal{H}_{k_p+1}, \dots, \mathcal{H}_{k-1}, \mathcal{F}_k), k \in \mathbb{N}\}$ es el espacio muestral que determina todos los posibles resultados durante experimentos, los cuales corresponden a todas las posibles secuencias tales que el sistema permanece saludable salvo en el último instante “ k ”, que es donde falla,
- $\mathcal{B}_{fail} \subseteq 2^{\Omega_{fail}}$ es una σ -álgebra,
- \mathcal{P}_{fail} es una función que asigna una medida de probabilidad a cada evento posible, que es elemento de la σ -álgebra \mathcal{B}_{fail} .

Teniendo en cuenta lo anterior, se persigue caracterizar la probabilidad de falla en un instante k -ésimo. En efecto, denotando $\mathcal{H}_{k_p:k} = (\mathcal{H}_{k_p}, \mathcal{H}_{k_p+1}, \dots, \mathcal{H}_k)$, y de acuerdo a la definición de probabilidad condicional, se obtiene que

$$\mathcal{P}(\mathcal{F}_k) = \frac{\mathcal{P}(\mathcal{F}_k, \mathcal{H}_{k_p:k-1})}{\mathcal{P}(\mathcal{H}_{k_p:k-1}|\mathcal{F}_k)} \quad (3.4)$$

Puesto que $\mathcal{P}(\mathcal{H}_{k_p:k-1}|\mathcal{F}_k)$ es la probabilidad de que el sistema haya estado saludable hasta antes de la falla. Es directo notar que $\mathcal{P}(\mathcal{H}_{k_p:k-1}|\mathcal{F}_k) = 1$, pues el sistema puede solo fallar una vez (en el instante k).

Aplicando la definición de probabilidad conjunta, la Ec. (3.4) puede ser reescrita como

$$\mathcal{P}(\mathcal{F}_k) = \mathcal{P}(\mathcal{F}_k|\mathcal{H}_{k_p:k-1})\mathcal{P}(\mathcal{H}_{k_p:k-1}) \quad (3.5)$$

Considerando que $\mathcal{P}(\mathcal{F}_k)$ es la probabilidad de que el sistema falle en el instante k -ésimo, se tiene lo siguiente:

- ▷ $\mathcal{P}(\mathcal{F}_k|\mathcal{H}_{k_p:k-1})$ corresponde a la medida de probabilidad de falla actualmente utilizada en la literatura, descrita en la Ec. (3.3). Esto es

$$\mathcal{P}(\mathcal{F}_k|\mathcal{H}_{k_p:k-1}) = \int_{\mathfrak{R}^{n_x}} p(failure|x_k)p(x_k|y_{1:k_p})dx_k \quad (3.6)$$

Los algoritmos de pronóstico actualmente empleados en el contexto de monitoreo en línea caracterizan $p(x_k|y_{1:k_p})$ mediante la propagación de la incertidumbre a instantes futuros sin considerar ninguna condicionante más que la dinámica del estado. Al propagar incertidumbre de un instante a otro futuro, se asume implícitamente que el sistema se encuentra saludable en el instante previo.

Es importante mencionar que $\mathcal{P}(\mathcal{F}_k|\mathcal{H}_{k_p:k-1})$ puede interpretarse como la esperanza de $p(failure|x_k)$ con respecto a la medida de probabilidad $p(x_k|y_{1:k_p})$.

- ▷ $\mathcal{P}(\mathcal{H}_{k_p:k-1})$ es la probabilidad de que el sistema se encuentre saludable hasta el instante

$(k - 1)$ -ésimo. En consecuencia,

$$\begin{aligned}
\mathcal{P}(\mathcal{H}_{k_p:k-1}) &= \mathcal{P}(\mathcal{H}_{k-1}|\mathcal{H}_{k_p:k-2})\mathcal{P}(\mathcal{H}_{k_p:k-2}) \\
&= \mathcal{P}(\mathcal{H}_{k-1}|\mathcal{H}_{k_p:k-2})\mathcal{P}(\mathcal{H}_{k-2}|\mathcal{H}_{k_p:k-3})\mathcal{P}(\mathcal{H}_{k_p:k-3}) \\
&\quad \vdots \\
&= \prod_{j=k_p}^{k-1} \mathcal{P}(\mathcal{H}_j|\mathcal{H}_{k_p:j-1})
\end{aligned} \tag{3.7}$$

Luego como $\mathcal{P}(\mathcal{H}_j|\mathcal{H}_{k_p:j-1}) = 1 - \mathcal{P}(\mathcal{F}_j|\mathcal{H}_{k_p:j-1})$ y como el proceso estocástico es de Bernoulli, sigue que

$$\mathcal{P}(\mathcal{H}_{k_p:k-1}) = \prod_{j=k_p}^{k-1} \left(1 - \mathcal{P}(\mathcal{F}_j|\mathcal{H}_{k_p:j-1})\right) \tag{3.8}$$

$$= \prod_{j=k_p}^{k-1} \left(1 - \int_{\mathfrak{R}^{n_x}} p(\text{failure}|x_j)p(x_j|y_{1:k_p})dx_j\right) \tag{3.9}$$

En síntesis, la probabilidad de falla descrita en la Ec. (3.5) queda definida como el producto entre $\mathcal{P}(\mathcal{F}_k|\mathcal{H}_{k_p:k-1})$ y $\mathcal{P}(\mathcal{H}_{k_p:k-1})$, donde el primer término corresponde a la verosimilitud de fallar en el instante k -ésimo (asumiendo que el sistema estuvo saludable en todos los instantes anteriores). El segundo término indica la probabilidad de que el sistema se haya encontrado saludable en todos los instantes hasta el $(k - 1)$ -ésimo.

Una observación importante es que $\mathcal{P}(\mathcal{F}_k|\mathcal{H}_{k_p:k-1})|_{k=k_p} = 0$, pues el sistema se asume saludable en el instante k_p . Consecuentemente, es directo notar que $\mathcal{P}(\mathcal{H}_{k_p:k-1})|_{k=k_p} = 1$.

Por completitud, se muestra que la medida de probabilidad $\mathcal{P}(\mathcal{F}_k)$ propuesta en la Ec. (3.5), efectivamente cumple con las propiedades de una medida de probabilidad, como una función del tiempo. Como primer paso, se demuestra el siguiente lema.

Lema 3.1 *Si $\mathcal{P}(\mathcal{H}_{k_p:k}) = \prod_{j=k_p}^k \left(1 - \mathcal{P}(\mathcal{F}_j|\mathcal{H}_{k_p:j-1})\right)$, entonces*

$$\sum_{k=k_p}^n \mathcal{P}(\mathcal{F}_k|\mathcal{H}_{k_p:k-1})\mathcal{P}(\mathcal{H}_{k_p:k-1}) = 1 - \mathcal{P}(\mathcal{H}_{k_p:n}) \tag{3.10}$$

DEMOSTRACIÓN. Notando que

$$\begin{aligned}
& \mathcal{P}(\mathcal{H}_{k_p:k}) = \prod_{j=k_p}^k \left(1 - \mathcal{P}(\mathcal{F}_j | \mathcal{H}_{k_p:j-1})\right) \\
\Leftrightarrow & \mathcal{P}(\mathcal{H}_{k_p:k}) = \left(1 - \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1})\right) \mathcal{P}(\mathcal{H}_{k_p:k-1}) \\
\Leftrightarrow & \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \mathcal{P}(\mathcal{H}_{k_p:k-1}) = \mathcal{P}(\mathcal{H}_{k_p:k-1}) - \mathcal{P}(\mathcal{H}_{k_p:k}) \\
\Rightarrow & \sum_{k=k_p}^n \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \mathcal{P}(\mathcal{H}_{k_p:k-1}) = \sum_{k=k_p}^n \mathcal{P}(\mathcal{H}_{k_p:k-1}) - \mathcal{P}(\mathcal{H}_{k_p:k}) \\
\Leftrightarrow & \sum_{k=k_p}^n \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \mathcal{P}(\mathcal{H}_{k_p:k-1}) = \mathcal{P}(\mathcal{H}_{k_p}) - \mathcal{P}(\mathcal{H}_{k_p:n})
\end{aligned}$$

$\mathcal{P}(\mathcal{H}_{k_p}) = 1$, debido a que se sabe de antemano que el sistema se encuentra saludable al momento de realizar pronóstico. Sigue que

$$\Rightarrow \sum_{k=k_p}^n \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \mathcal{P}(\mathcal{H}_{k_p:k-1}) = 1 - \mathcal{P}(\mathcal{H}_{k_p:n})$$

□

Ahora se muestra que la medida de probabilidad $\mathcal{P}(\mathcal{F}_k)$ está bien definida.

Teorema 3.2 *Considerando la definición de $\mathcal{P}(\mathcal{F}_k)$ expuesta en la Ec. (3.5), se cumple que*

1. $0 \leq \mathcal{P}(\mathcal{F}_k) \leq 1, \forall k \in \mathbb{N}$
2. $\mathcal{P}(\mathcal{F}_k) = 0, \forall k \leq k_p$
3. $\{\mathcal{P}(\mathcal{F}_k)\}_{k \in \mathbb{N}} \in \ell^1$. Más aún, $\sum_{k=-\infty}^{\infty} \mathcal{P}(\mathcal{F}_k) = 1$ si que se falla en tiempo finito o si es que el soporte de la PDF de los estados del sistema tiene intersección no vacía con la hazard zone hasta el infinito.

DEMOSTRACIÓN.

1. Sigue de la Ley de Probabilidades Totales.
2. Se desprende del hecho que los algoritmos de pronóstico de fallas solamente proveen información útil cuando son utilizados para predecir tiempo de falla (y por ende, se asume que el sistema no ha fallado todavía). Si ya ha ocurrido una falla catastrófica en el sistema, entonces el pronóstico de la falla se convertiría en un problema mal definido.
3. Dado que

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} |\mathcal{P}(\mathcal{F}_k)| &\stackrel{\text{(Teorema 3.2.1)}}{=} \sum_{k=-\infty}^{\infty} \mathcal{P}(\mathcal{F}_k) \\
&\stackrel{\text{(Teorema 3.2.2)}}{=} \sum_{k=k_p}^{\infty} \mathcal{P}(\mathcal{F}_k) \\
&= \sum_{k=k_p}^{\infty} \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \mathcal{P}(\mathcal{H}_{k_p:k-1}) \\
&= \lim_{n \rightarrow \infty} \sum_{k=k_p}^n \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \mathcal{P}(\mathcal{H}_{k_p:k-1}) \\
&\stackrel{\text{(Lema 3.1)}}{=} 1 - \lim_{n \rightarrow \infty} \mathcal{P}(\mathcal{H}_{k_p:n})
\end{aligned}$$

Luego $\{\mathcal{P}(\mathcal{F}_k)\}_{k \in \mathbb{N}} \in \ell^1$ pues $\mathcal{P}(\mathcal{H}_{k_p:n}) \leq 1, \forall n \geq k_p$. Finalmente, considerando que existe un $n \in \mathbb{N}$ grande tal que asegure que el soporte de $p(x_k | y_{1:k_p})$ se interseque con el soporte de $p(\text{failure} | x_k)$ por un periodo de tiempo suficientemente extenso, sigue que

$$\lim_{n \rightarrow \infty} \mathcal{P}(\mathcal{H}_{k_p:n}) = 0$$

La ecuación anterior se cumple trivialmente cuando el RUL del sistema es acotado en el tiempo, un hecho que puede ser asumido con seguridad para todos los propósitos prácticos. □

En el caso de sistemas degenerativos; es decir, aquellos cuya salud decrece monótonamente conforme transcurre el tiempo (sin presentar fenómenos de regeneración), se espera que presenten falla en un periodo finito de tiempo. Incluso en casos donde el sistema pudiese permanecer saludable por un largo periodo de tiempo, por ejemplo cuando exhibe fenómenos regenerativos, la probabilidad de fallar luego de una gran cantidad de tiempo debería ser cercana a cero. De hecho,

$$\begin{aligned}
\lim_{k \rightarrow \infty} \mathcal{P}(\mathcal{F}_k) &= \lim_{k \rightarrow \infty} \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \mathcal{P}(\mathcal{H}_{k_p:k-1}) \\
&= \lim_{k \rightarrow \infty} \mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1}) \prod_{j=k_p}^{k-1} \mathcal{P}(\mathcal{H}_j | \mathcal{H}_{k_p:j-1}) \\
&= \lim_{k \rightarrow \infty} \underbrace{\mathcal{P}(\mathcal{F}_k | \mathcal{H}_{k_p:k-1})}_{\leq 1} \prod_{j=k_p}^{k-1} \left(1 - \mathcal{P}(\mathcal{F}_j | \mathcal{H}_{k_p:j-1})\right) \\
&\leq \lim_{k \rightarrow \infty} \prod_{j=k_p}^{k-1} \left(1 - \mathcal{P}(\mathcal{F}_j | \mathcal{H}_{k_p:j-1})\right) \\
&\leq \lim_{k \rightarrow \infty} \underbrace{\left(\max_{k_p \leq j \leq k-1} \left(1 - \mathcal{P}(\mathcal{F}_j | \mathcal{H}_{k_p:j-1})\right) \right)^{k-k_p}}_{< 1, \text{ pues no falla en tiempo finito}} \\
&= 0
\end{aligned}$$

Capítulo 4

Pronóstico de Fallas Basado en Filtro de Partículas y el Uso de Sigma-Points

Uno de los aspectos más críticos a considerar en la elaboración de procesos de toma de decisiones es la manera en la cual se cuantifica tanto la incertidumbre como el riesgo. Éste último, debido al impacto que las fuentes de incertidumbre puedan tener en los costos finales asociados a cada alternativa; costos que en algunos casos están fuertemente relacionados a la continuidad operacional de procesos industriales. En este sentido, variadas técnicas intentan abordar este problema de caracterización de incertidumbre desde distintos puntos de vista. Algunos métodos siguen enfoques epistémicos o basados en posibilidad, tales como la lógica difusa o la teoría de la evidencia [13, 52, 57, 58]. No obstante, la mayoría de los investigadores han preferido métodos basados en probabilidades para la implementación de algoritmos de pronóstico de fallas en tiempo-real [29], principalmente porque la forma en que se abordan los problemas desde esta perspectiva permite incluir una noción de incertidumbre naturalmente bajo una bien conocida y aceptada formalización matemática.

En el contexto de teoría de probabilidad, los enfoques Bayesianos [11] se alzan como una opción ideal para la mayoría de los sistemas de aprendizaje en línea, pues permiten la implementación de algoritmos de filtrado estadístico (también conocidos como procesadores Bayesianos) que permiten la caracterización de incertidumbre en procesos dinámicos no lineales. Esta tarea es realizada mediante el cómputo de estimaciones a posteriori del vector de estados, donde tanto la información a priori (provista por el modelo del sistema) y las mediciones (adquiridas en tiempo-real) son usadas eficientemente. En pronóstico de fallas, estos estados típicamente corresponden a variables críticas cuya evolución futura a lo largo del tiempo puede afectar significativamente la salud del sistema, por lo que esto podría conducir a una condición de falla. Sin embargo, a excepción de un número reducido de casos, no es factible obtener una solución analítica para el problema de filtrado. Una solución bien conocida para sistemas lineales Gaussianos es el filtro de Kalman (KF - del inglés *Kalman Filter*) [25, 16], el cual provee un estimador óptimo en el sentido que minimiza el error medio cuadrático. Aunque algunos sistemas no lineales pueden ser aproximados por versiones linealizadas en torno a ciertas regiones de operación, las implementaciones del KF son en general subóptimas en la práctica. Más aún, la incertidumbre subyacente puede poseer una PDF distinta de la clásica distribución Gaussiana. En este sentido, los métodos SMC, a.k.a.

PFs [11], proveen una solución interesante al problema de filtrado estadístico en sistemas no lineales y no Gaussianos. Los PFs aproximan la PDF subyacente del estado del sistema mediante un conjunto de muestras asociadas a pesos, las cuales van siendo actualizadas recursivamente a medida que se van registrando mediciones, lo cual permite ir generando a su vez una distribución empírica desde la cual es posible realizar inferencia estadística.

Aunque los PFs permiten caracterizar incertidumbre asociada al vector de estados en el problema de filtrado, el interés en pronóstico se enfoca en describir apropiadamente la evolución futura de la incertidumbre [15], todo para poder tomar medidas preventivas y evitar eventos catastróficos [39]. La simulación de Monte Carlo (MC) [32] ofrece una solución a este problema, incluso aunque el costo computacional asociado complica la implementación en tiempo-real de esquemas de pronóstico basados en MC. La comunidad de *Prognostics and Health Management* (PHM) ha escogido a los algoritmos basados en PFs [37] como los algoritmos *de facto* del estado-del-arte en esta materia, pues ofrecen una alternativa eficiente frente a la simulación de MC en la implementación de esquemas de pronóstico en tiempo-real.

Hasta el momento en la literatura el algoritmo de pronóstico basado en PFs que se ha ido consolidando es el desarrollado en [39]; teniendo diversas aplicaciones [10] como predicción del crecimiento de fallas o de daños a lo largo del tiempo [61], circuitos electrónicos analógicos [50], diseño de componentes [22], predicción de fallas de máquinas eléctricas [56], y pronóstico de SoC y SoH de baterías de ion-litio [36, 35, 42, 2]. Sin embargo, éste algoritmo no contempla caracterizaciones de entradas exógenas que puedan incidir en las predicciones, y adicionalmente opera con un número fijo de partículas de igual peso que se ubican aleatoriamente en el espacio-estados cuyo número puede ser insuficiente para caracterizar apropiadamente PDFs futuras de los estados, en especial sus colas. Éstos antecedentes son los que motivan a desarrollar nuevos algoritmos de pronóstico como el presentado en éste capítulo para presentar metodologías alternativas que puedan mejorar algunos de los aspectos ya mencionados u otros.

A continuación se presenta uno de los métodos propuestos en esta Tesis para la caracterización probabilista de la incertidumbre asociada con la evolución futura en el tiempo de sistemas dinámicos no lineales, dentro de un marco de trabajo definido como problema de pronóstico. El método que se propone busca mejorar los esquemas de pronóstico basados en PFs hasta la fecha conocidos al caracterizar de una manera más apropiada las colas de la PDF asociada al vector de estados del sistema en cuestión y, de esta manera, mejorar también la PMF del ToF de éste mediante la propagación consecutiva de sigma-points [23, 24, 18] (un conjunto de puntos definidos determinísticamente asociados a pesos, que preservan los primeros dos momentos de una PDF) y el cómputo de GMMs. Para evaluar el desempeño del algoritmo propuesto se realiza una comparación de la predicción de ToF entre una solución analítica aproximada y el algoritmo mismo, considerando distintas condiciones y con una gran cantidad de realizaciones. Todo esto se realiza en el contexto de predicción del SoC de baterías de ion-litio.

4.1. Propagación de Incertidumbre Mediante Sigma-Points

El método supone una representación en espacio-estados del sistema:

$$x_{k+1} = f(x_k, u_k, \omega_k) \quad (4.1)$$

$$y_k = g(x_k, u_k, \eta_k) \quad (4.2)$$

donde $x_k \in \mathfrak{R}^n$ es el vector de estados, $u_k \in \mathfrak{R}^m$ es una entrada exógena, $y_k \in \mathfrak{R}^p$ es la salida, $\omega_k \in \mathfrak{R}^n$ y $\eta_k \in \mathfrak{R}^q$ son vectores aleatorios correspondientes al ruido de proceso y observación, respectivamente, y $f(\cdot)$ y $g(\cdot)$ son funciones no lineales de variable real. Los requerimientos considerados en el diseño del algoritmo de pronóstico propuesto son:

- i) Aplicable a sistemas no lineales.
- ii) Bajo costo computacional para aplicaciones en tiempo-real.
- iii) Garantizar cierta noción de invarianza en las aproximaciones.
- iv) Considerar entradas exógenas.

Dado que los PFs son la herramienta *de facto* del estado-del-arte para inferencia estadística [37] en filtrado estadístico no lineal, el algoritmo que se propone cumple con el primer requerimiento al tomar las muestras asociadas a pesos (partículas) de PFs como entrada para la rutina de pronóstico. Como la propagación de incertidumbre a lo largo del tiempo podría general PDFs complejas, tanto el segundo como el tercer requerimiento son difíciles de alcanzar. La cuarta condición permite incluir diferentes posibles escenarios para la operación futura del sistema cuando se computan los resultados de pronóstico.

Los algoritmos de pronóstico deben ser capaces de caracterizar la evolución futura de la incertidumbre asociada al vector de estados, solamente utilizando información recolectada desde las salidas de los sistemas hasta el tiempo presente. Si se asume la entrada exógena como conocida, o al menos caracterizada probabilísticamente, el problema de propagación de incertidumbre debería poder ser resuelto, en teoría, a través de la aplicación sistemática de la Ec. (4.1) de transición de estados. Sin embargo, el cómputo en tiempo-real de la PDF para la predicción de estados requiere esfuerzos computacionales significativos cuando se lidia con sistemas no lineales y no Gaussianos. Para resolver esta problemática, el enfoque propone una aproximación de la PDF futura propagando consecutivamente sigma-points y computando también GMMs. El procedimiento se describe a continuación.

Para motivar la intuición y sólo con fines ilustrativos, consideremos que el sistema posee ruido Gaussiano aditivo

$$x_{k+1} = f(x_k, u_k) + \omega_k \quad (4.3)$$

$$y_k = g(x_k, u_k, \eta_k) \quad (4.4)$$

Se supone que la etapa estimación basada en PF ha procesado información hasta el instante k_p , que es el tiempo de pronóstico; proveyendo una distribución a posteriori empírica para los estados del sistema a través de un conjunto de duplas que corresponden a muestras asociadas

a pesos (partículas) $\{x_{k_p}^{(i)}, W_{k_p}^{(i)}\}_{i=1}^{N_p}$ denotada por

$$p(x_{k_p} | y_{1:k_p}) \approx \sum_{i=1}^{N_p} W_{k_p}^{(i)} \delta_{x_{k_p}^{(i)}}(x_{k_p}) \quad (4.5)$$

Se asume también que la PDF del ruido de proceso ω_k puede caracterizarse (o aproximarse) por una GMM:

$$\omega_k \sim \sum_{l=1}^{N_c} \gamma_k^{(l)} \mathcal{N}(\mu_k^{(l)}, \Sigma_k^{(l)}) \quad (4.6)$$

con $\gamma_k^{(l)} > 0$, $\mu_k^{(l)}$ y $\Sigma_k^{(l)}$ el l -ésimo peso, vector media y matriz de covarianza, respectivamente, cumpliéndose $\sum_{l=1}^{N_c} \gamma_k^{(l)} = 1$. Luego, aplicando la Ec. (4.3) a la Ec. (4.5), la PDF a priori puede ser aproximada como

$$p(x_{k_p+1} | y_{1:k_p}) \approx \left(\sum_{i=1}^{N_p} W_{k_p+1}^{(i)} \delta_{f(x_{k_p}^{(i)})}(x_{k_p+1}) \right) \otimes \omega_{k_p} \quad (4.7)$$

Por lo tanto $p(x_{k_p+1} | y_{1:k_p})$ es aproximada por la convolución (denotada por el operador \otimes) entre una PMF (definida por partículas $\{f(x_{k_p}^{(i)}), W_{k_p+1}^{(i)}\}_{i=1}^{N_p}$) y una PDF continua de ω_k . Notar que el modelo de transición de estados es usado para reubicar las posiciones de las partículas en el espacio-estados mientras que $W_{k_p+1}^{(i)} = W_{k_p}^{(i)}$, debido a que el incremento de la incertidumbre relacionado con la predicción a un paso es modelado por la convolución con el ruido de proceso.

Ya que la PDF del ruido de proceso es una GMM, la Ec. (4.7) puede ser reescrita como:

$$p(x_{k_p+1} | y_{1:k_p}) \approx \sum_{i=1}^{N_p} W_{k_p+1}^{(i)} \sum_{l=1}^{N_c} \gamma_{k_p}^{(l)} \mathcal{N}(x_{k_p+1}; \mu_{k_p}^{(l)} + f(x_{k_p}^{(i)}), \Sigma_{k_p}^{(l)}) \quad (4.8)$$

La Ec. (4.8) provee una representación continua para la PDF del vector de estados en el instante de tiempo $k_p + 1$. Desafortunadamente, éste resultado no puede ser extendido más allá de ése instante en particular, pues $f(\cdot)$ es una función no lineal y entonces la PDF predicha para el instante de tiempo $k_p + 2$ ni siquiera sería una GMM en general. Además, en la mayoría de los casos, las PDFs futuras más allá del tiempo $k_p + 2$ generan distribuciones insolubles computacionalmente para aplicaciones en tiempo-real, debido a que no hay expresiones analíticas para ellas, haciendo difícil llevar a cabo inferencia estadística. Inspirado en la idea de MC, una alternativa para resolver ésta problemática sería representar $\mathcal{N}(x_{k_p+1}; \mu_{k_p}^{(l)} + f(x_{k_p}^{(i)}), \Sigma_{k_p}^{(l)})$ por un conjunto de muestras asociadas a pesos. No obstante, ésta opción inevitablemente daría lugar a un incremento exponencial en el número de muestras a medida que continúa el procedimiento de pronóstico.

El siguiente algoritmo que se describe en cuatro pasos y que se inspira en otro algoritmo denominado *Unscented Transform Approximation Clustering* (UTAC) [18], apunta a superar estas dificultades, proponiendo un procedimiento que puede ser implementado en tiempo-real.

Ayuda a generar versiones aproximadas de la PDF del estado en tiempos futuros $k > k_p$, basándose en el hecho que la etapa de filtrado es resuelta mediante una implementación de PF (es decir, se utiliza una PMF basada en PF; ver Ec. 1.7, es entrada o condición inicial para la etapa de pronóstico). Se supone que la etapa de estimación finaliza en el instante de tiempo k_p , momento en el cual el algoritmo de pronóstico es ejecutado.

• Paso 1: Transición de Estados con Sigma-Points

Se aproxima la PDF del ruido de proceso $\omega_k \sim \sum_{l=1}^{N_c} \gamma_k^{(l)} \mathcal{N}(\mu_k^{(l)}, \Sigma_k^{(l)})$ por una suma ponderada de Delta-dirac al representar cada componente de la GMM por un conjunto de sigma-points. En efecto, considerando que las componentes de la GMM se indexan con l y los sigma-points se indexan con j

$$\sum_{l=1}^{N_c} \gamma_k^{(l)} \mathcal{N}(x_k; \mu_k^{(l)}, \Sigma_k^{(l)}) \approx \sum_{l=1}^{N_c} \sum_{j=1}^{2n_x+1} \gamma_k^{(l)} \mathcal{W}_k^{(l,j)} \delta_{\mathcal{X}_k^{(l,j)}}(x_k) \quad (4.9)$$

donde $\{\mathcal{X}_k^{(l,j)}, \mathcal{W}_k^{(l,j)}\}$ es el conjunto de sigma-points con $l = 1, \dots, N_n$ y $j = 0, \dots, 2n_x$. Así, ahora es posible representar el ruido de proceso mediante realizaciones fijadas determinísticamente a través de la asociación de pesos, que da lugar al conjunto $\{\mathcal{X}_k^{(l,j)}, \gamma_k^{(l)} \mathcal{W}_k^{(l,j)}\}$.

Considerando una población de N_k partículas para el k -ésimo instante de pronóstico, se aplica la transición de estado a cada de ellas con lo que se genera una distribución a priori

$$p(x_k | y_{1:k_p}) \approx \sum_{i^*=1}^{N_k N_c (2n_x+1)} \tilde{W}_k^{(i^*)} \delta_{\tilde{x}_k^{(i^*)}}(x_k) \quad (4.10)$$

con

$$\tilde{x}_k^{(i^*)} = f(x_{k-1}^{(i)}, u_{k-1}, \mathcal{X}_{k-1}^{(l,j)}) \quad (4.11)$$

$$\tilde{W}_k^{(i^*)} = W_{k-1}^{(i)} \gamma_{k-1}^{(l)} \mathcal{W}_{k-1}^{(l,j)} \quad (4.12)$$

donde $i^* = i^*(i, j, l)$. Es decir, a cada i^* se le asigna una única combinación de i, j y l . Por ende, como $i = 1, \dots, N_k$, se tiene entonces que $i^* = 1, \dots, N_k \cdot N_c \cdot (2n_x + 1)$.

• Paso 2: Ajustar una GMM

Considerando que la representación empírica computada en el Paso 1 incorpora la incertidumbre con un incremento en la población de partículas, se aplica el algoritmo WEM (ver Sección 1.6) para ajustar una GMM de $M_k \in \mathbb{N}$ componentes, denotada por GMM_k .

$$GMM_k(x_k) = \sum_{j^*=1}^{M_k} \alpha_{WEM_k}^{(j^*)} \mathcal{N}(x_k; \mu_{WEM_k}^{(j^*)}, \Sigma_{WEM_k}^{(j^*)}) \quad (4.13)$$

Si $k = k_p$, se puede usar el algoritmo k -Means (u otra opción afín) para obtener una condición inicial para la estimación de parámetros de la GMM. Si en cambio $k > k_p$, una alternativa es usar GMM_{k-1} y aplicar la esperanza de $f(\cdot)$ al vector media de cada kernel.

• Paso 3: Simplificación usando Sigma-Points

En éste paso se vuelven a utilizar sigma-points para representar ahora $GMM_k(x_k)$. La idea del Paso 2 es que manipulando el valor de M_k , puede variarse (reducirse) el tamaño de la población de partículas incrementada del Paso 1 de acuerdo a la siguiente relación

$$N_{k+1} = M_k \cdot (2n_x + 1) \quad (4.14)$$

• Paso 4: Repetición

Dado que los Pasos 1-3 definen un invariante (se tiene una población de partículas se tamaño controlado), aquellos pasos pueden ser repetidos de manera indefinida propagando incertidumbre hasta cubrir completamente el horizonte de predicción estipulado.

4.2. Comentarios Sobre el Algoritmo Propuesto

Luego de haber explicado el algoritmo propuesto para propagación de incertidumbre en sistemas no lineales, es importante notar cómo los requerimientos para su diseño son cumplidos así como también precisar ciertas condiciones para la convergencia del algoritmo WEM.

4.2.1. En Cuanto a los Requerimientos

El primer y el cuarto requerimiento se satisfacen claramente por el algoritmo ya que éste toma como base la distribución empírica que suministran los PFs durante la etapa de estimación y además permite una caracterización de entradas exógenas en la aplicación de la función no lineal de transición de estados. El segundo requerimiento se alcanza de manera satisfactoria ya que el Paso 2 provee flexibilidad en cuanto a la simplificación de la PDF a priori en cada instante de tiempo. En el caso del tercer requerimiento, el uso intensivo de sigma-points permite preservar los primeros dos momentos de la distribución, propagando la media y la covarianza de la distribución original con cierta precisión.

La característica más importante del algoritmo propuesto está en realidad relacionada con el cumplimiento del segundo y tercer requerimiento. El algoritmo se muestra elegante en la manera en que localiza muestras de forma determinística en el espacio-estados, de modo que una cantidad finita de ellas son esparcidas para aproximar una PDF cuyo soporte se amplía a medida que se avanza temporalmente hasta el horizonte de predicción. Además, éstas muestras son correctamente ubicadas y asociadas a pesos apropiados bajo el concepto de sigma-points.

El Paso 3 permite ahorrar recursos computacionales dado que provee una cantidad acotada de partículas desde las cuales puede aplicar directamente inferencia estadística. Una alternativa razonable es muestrear desde la GMM ajustada en el Paso 2 pero, como la masa

de probabilidad se esparce sobre el espacio-estados, sería necesario muestrear cada vez más y más dependiendo de cuánto falte para alcanzar el horizonte de predicción, lo cual volvería ineficiente al algoritmo en cuanto a costo computacional.

4.2.2. Convergencia del Algoritmo WEM

Claro está que el algoritmo que se propone depende fuertemente de la convergencia del algoritmo WEM en el Paso 2. Dado que el algoritmo WEM es iterativo y busca una solución óptima local, éste puede converger a soluciones erradas o sencillamente no converger si es que se da alguna de las siguientes condiciones:

- 1) Las condiciones iniciales están muy alejadas de los verdaderos parámetros.
- 2) La estructura cuyos parámetros se busca ajustar se encuentra sobreparametrizada.
- 3) El número de muestras es insuficiente para ajustar cierto número de parámetros.

El uso de muestras asociadas a pesos (que contienen tanto el pesaje proveniente de una descomposición en sigma-points como también de la repartición de masa probabilista que se muestra en la Ec. (4.10)) del Paso 1 permite utilizar menos muestras para caracterizar una PDF. El motivo de esto es que los pesos pueden ser interpretados desde un punto de vista *frecuentista*, asignando cierta frecuencia de aparición a cada muestra, característica carente en un enfoque de muestreo convencional, donde es necesario que una muestra resulte en reiteradas realizaciones en una posición o zona determinada para poder recién acercarse a la medición de su frecuencia de aparición. Es decir, el peso de las muestras contiene valiosa información con respecto a cada muestra que permite converger más rápidamente (pues hay un menor uso de muestras).

Aunque en esta Tesis no se contempla formalmente la convergencia del algoritmo, existen ciertas consideraciones que pueden tomarse de modo que el algoritmo no deje de funcionar por motivos de convergencia. Si por algún motivo ése fuese el caso, lo más probable es que se tenga un escenario en el cual se pretende ajustar una GMM de cierto número de componentes en el Paso 2 con muy pocas muestras provenientes del Paso 1. Una medida para sobrepasar éste problema es incrementar el número de sigma-points producto de la descomposición de GMMs del Paso 3.

En la Sección 1.5 se ve cómo se computan los sigma-points. Éstos se obtienen a partir de la matriz de covarianza del vector de estados, lo cual lleva a que los sigma-points terminen ubicándose a lo largo de los vectores propios de dicha matriz. Recordando que las ecuaciones que determinan las posiciones de los sigma-points quedan dadas por

$$\begin{aligned}
 \mathcal{X}^{(0)} &= \bar{x} & i &= 0 \\
 \mathcal{X}^{(i)} &= \bar{x} + \left(\sqrt{(n_x + \kappa) \cdot P_x} \right)_i & i &= 1, \dots, n_x \\
 \mathcal{X}^{(i)} &= \bar{x} - \left(\sqrt{(n_x + \kappa) \cdot P_x} \right)_i & i &= n_x + 1, \dots, 2n_x
 \end{aligned}$$

y además como la matriz de covarianza del estado puede descomponerse en $P_x = S_x S_x^T$, en [14] se muestra que

$$P_x = S_x S_x^T = S_x I S_x^T = S_x C C^T S_x^T = S_x^r S_x^{rT} \quad (4.15)$$

donde I es una matriz identidad y C es una matriz de rotación, ambas de dimensión n_x . La matriz de rotación C permite obtener sigma-points en direcciones distintas de los vectores propios de la matriz de covarianza P_x manteniendo las propiedades de la UT y, por lo tanto, permite conseguir un mayor número de sigma-points (cuyos pesos deben ser definidos de manera apropiada).

En el caso donde $n_x = 2$, la matriz de rotación $C = C(\theta)$ sería

$$C(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad (4.16)$$

Un ejemplo ilustrativo se muestra en la Fig. 4.1, donde la posición convencional de los sigma-points -que es la misma descrita en la Sección 1.5- queda definida por la dirección de los vectores propios, que corresponde al caso de $\theta = 0$.

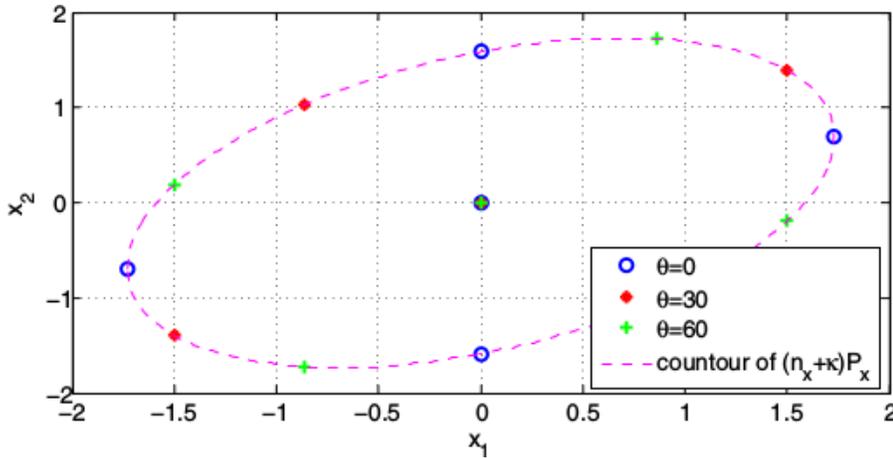


Figura 4.1: Ejemplo de rotación de sigma-points para el caso de dos dimensiones, [14].

4.3. Análisis del Algoritmo: Estimación de EoD

En ésta sección se provee un análisis del desempeño del algoritmo propuesto en un caso de estudio simplificado en particular que trata sobre el problema de pronóstico de SoC de baterías de ion-litio (estimación del tiempo de EoD). Éste estudio se basa en el procedimiento descrito en [35] para el análisis de esquemas de pronóstico basados en PFs para SoH con caracterización estadística de fenómenos de regeneración. La intención detrás de la utilización de casos de estudio simplificados es contribuir a un mejor entendimiento de las fortalezas de algoritmos en casos donde es posible computar la solución analítica para la predicción de la PDF del vector de estados.

Considerando aspectos básicos de baterías de ion-litio que se muestran en la Sección 1.1, se presenta a continuación un modelo lineal bajo una representación en espacio-estados que caracteriza el comportamiento de baterías en cuanto a SoC en un rango de operación interesante (es donde la mayoría de las baterías opera). Dicho modelo es posteriormente usado para obtener una aproximación de la solución analítica para la PDF futura del SoC que es considerada para análisis de desempeño.

Es importante mencionar que para evaluar el desempeño del algoritmo se utilizan métricas definidas a partir de la PMF generada para el ToF. Dicha PMF se construye utilizando la medida de probabilidad para ToF introducida en el Capítulo 3 donde, en el contexto de ésta aplicación, se entiende indistintamente que el tiempo de ToF es equivalente al de EoD. De hecho, en el caso de problemas de estimación y pronóstico de SoC, la escasez energética representa un modo de falla del sistema.

A pesar de las características no lineales asociadas a los problemas de estimación y pronóstico del SoC de baterías, para la mayor parte del rango de operación del SoC la relación entre SoC y la curva de voltaje de circuito abierto (OCV - del inglés *Open Circuit Voltage*) puede ser bien caracterizada por una función afín; ver “zona 2” en la Fig. 4.2.

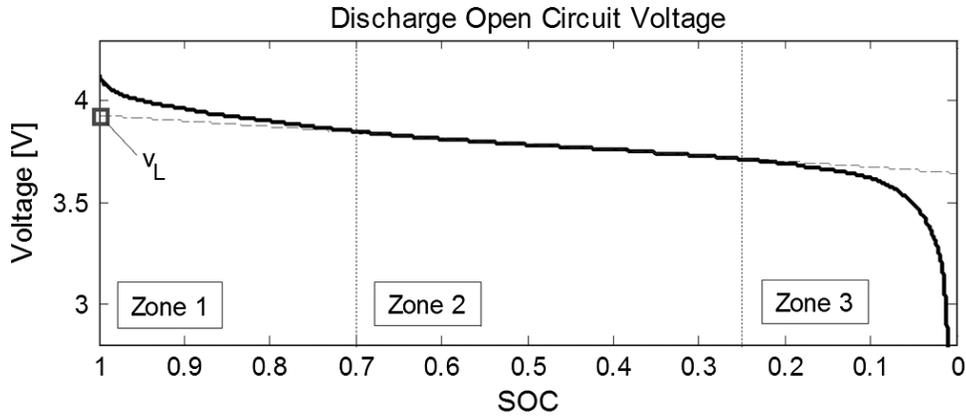


Figura 4.2: Curva de OCV de una celda de ion-litio (línea negra) y proyección de zona lineal (línea segmentada transversal a las zonas) como función del SoC, [41, 42].

Como se establece en [42], el modelo de la dinámica asociada al SoC permanece lineal y Gaussiana a lo largo de éste rango de operación, el cual queda determinado por los valores de SoC que van desde un 25 % hasta un 70 % aproximadamente. Por lo tanto, en la “zona 2” el voltaje de salida puede ser expresado como una función del OCV, como en [42]

$$v(k) = v_{ocv}(k) - i(k)|z_{imp}(k)| + \eta(k) \quad (4.17)$$

donde $i(k)$ es la corriente de descarga (medida en Amperes), $|z_{imp}(k)|$ es la magnitud de la impedancia interna (medida en mili Ohms) y $\eta(k)$ representa ruido de medición Gaussiano de varianza σ_η^2 . Aunque $v_{ocv}(k)$ posee un comportamiento no lineal y $|z_{imp}(k)|$ es en general un parámetro variante en el tiempo de la batería, si se delimita el análisis a la “zona 2”, la estimación de SoC puede ser llevada a cabo en un marco de filtrado de un sistema lineal a través de la siguiente representación en el espacio-estados de batería.

Modelo de Transición de Estados

$$x_1(k+1) = x_1(k) + \omega_1(k) \quad (4.18)$$

$$x_2(k+1) = x_2(k) - \frac{\hat{v}(x_1(k), x_2(k), u(k)) \cdot u(k) \cdot T_s}{E_{crit}} + \omega_2(k) \quad (4.19)$$

Modelo de Observación

$$y(k) = \hat{v}(x_1(k), x_2(k), u(k)) + \eta(k) \quad (4.20)$$

con

$$\hat{v}(x_1(k), x_2(k), u(k)) = v_{ocv}(x_2(k)) - u(k)x_1(k) \quad (4.21)$$

$$= v_L + \alpha v_L(x_2(k) - 1) - u(k)x_1(k) \quad (4.22)$$

En ésta representación se define $u(k) = i(k)$ como la corriente de descarga que viene a ser la entrada del sistema, mientras que $y(k) = v(k)$ es el voltaje producto de la diferencia de potencial en bornes de la batería y la vez es la salida del sistema. El OCV modelado en la región de operación lineal antes mencionada queda caracterizado como

$$v_{ocv}(x_2(k)) = v_L + \alpha v_L(x_2(k) - 1), \quad (4.23)$$

donde v_L y α son parámetros del modelo determinados bajo un procedimiento fuera de línea descrito en [42].

El modelo en espacio-estados en este caso es similar al presentado en la Sección 1.1.1. La diferencia radica en que originalmente $\hat{v}(x_1(k), x_2(k), u(k))$ modela el voltaje en los bornes de la batería a lo largo de todo un ciclo de descarga, que incluye las zonas 1-3, pero que en éste caso se limita a la caracterización sólo a la “zona 2”, que es la zona lineal de operación. De éste modo el modelo de baterías es un *Hidden Markov Model* (HMM) [43], dependiendo así la transición de estados solamente del estado anterior y de la entrada anterior del sistema.

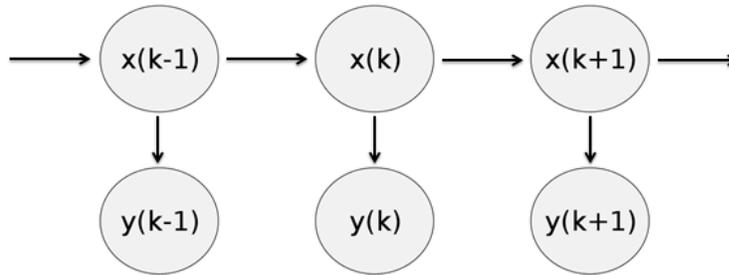


Figura 4.3: El modelo corresponde a un HMM, pues la transición de los estados sólo depende del instante anterior y deben ser inferidos indirectamente a través de mediciones.

Pronóstico es básicamente el estudio de la propagación de incertidumbre a través del tiempo. Sin pérdida de generalidad, el análisis presentado aquí considera la situación donde la evolución futura de las entradas exógenas (corriente de descarga en este caso) es caracterizada

de una forma probabilística, para enfocar la propagación a lo largo del tiempo de las fuentes de incertidumbre asociadas con la etapa de estimación. Notar que aunque las ecuaciones de transición de estados dependen no linealmente de $u(k)$, el problema se vuelve lineal en los estados y por ende es posible computar analíticamente las predicciones para el vector de estados y matriz condicionales:

$$\hat{x}(k+1) = A(k)\hat{x}(k) + B(k) \quad (4.24)$$

$$P(k+1) = A(k)P(k)A^T(k) + R_w(k), \quad (4.25)$$

con matrices variantes en el tiempo definidas como

$$A(k) = \begin{bmatrix} 1 & 0 \\ \frac{u(k)^2}{E_{crit}}T_s & 1 - \frac{\alpha v_L u(k)}{E_{crit}}T_s \end{bmatrix} \quad (4.26)$$

$$B(k) = \begin{bmatrix} 0 \\ (\alpha - 1)\frac{v_L u(k)}{E_{crit}}T_s \end{bmatrix} \quad (4.27)$$

$$R_w(k) = \begin{bmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{bmatrix} \quad (4.28)$$

Asumiendo que las entradas exógenas son caracterizadas probabilísticamente a través de una cadena de Markov de dos estados (primer orden) [34, 42], cada transición de estados genera dos casos probables $u_0(k)$ y $u_1(k)$ y, por lo tanto, la distribución del estado x_2 (SoC) puede ser descrita por una mezcla de Gaussianas como se muestra a continuación

$$x_2(k) \sim \sum_{j=1}^{2^k} \omega_{j,k} \mathcal{N}(\mu_{x_2,j}(k), \sigma_{x_2,j}^2(k)) \quad (4.29)$$

donde cada parámetro es evaluado dependiendo de si j es impar o par, respectivamente.

Dado que el número de componentes de la mezcla de Gaussianas crece exponencialmente a medida que transcurre el tiempo, su solución analítica se vuelve costosa computacionalmente. Como el número de componente se duplica en cada paso de predicción, por simplicidad se mezclan algunos kernels Gaussianos [7], emparejando aquellas cuya vector media sea más cercano de acuerdo a la distancia Mahalanobis de la siguiente forma:

$$w_{mix} = \sum_j w_j \quad (4.30)$$

$$\mu_{mix} = \sum_j \frac{w_j}{w_{mix}} \mu_j \quad (4.31)$$

$$\Sigma_{mix} = \sum_j \frac{w_j}{w_{mix}} (\Sigma_j + (\mu_j - \mu_{mix})(\mu_j - \mu_{mix})^T) \quad (4.32)$$

donde w_{mix} , μ_{mix} y Σ_{mix} son el peso, el vector media y la matriz de covarianza de la nueva Gaussiana producto de una mezcla, respectivamente.

4.3.1. Métricas de Evaluación de Desempeño

Con el objeto de estudiar el desempeño del algoritmo propuesto, el análisis se establece en torno a las siguientes métricas: esperanza, varianza y $JITP_{\alpha\%}$. Las primeras dos métricas son bastante conocidas en general, no obstante es importante detallar bien de qué se trata la última [35].

El $JITP_{\alpha\%}$ busca caracterizar el riesgo asociado a las decisiones basándose en el resultado que ofrecen los algoritmos de pronóstico. De hecho, el riesgo es una función de las colas de la PMF asociada a la falla. En éste sentido, el concepto de *Just In Time Point* (JITP) se define como

$$JITP_{\alpha\%} = \underset{k}{\operatorname{argmin}}(\mathcal{P}(ToF \leq k) \geq \alpha\%) \quad (4.33)$$

Como puede apreciarse, el JITP genera la noción de un umbral para la probabilidad de fallar, la cual a su vez permite definir el índice α -crítico como una medida de la aversión al riesgo. En efecto,

$$\alpha_{crit} = \underset{\alpha}{\operatorname{argmax}}(JITP_{\alpha\%}(k_p) \leq ToF)_{\forall k_p \in [1, ToF]} \quad (4.34)$$

Éste valor se define como el máximo $\alpha \in [0, 100]$ que garantiza que el valor de $JITP_{\alpha\%}(k_p)$ es más pequeño que el verdadero valor del ToF, $\forall k_p \in [1, ToF]$, donde k_p es el instante de tiempo en el cual se inicia el pronóstico.

4.3.2. Resultados Experimentales

El algoritmo propuesto es aplicado al problema de pronóstico de EoD de baterías de ion-litio, estableciendo una comparativa en base a métricas entre la solución provista por el algoritmo y la aproximación de una solución analítica. Por ende, para que sea posible obtener la solución analítica aproximada, se realiza pronóstico a lo largo de una región lineal de operación entre un 70% y un 20% de SoC (ver “zona 2” en Fig. 4.2). Para la implementación del algoritmo, algunas funciones de Matlab Toolkit ReBeL-0.2.7 desarrollado por Rudolph van der Merwe y Eric A. Wan son utilizadas. Los valores para los parámetros del modelo descrito por las Ecs. (4.18)-(4.20) son los siguientes: $\alpha=0.15$, $\beta=12$, $\gamma=6.6061$, $v_0=4$, $v_L=3.8126$, y $E_{crit}=19865$. Las mediciones en línea son provistas por datos experimentales de voltaje y corriente de descarga de una celda de ion-litio 18650. El conjunto de datos está conformado por 3638 observaciones sucesivas con un tiempo de muestreo de $T_s = 1[s]$. El experimento de adquisición de datos se diseña siguiente un perfil de corriente de descarga equivalente a la realización de una cadena de Markov homogénea de dos estados, los que representan cambios aleatorios entre niveles altos y bajos de corriente; ver *Set #2* de la Sección 3.2 en [34]. Los estados de la cadena de Markov son $u_0 = 1[A]$ y $u_1 = 3[A]$, y el kernel de transición es

$$K = \begin{bmatrix} 0,5709 & 0,4291 \\ 0,5577 & 0,4423 \end{bmatrix} \quad (4.35)$$

Como se poseen dos posibles entradas, se tiene entonces que

$$\hat{p}(x_k|y_{1:k_p}) \approx \sum_{i=0}^1 p(u_i)\hat{p}(x_k, u_i|y_{1:k_p}), \quad \forall k \geq k_p \quad (4.36)$$

donde cada PDF $\hat{p}(x_k, u_i|y_{1:k_p})$, con $i = 0, 1$, se computa por separado aplicando el Paso 1, para posteriormente unir ambos resultados en el Paso 2. En cada instante se actualiza la probabilidad de cada entrada aplicando el kernel de transición de la Ec. (4.35).

Por otra parte, la *hazard zone* [51, 53] se define en éste caso como

$$p(failure|x_k) = \begin{cases} 1, & SoC(k) \leq 40\% \\ 0, & \sim \end{cases} \quad (4.37)$$

Para poder evaluar el desempeño del algoritmo de pronóstico basado en sigma-points, se consideran tres escenarios de estudio. En cada uno de ellos se utiliza un número constante de partículas y de componentes en las GMMs del Paso 2, así como también se asumen valores constantes para el valor de $\kappa > 0$ (ver Sección 1.5) tanto en el Paso 1 como en el Paso 3 del algoritmo.

Denotando el valor de κ para el cómputo de sigma-points del Paso 1 como κ_1 y como κ_3 al del Paso 3, se tiene lo siguiente. Primeramente, luego de estudiar el impacto de variar el valor de κ_1 en una serie de experimentos se encuentra que para valores menores o iguales a 1, los resultados no se ven afectados, por lo que se asume en todos los casos de estudio que $\kappa_1 = 0,1$. Esto se debe principalmente a que los sigma-points del Paso 1 son empleados netamente en el ajuste de GMM del Paso 2. Por otra parte, se obtiene que el algoritmo es muy sensible al valor de κ_3 ya que éste determina la densidad en el espacio-estados de la población de partículas empleada como entrada del algoritmo en el siguiente tiempo de predicción, con lo que se pierde gran parte de la aproximación si es que las muestras se alejan mucho del soporte de la verdadera PDF y, por consiguiente, el valor de κ_3 debe ser pequeño. Se encuentra que los resultados se ven afectados a partir de un valor de 0.15, por lo que se opta por definir $\kappa_3 = 0,1$ en todas las simulaciones.

Los tres casos con los que se estudia el nuevo algoritmo de pronóstico y se contrasta con la solución analítica consideran, de acuerdo a la notación utilizada en la descripción de los pasos del algoritmo en la Sección 4.1, que N_k (tamaño de la población de partículas en el instante k -ésimo) y que M_k (número de componentes para la GMM ajustada en el instante k -ésimo) permanecen constantes y son los siguientes

- Caso 1: $N_k=45$ y $M_k=5$.
- Caso 2: $N_k=90$ y $M_k=10$.
- Caso 3: $N_k=135$ y $M_k=15$.

Es importante notar que de acuerdo a la Ec. (4.14), se tiene que

$$\frac{N_k}{M_k} = 9 \text{ (número de sigma-points del Paso 3)}$$

pero de acuerdo a la Sección 1.5 el número de sigma-points debería ser $2n_x + 1 = 5$ (pues $n_x = 2$). Lo que ocurre en éste caso de estudio es que asumiendo un total de cinco sigma-points en el Paso 3 existen problemas de convergencia, por lo que se incrementa dicho número; ver Sección 4.2.2. Por lo tanto, en el Paso 3 los sigma-points se calculan como se muestra a continuación (ver Sección 1.5 como referencia).

$$\mathcal{X}^{(0)} = \bar{x} \quad i = 0 \quad (4.38)$$

$$\mathcal{X}^{(i)} = \bar{x} + \left(\sqrt{(n_x + \kappa) \cdot P_x} \right)_i \quad i = 1, \dots, n_x \quad (4.39)$$

$$\mathcal{X}^{(i)} = \bar{x} - \left(\sqrt{(n_x + \kappa) \cdot P_x} \right)_i \quad i = n_x + 1, \dots, 2n_x \quad (4.40)$$

$$\mathcal{X}^{(i)} = \bar{x} + \left(\sqrt{(n_x + \kappa) \cdot P_x \cdot C(\theta)} \right)_i \quad i = 2n_x + 1, \dots, 3n_x \quad (4.41)$$

$$\mathcal{X}^{(i)} = \bar{x} - \left(\sqrt{(n_x + \kappa) \cdot P_x \cdot C(\theta)} \right)_i \quad i = 3n_x + 1, \dots, 4n_x \quad (4.42)$$

donde

$$\mathcal{W}^{(0)} = \frac{\kappa}{n_x + \kappa} \quad i = 0 \quad (4.43)$$

$$\mathcal{W}^{(i)} = \frac{1}{4(n_x + \kappa)} \quad i = 1, \dots, 4n_x \quad (4.44)$$

y $C(\theta)$ se define al igual que en la Ec. (4.16). Además se asume $\theta = 45^\circ$ de modo que los sigma-points queden equidistantes. Con lo anterior se alcanza un total de $4n_x + 1 = 9$ sigma-points.

A continuación se muestran los resultados obtenidos. En cada caso se obtienen 30 realizaciones que comprenden tanto la etapa de estimación como la de pronóstico.

Caso 1: $N_k=45$ y $M_k=5$, $\forall k \geq k_p$

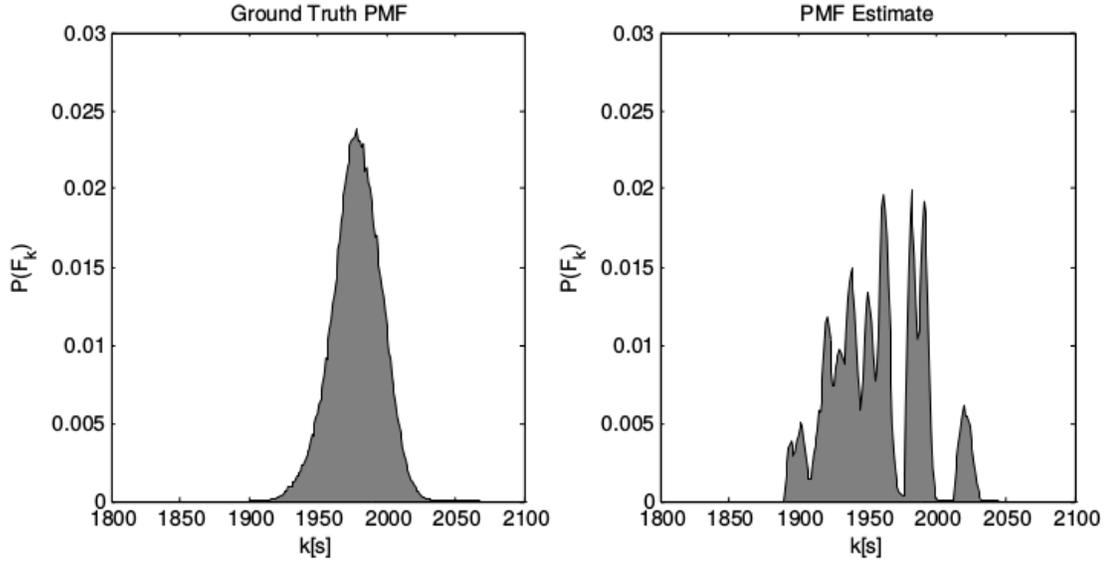


Figura 4.4: Curvas promedio de PMFs para el EoD producto de 30 realizaciones considerando una población de 45 partículas y el ajuste de GMMs de 5 componentes generadas a partir de la solución analítica aproximada (izquierda) y del algoritmo basado en sigma-points (derecha).

Promedios de Estadísticos de la PMF (Solución Analítica Aproximada)					Promedios de Estadísticos de la PMF (Algoritmo Propuesto)				
$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$	$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$
1978.3	125.3	1959.0	1963.4	1966.4	1957.3	13.6	1952.7	1953.4	1954.0

Tabla 4.1: Valores promedio de los estadísticos producto de 30 realizaciones.

Desviaciones Estándar de Estadísticos de la PMF (Solución Analítica Aproximada)					Desviaciones Estándar de Estadísticos de la PMF (Algoritmo Propuesto)				
$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$	$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$
13.9207	47.3337	13.4792	13.4048	13.4154	31.9881	8.3891	32.5283	32.5285	32.5083

Tabla 4.2: Desviación estándar de los estadísticos producto de 30 realizaciones.

Caso 2: $N_k=90$ y $M_k=10$, $\forall k \geq k_p$

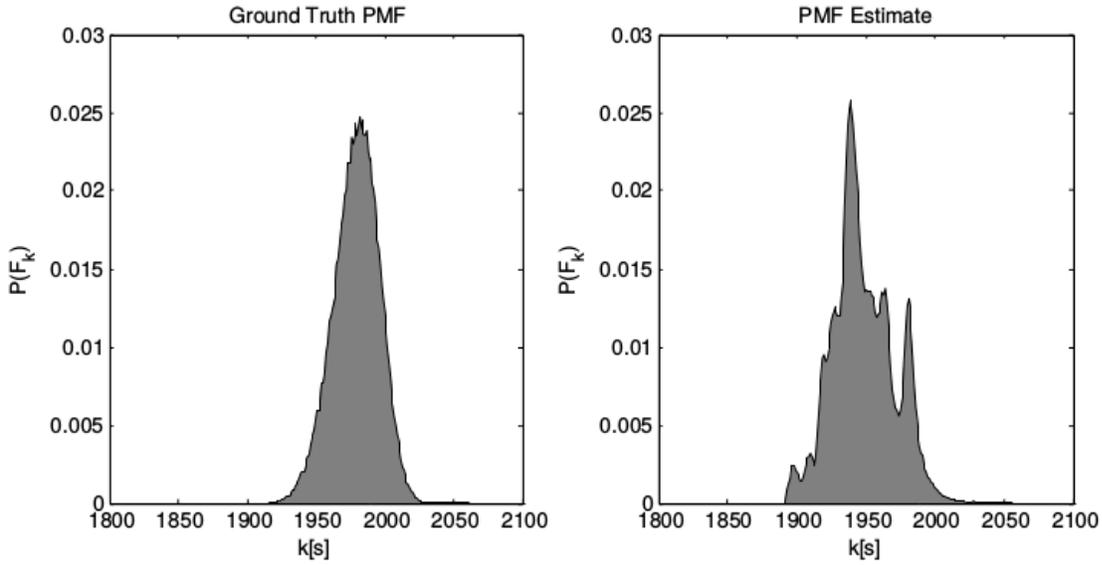


Figura 4.5: Curvas promedio de PMFs para el EoD producto de 30 realizaciones considerando una población de 90 partículas y el ajuste de GMMs de 10 componentes generadas a partir de la solución analítica aproximada (izquierda) y del algoritmo basado en sigma-points (derecha).

Promedios de Estadísticos de la PMF (Solución Analítica Aproximada)					Promedios de Estadísticos de la PMF (Algoritmo Propuesto)				
$\mathbb{E}\{\cdot\}$	$\mathbb{V}ar\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$	$\mathbb{E}\{\cdot\}$	$\mathbb{V}ar\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$
1979.2	141.1	1958.7	1963.2	1966.5	1948.3	49.9	1940.4	1941.6	1942.4

Tabla 4.3: Valores promedio de los estadísticos producto de 30 realizaciones.

Desviaciones Estándar de Estadísticos de la PMF (Solución Analítica Aproximada)					Desviaciones Estándar de Estadísticos de la PMF (Algoritmo Propuesto)				
$\mathbb{E}\{\cdot\}$	$\mathbb{V}ar\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$	$\mathbb{E}\{\cdot\}$	$\mathbb{V}ar\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$
11.6282	29.1979	11.7042	11.5600	11.5810	21.4768	32.6529	22.6816	22.8360	22.6829

Tabla 4.4: Desviación estándar de los estadísticos producto de 30 realizaciones.

Caso 3: $N_k=135$ y $M_k=15$, $\forall k \geq k_p$

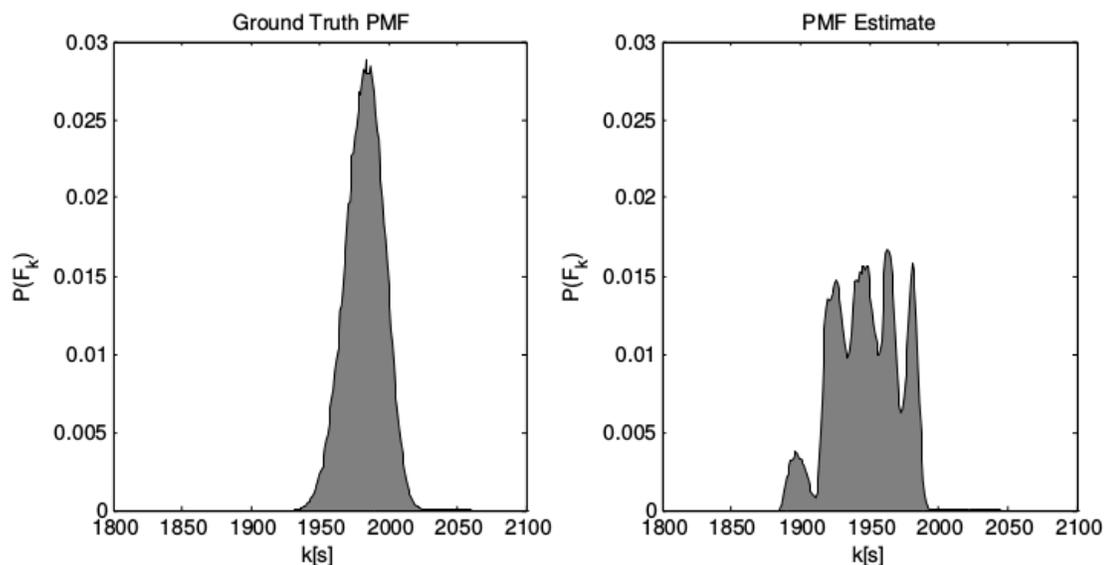


Figura 4.6: Curvas promedio de PMFs para el EoD producto de 30 realizaciones considerando una población de 135 partículas y el ajuste de GMMs de 15 componentes generadas a partir de la solución analítica aproximada (izquierda) y del algoritmo basado en sigma-points (derecha).

Promedios de Estadísticos de la PMF (Solución Analítica Aproximada)					Promedios de Estadísticos de la PMF (Algoritmo Propuesto)				
$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$	$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$
1982.6	134.9	1962.4	1966.9	1970.2	1920.9	82.5	1911.2	1912.3	1913.4

Tabla 4.5: Valores promedio de los estadísticos producto de 30 realizaciones.

Desviaciones Estándar de Estadísticos de la PMF (Solución Analítica Aproximada)					Desviaciones Estándar de Estadísticos de la PMF (Algoritmo Propuesto)				
$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$	$\mathbb{E}\{\cdot\}$	$\text{Var}\{\cdot\}$	$JITP_{5\%}$	$JITP_{10\%}$	$JITP_{15\%}$
7.5759	23.5448	7.8717	7.7768	7.6369	145.1320	77.7959	147.2757	147.2730	147.0628

Tabla 4.6: Desviación estándar de los estadísticos producto de 30 realizaciones.

Solamente en términos de la capacidad del algoritmo propuesto de aproximar la verdadera PMF para el EoD, es posible notar dos cosas en general. Por una parte, en los tres casos, se obtienen PMFs que dependen fuertemente de las condiciones iniciales del algoritmo. Dichas condiciones iniciales son provistas por la caracterización basada en PF de la posterior PDF del vector de estados. Por otra parte, se puede notar que la desviación estándar de los estadísticos decrece a medida que se considera un número de partículas mayor. Esto es de esperarse por la teoría que sustenta el uso de PFs, que sostiene que una de las condiciones que tiene que tenerse para que la PDF a posteriori empírica para los estados converja a la PDF buscada es que el número de partículas tienda al infinito.

De las Figs. 4.4-4.6 y las Tablas 4.1-4.6 es posible notar que el Caso 2 muestra un mejor desempeño respecto de los Casos 1 y 3. Considerando la literatura sobre algoritmos de pronóstico basados en PF, la cual establece que la diferencia entre 50 y 100 partículas es despreciable para este tipo de aplicaciones, se puede establecer que el número de componentes kernel usados en las GMMs es crítico para un diseño apropiado. En el Caso 1, sólo 5 componentes kernel son usados en las GMMs (insuficientes para obtener una adecuada, robusta caracterización). En el Caso 3 (15 componentes kernel en las GMMs) hay evidencia que indica una sobre-parametrización.

En cuanto al estadístico de la esperanza, en todos los casos el algoritmo presenta un sesgo que se antepone al valor esperado de la solución analítica aproximada, siendo el Caso 1 el menos sesgado. Por otra parte, en cuanto al estadístico de la varianza se tiene en todos los casos un valor inferior al analítico, lo cual evidencia que se requiere una mejor parametrización del algoritmo, siendo el Caso 3 el que posee mejor desempeño.

A pesar de que en general es muy importante la estimación tanto de la esperanza como de la varianza de la PMF para el ToF (que se interpreta como EoD en este caso de estudio), hay que notar un hecho importantísimo. La PMF para el ToF es construida secuencialmente al modelar el ToF como un proceso estocástico de Bernoulli no estacionario; ver Capítulo 3. En consecuencia, la PMF para el ToF se puede definir como una *secuencia infinita de medidas aleatorias cuya precisión decae en el tiempo*. Cada una de éstas medidas aleatorias corresponden al parámetro “ p ” de un proceso de Bernoulli que debe calcularse en base a todo el historial antecedente de pronóstico, y por tanto, mientras más lejos del instante inicial de pronóstico k_p se estimen dichos parámetros, éstos se calculan con menor precisión. Es decir, la PMF para el ToF es más verosímil en los instantes más cercanos a k_p y se vuelve menos confiable conforme se avanza hasta el horizonte de predicción. Éste hecho atribuye gran relevancia a los estadísticos $JITP_{\alpha\%}$, pues sólo se calculan en base a los valores de la PMF calculados con mayor confiabilidad e ignoran el resto, que no es el caso de estadísticos como la esperanza y la varianza. En cuanto a los estadísticos $JITP_{\alpha\%}$, los Casos 1 y 2 son los que mejor desempeño muestran. El Caso 1 es el que tiene mayor exactitud, sin embargo en el Caso 2 se tienen resultados más robustos al tener una mejor precisión. Las diferencias en términos del sesgo de las PMFs para el EoD son despreciables teniendo en cuenta el horizonte de predicción (1500 instantes de tiempo), en todos los casos, pues es un hecho que debe ser enfatizado dado que permite una implementación confiable en el marco de pronóstico en línea basado en el algoritmo propuesto.

Conclusión

En la presente Tesis se consigue satisfactoriamente cumplir con los objetivos generales, al proponer metodologías de caracterización de incertidumbre no Gaussiana y que, entre ellas, permiten propagar incertidumbre en el marco de pronóstico. En cuanto los objetivos específicos se tiene lo siguiente.

Primeramente se propone un esquema basado en PF con imputaciones múltiples para el manejo de incertidumbre cuando hay pérdida parcia de datos en la etapa de estimación. Bajo éste enfoque, la Teoría de Imputaciones Múltiples es el núcleo central para la caracterización de incertidumbre. Se presenta una implementación en particular para la estimación de SoC cuando se pierden mediciones de voltaje que se pierden consecutivamente en ventanas de tiempo. Con ello se presentan resultados preliminares que evidencian que la metodología logra incorporar exitosamente la incertidumbre al incrementar el número de partículas del algoritmo, incorporando un paso de reducción o simplificación para evitar su incremento exponencial. El caso de estudio para probar el algoritmo incluye una ventana de pérdida de datos cuando el SoC se encuentra por sobre un 20% de la capacidad de almacenamiento de la batería. Ésto permite adoptar un modo simplificado de reducción de partículas en el algoritmo basándose en la hipótesis de que la impedancia interna de la batería permanece constante. La estrategia MI es comparada al caso donde el filtro tiene acceso al conjunto de datos completo, y también se compara con un algoritmo de predicción basado en PFs que predice a 1000-pasos cuando se pierden datos. Aunque las estimaciones del MI PF son similares a las generadas usando el algoritmo de predicción a 1000-pasos en términos de valor esperado, las diferencias principales entre éstas dos metodologías se encuentran en el modo en que se caracteriza la incertidumbre. Si la PDF a priori se caracteriza bajo el concepto de *evolución artificial* (como en el caso de la predicción a 1000-pasos), la incertidumbre crece a lo largo del tiempo aunque se disponga de información parcial recolectada desde sensores activos. Aunque en este caso se asume un sólo sensor de voltaje para el pack completo de celdas, el potencial una estrategia MI PF no depende solamente en el modelo de evolución artificial, sino que se vuelve más evidente en un caso donde se adquiere información parcial a nivel de las diferentes celdas que componen un pack. En cuanto al impacto en rutinas de pronóstico, el método propuesto no afecta negativamente la precisión de la predicción. Más aún, en casos donde la evolución de que al menos uno de los estados se base en el concepto de evolución artificial, ayuda a acotar la incertidumbre asociada a predicciones a largo plazo, mejorando así la precisión de resultados de pronóstico.

Como MI ha sido desarrollado originalmente para aplicaciones fuera de línea, existen muchos aspectos que tienen que ser considerandos antes de implementarlo en aplicaciones en

línea. Algunos aspectos incluyen mejoras en el modelo de imputación, estimación adaptativa de un número óptimo de partículas y de cantidad de imputaciones, métodos alternativos de reducción de la población de partículas, mejores formas de caracterizar la evolución temporal de la impedancia interna, evaluación de riesgo, entre otros. Adicionalmente, el desarrollo de una reducción óptima de partículas puede permitir la conexión de redes asíncronas, tratamiento de mediciones perdidas, y pronóstico, por dar algunos ejemplos.

En esta Tesis se presenta una importante corrección en la manera en la que se calcula la PMF del EoL de sistemas, además de su interpretación en el contexto de monitoreo en línea. En éste sentido, la verdadera caracterización probabilística para el EoL se muestra como una PMF la cual da cuenta de la probabilidad de falla de sistemas en general, incluyendo por supuesto a los sistemas no lineales y de ruido no Gaussiano, o incluso sistemas con fenómeno de regeneración. Una demostración de un conjunto de condiciones concluye el hecho de que efectivamente se trata de una medida de probabilidad y, por ende, es posible computar parcialmente la PMF del EoL sin necesidad de normalizaciones.

Finalmente, se presenta además un novedoso algoritmo para pronóstico de fallas basado en la propagación de sigma-points y GMMs. El nuevo algoritmo se construye proponiendo una mejora con respecto a los algoritmos de pronóstico basados en PFs; dado que provee una mejor caracterización de las colas de la PDF subyacente de los estados, mientras busca al mismo tiempo preservar al menos los primeros dos momentos de la distribución y considera una caracterización probabilista de las entradas. Es interesante la idea del empleo de sigma-points ya que permite propagar la incertidumbre posicionando partículas de manera determinística en el espacio-estados atribuyéndoles un peso o importancia apropiada para conservar momentos de la distribución. Se mencionan diversos aspectos a considerar en su implementación y se muestra además una comparación entre el algoritmo propuesto y una versión aproximada de la solución analítica para pronóstico de EoD de celdas de ion-litio. Los resultados muestran que el algoritmo propuesto es sensible tanto a la condición inicial de pronóstico, como también de sus parámetros. Los resultados se estudian en base a estadísticos como la esperanza y $JITP_{\alpha\%}$, encontrándose sesgo y deficiencia en cuanto a la precisión. Sin embargo, los estadísticos $JITP_{\alpha\%}$ muestran valores no tan lejanos de sus versiones analíticas aproximadas y muestran así resultados positivos al ser de alta relevancia para la evaluación de riesgo; inclusive más que la exactitud y la precisión, pues la PMF del EoL es calculada con mayor precisión en la medida en que se computa en instantes más cercanos al del inicio de pronóstico. A lo anterior hay que considerar que el pronóstico se hace a 1500 pasos, por lo que también es un factor no despreciable en el desempeño del algoritmo.

Como trabajo futuro se puede mencionar la extensión del algoritmo MI PF a sistemas más generales, el desarrollo teórico de cotas fundamentales que rijan la estimación de la PMF del EoL para medir desempeño de algoritmos de pronóstico, y también el establecimiento de reglas, ya sea en línea (adaptativas) o fuera de línea, para una correcta parametrización del algoritmo propuesto para realizar pronóstico en base a sigma-points.

Bibliografía

- [1] M. Ackerman, S. Ben-David, S. Brânzei, and D. Loker. Weighted clustering. *Proc. 26th AAAI Conference on Artificial Intelligence*, 2012.
- [2] D. Acuña, M. Orchard, J. Silva, and A. Pérez. Multiple-imputation-particle-filtering for uncertainty characterization in battery state-of-charge estimation problems with missing measurement data: Performance analysis and impact on prognostic algorithms. *International Journal of Prognostics and Health Management*, 6(008):1–12, 2015.
- [3] J. Candy. *Bayesian Signal Processing: Classical, Modern and Particle Filtering Methods*. Wiley, 2009.
- [4] C. Chen, G. Vachtsevanos, and M. Orchard. Machine remaining useful life prediction: an integrated adaptive neuro-fuzzy and high-order particle filtering approach. *Mechanical Systems and Signal Processing*, 28:597–607, 2012.
- [5] A. Coppe, R. T. Haftka, N. H. Kim, and F.-G. Yuan. Uncertainty reduction of damage growth properties using structural health monitoring. *Journal of Aircraft*, 47(6):2030–2038, 2010.
- [6] D. Crisan and A. Doucet. A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50(3):736–746, 2002.
- [7] D. Crouse, P. Willett, K. Pattipati, and L. Svensson. A look at gaussian mixture reduction algorithms. *2011 Proceedings of the 14th International Conference on Information Fusion (FUSION)*, pages 1–8, 2011.
- [8] M. Daigle and K. Goebel. Model-based prognostics under limited sensing. *2010 IEEE Aerospace Conference*, pages 1–12, 2010.
- [9] M. Daigle, A. Saxena, and K. Goebel. An efficient deterministic approach to model-based prediction uncertainty estimation. *Annual Conference of the Prognostics and Health Management Society 2012*, pages 326–335, 2012.
- [10] A. Dawn, J.-H. Choi, and N. H. Kim. Prognostics 101: a tutorial for particle filter-based prognostics algorithm using matlab. *Reliability Engineering and System Safety*, pages 161–169, 2013.
- [11] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in*

Practice. New York: Springer-Verlag, 2001.

- [12] A. Doucet, S. J. Godsill, and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.
- [13] D. Dubois. Possibility theory and statistical reasoning. *Computational Statistics & Data Analysis*, 51:47–96, 2006.
- [14] J. Duník, O. Straka, and M. Šimandl. Sigma-point set rotation in unscented kalman filter: Analysis and adaptation. *Proceedings of the 19th IFAC World Congress*, pages 5951–5956, 2014.
- [15] S. Engel, B. Gilmartin, K. Bongort, and A. Hess. Prognostics, the real issues involved with predicting life remaining. *Proc. 2000 IEEE Aerospace Conf.*, 6:457–469, 2000.
- [16] R. Faragher. Understanding the basis of the kalman filter via a simple and intuitive derivation [lecture notes]. *IEEE Signal Processing Magazine*, 29(5):128–132, 2012.
- [17] C. R. Farrar and N. A. Lieven. Damage prognosis: the future of structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):623–632, 2007.
- [18] J. Goldberger, H. Greenspan, and J. Dreyfuss. Simplifying mixture models using the unscented transform. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 30(8):1–7, 2008.
- [19] J. Graham, A. Olchowski, and T. Gilreath. How many imputations are really needed? some practical clarifications of multiple imputation theory. *Prevention Science*, 8:206–213, 2007.
- [20] J. Gu, D. Barker, and M. Pecht. Uncertainty assessment of prognostics of electronics subject to random vibration. *AAAI fall symposium on artificial intelligence for prognostics*, pages 50–57, 2007.
- [21] A. Housfater, X. Zhang, and Y. Zhou. Nonlinear fusion of multiple sensors with missing data. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 4:961–964, 2006.
- [22] Y. Hu, P. Baraldi, F. Di Maio, and E. Zio. A particle filtering and kernel smoothing-based approach for new design component prognostics. *Reliability Engineering and System Safety*, pages 19–31, 2014.
- [23] S. J. Julier. The scaled unscented transformation. *Proc. Am. Control Conf. (ACC 02)*, pages 4555–4559, 2002.
- [24] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.
- [25] R. E. Kalman. A new approach to linear filtering and prediction problems. *J. Basic*

Eng., 82(1):35–45, 1960.

- [26] T. Li, S. Sun, , and J. Duan. Monte carlo localization for mobile robot using adaptive particle merging and splitting technique. *IEEE Int. Conf. Information and Automation*, pages 1913–1918, 2010.
- [27] T. Li, S. Sun, T. P. Sattar, and J. M. Corchado. Fight sample degeneracy and impoverishment in particle filters: A review of intelligent approaches. *Expert Systems With Applications*, 41(8):3944–3954, 2014.
- [28] H. Liao, W. Zhao, and H. Guo. Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model. *Reliability and Maintainability Symposium, 2006. RAMS'06*, pages 127–132, 2006.
- [29] D. Liu, V. Luo, and Y. Peng. Uncertainty processing in prognostics and health management: An overview. *2012 IEEE Conference on Prognostics and System Health Management (PHM)*, 2012.
- [30] J. Liu, A. Kong, and W. Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, 1994.
- [31] J. Liu and M. West. Combined parameter and state estimation in simulation-based filtering. *Sequential Monte Carlo Methods in Practice*, 2001.
- [32] D. J. C. MacKay. Introduction to monte carlo methods. *Learning in Graphical Models*, 1998.
- [33] G. J. McLachlan and T. Krishnan. *The EM algorithm and extensions*. NJ: Wiley, Hoboken, 2008.
- [34] H. Navarrete. *Caracterización Estadística del Perfil de Uso de Baterías para el Prognóstico del Estado-de-Carga*. Memoria de Título. Department of Electrical Engineering, Universidad de Chile, 2014.
- [35] B. E. Olivares, M. A. Cerda, M. E. Orchard, and J. F. Silva. Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. *IEEE Transactions on Instrumentation and Measurement*, 62(2):364–376, 2013.
- [36] M. Orchard, M. Cerda, B. Olivares, and J. Silva. Sequential monte carlo methods for discharge time prognosis in lithium-ion batteries. *International Journal of Prognostics and Health Management*, 3:1–12, 2012.
- [37] M. Orchard, G. Kacprzyński, K. Goebel, B. Saha, and G. Vachtsevanos. Advances in uncertainty representation and management for particle filtering applied to prognostics. *Proc. Int. Conf. Prognostics and Health Manag.*, pages 1–6, 2008.
- [38] M. Orchard, F. Tobar, and G. Vachtsevanos. Outer feedback correction loops in particle filtering-based prognostic algorithms: Statistical performance comparison. *Studies in*

Informatics and Control, 18(4):295–304, 2009.

- [39] M. Orchard and G. Vachtsevanos. A particle-filtering approach for on-line fault diagnosis and failure prognosis. *Trans. Inst. Meas. Control* 2009, 31:221–246, 2009.
- [40] B. Pattipati, C. Sankavaram, and K. Pattipati. System identification and estimation framework for pivotal automotive battery management system characteristics. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 41(6):869–884, 2011.
- [41] D. Pola. *An improved prognosis strategy with temperature-dependent state space models for the analysis of the state-of-health and state-of-charge in lithium-ion batteries*. M.Sc. Tesis. Department of Electrical Engineering, Universidad de Chile, 2014.
- [42] D. A. Pola, H. F. Navarrete, M. E. Orchard, R. S. Rabié, M. A. Cerda, B. E. Olivares, J. F. Silva, P. A. Espinoza, and A. Pérez. Particle-filtering-based discharge time prognosis for lithium-ion batteries with a statistical characterization of use profiles. *IEEE Transactions on Reliability*, 64(2):710–721, 2015.
- [43] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [44] D. Rubin. *Multiple Imputation for Nonresponse in Surveys*. Wiley, 1987.
- [45] B. Saha and K. Goebel. Uncertainty management for diagnostics and prognostics of batteries using bayesian techniques. *2008 IEEE Aerospace Conference*, pages 1–8, 2008.
- [46] S. Sankararaman, M. Daigle, A. Saxena, and K. Goebel. Analytical algorithms to quantify the uncertainty in remaining useful life prediction. *2013 IEEE Aerospace Conference*, pages 1–11, 2013.
- [47] S. Sankararaman and K. Goebel. Why is the remaining useful life prediction uncertain? *Annual Conference of the Prognostics and Health Management Society 2013*, 2013.
- [48] S. Sankararaman, Y. Ling, and S. Mahadevan. Uncertainty quantification and model validation of fatigue crack growth prediction. *Engineering Fracture Mechanics*, 78(7):1487–1504, 2011.
- [49] S. Sankararaman, Y. Ling, C. Shantz, and S. Mahadevan. Uncertainty quantification in fatigue crack growth prognosis. *International Journal of Prognostics and Health Management*, 2(1), 2011.
- [50] A. S. Sarathi, B. Long, and M. Pecht. Prognostics method for analog electronic circuits. *Annual Conference of the Prognostics and Health Management Society 2012*, pages 1–7, 2012.
- [51] A. Saxena, J. Celaya, B. Saha, S. Saha, and K. Goebel. Evaluating prognostics performance for algorithms incorporating uncertainty estimates. *Proc. 2010 IEEE Aerospace Conf.*, pages 1–11, 2010.

- [52] G. Shafer. *A Mathematical Theory of Evidence*. N.J: Princeton University Press, Princeton, 1976.
- [53] L. Tang, M. E. Orchard, K. Goebel, and G. Vachtsevanos. Novel metrics and methodologies for the verification and validation of prognostic algorithms. *Proc. IEEE Aerosp. Conf.*, pages 5–12, 2011.
- [54] R. van der Merwe, A. Doucet, N. de Freitas, and E. Wan. The unscented particle filter. Technical report CUED/F-INFENG/TR 380, Engineering Department, Cambridge University, 2000.
- [55] R. van der Merwe and E. Wan. Gaussian mixture sigma-point particle filters for sequential probabilistic inference in dynamic state-space models. *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 6:701–704, 2003.
- [56] M. Yu, D. Wang, A. Ukil, V. Vaiyapuri, and N. Sivakumar. Model-based failure prediction for electric machines using particle filter. *13th International Conference on Control, Automation, Robotics & Vision*, pages 1811–1816, 2014.
- [57] L. A. Zadeh. Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications*, 23(2):421–427, 1968.
- [58] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, pages 9–34, 1977.
- [59] X. Zhang, A. Khwaja, J. Luo, and A. Housfater. Convergence analysis of multiple imputations particle filters for dealing with missing data in nonlinear problems. *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2014.
- [60] X. Zhang, A. Khwaja, J. Luo, A. Housfater, and A. Anpalagan. Multiple imputations particle filters: Convergence and performance analyses for nonlinear state estimation with missing data. *IEEE Journal of Selected Topics in Signal Processing*, pages 1–12, 2015.
- [61] E. Zio and G. Pelsoni. Particle filtering prognostic estimation of the remaining useful life of nonlinear components. *Reliability Engineering and System Safety*, pages 403–409, 2010.