Reconstructing Semi-Markov Maps with Arbitrary Branch Monotonicity from Density Sequences

André M. McDonald, *Member, IEEE* and Michaël A. van Wyk, *Senior Member, IEEE*

***Abstract*—This paper presents a novel solution to the inverse Frobenius-Perron problem of reconstructing an unknown nonlinear and ergodic map from causal sequences of probability density functions generated by the map. The original solution to this problem successfully reconstructs members of the *canonical map class* (i.e., a subset of the piecewise linear semi-Markov maps), provided all the map's branches are monotonically increasing. The original solution constructs a matrix estimate of the map's Frobenius-Perron operator, which governs the evolution of density functions under iteration of the map, from the density sequences. The one-dimensional map is reconstructed from this matrix. In contrast, the proposed solution constructs a higher-order matrix estimate of the Frobenius-Perron operator. A member of the newly proposed class of *generalized hat maps*, a superset of the canonical maps, is constructed from this matrix estimate. The proposed solution successfully distinguishes between increasing and decreasing map branches and enlarges the class of maps that can be successfully reconstructed to canonical maps with any subset of decreasing branches. When used to reconstruct any piecewise linear semi-Markov map, the proposed solution generates a map with consistent invariant density and power spectrum mode characteristics, regardless of the unknown map's canonicity or branch monotonicity. Numerical examples that illustrate the proposed solution's characteristics are presented.**

***Index Terms*—Inverse Frobenius-Perron problem, system identification, nonlinear system, piecewise defined maps, Markov map, Frobenius-Perron operator, power spectral density.**

I. Introduction

THE probabilistic approach to characterizing nonlinear and deterministic dynamical systems with complex and ostensibly random behavior gained popularity in the latter half of the 20$^{\text{th}}$ century [1], [2]. Advances in this domain spurred the development of practical and accurate methods for estimating the invariant state density and power spectral density of ergodic systems [3], [4]. However, the probabilistic modeling problem has attracted little attention compared to the probabilistic analysis problem. This related problem requires the construction of an ergodic dynamical system model with prescribed statistical properties, and is referred to collectively as the inverse Frobenius-Perron problem (IFPP) [5].

A. M. McDonald is with the Defence and Security Cluster, Council for Scientific and Industrial Research, Pretoria, South Africa and the School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa.

M. A. van Wyk is with the School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa and the Department of Electrical Engineering, City University of Hong Kong, Hong Kong SAR, China.

The IFPP first appeared in the literature in the 1970s [6]. Initial work towards solving the IFPP focused on a particular problem formulation that requires the construction of an ergodic one-dimensional map with a prescribed invariant state density function [7], [8]. Over time, it was discovered that solutions to this original problem often provide insight into bifurcation routes, the transition from the non-chaotic to the chaotic regime, and how structural features of a map may give rise to certain statistical characteristics [9], [10]. Mature solutions to the original problem have been used not only to model the motion of fluids [11] and neural processes pertaining to the olfactory system [12], but also to design more robust cogged bits for rock drills [13] and more efficient pseudorandom signal generators [14].

A new formulation of the IFPP recently appeared in the literature [15], [16]; we refer to this problem as IFPP-III for consistency with [5]. IFPP-III requires the reconstruction of an unknown map that coincides with the evolution rule of an ergodic one-dimensional dynamical system. The map is to be reconstructed using measurements of the system state, which are collected during several experiments wherein the system is probed. In each of these *probing experiments*, an ensemble of states with values drawn from a selected probability density (the *probing density*) is permitted to evolve over time, and all states that constitute the ensemble are measured at several evenly spaced instants in time. However, the problem assumes that it is either impractical or physically impossible to isolate and follow individual state trajectories from one time instant to the next. IFPP-III thus mandates the estimation of the probability density function corresponding to each collection of measurements, thereby producing a causal sequence of state densities for each probing experiment. The unknown map is to be reconstructed from these density sequences. Solutions to IFPP-III are relevant in applications where state trajectories cannot be distinguished from one another, such as the modeling of emulsion polymerization processes [17], papermaking systems [18], and certain communication systems [19].

Nie and Coca [15], [16] proposed the first solution to IFPP-III. This solution is used to reconstruct ergodic piecewise linear (PWL) maps that are semi-Markov [8] over a partition of the map domain (the *Markov partition*); this is referred to as the class of ergodic PWL *semi-Markov maps*. The operator that governs the evolution of state density functions under iteration of a given map, known as the *Frobenius-Perron operator* [1], may be represented as a matrix if attention is restricted to PWL semi-Markov maps and density functions

that are piecewise constant (PWC) over the Markov partition. This matrix, referred to as the *Frobenius-Perron matrix* (FPM) of the map, is central to the solution of IFPP-III.

In the solution of [15], [16], the Markov partition of the unknown map is first estimated. Probing experiments are then carried out with probing density functions that are PWC over the estimated Markov partition. Using the density sequences obtained from measurements collected in these experiments, an estimate of the unknown map's FPM is constructed. IFPP-III is then solved by carrying out the final step of constructing a PWL map that is semi-Markov over the estimated Markov partition, under the condition that the FPM of this map matches the estimate of the unknown map's FPM.

The motivation for the work presented in this paper follows from the observation that multiple distinct PWL semi-Markov maps can generally be constructed to possess the same prescribed FPM [8]. For example, if any branch of a map in this class is redefined such that its slope changes from positive to negative (or vice versa), while simultaneously preserving its domain and codomain, the map retains its FPM. Thus, the problem of constructing a map with a prescribed FPM inherently provides freedom to choose certain structural characteristics of the map. When using any particular method for constructing such a map to solve IFPP-III, the underlying choices that were made regarding these characteristics become assumptions regarding the unknown map's structure and restrict the class of maps that can be successfully reconstructed, regardless of the accuracy of the Markov partition and density estimates. The method used in [15], [16] constructs a map with a specific branch count and order (we refer to these maps as *canonical*, and define this class formally in section III-A3), and selects a positive slope for each branch function. Hence, if we disregard estimation error and assume that the estimate of the unknown map's FPM is perfect, the class of maps reconstructible using this solution is the set of PWL semi-Markov maps that are canonical and possess only increasing branches. However, canonicity and possessing only increasing branches are not necessary conditions for the consistency of the constructed and unknown maps' invariant densities.

This paper presents a novel solution to IFPP-III. The solution we propose is based on the observation that the original solution's inability to discern the monotonicity of each map branch (i.e., increasing or decreasing) is due to the elements of the FPM being dependent only on the absolute values of the branch slopes. We show that a higher-order matrix representation of the map's Frobenius-Perron operator, which characterizes the evolution of densities that are PWL over the Markov partition, contains elements that are dependent on the monotonicity of the branches. The proposed solution estimates a predetermined subset of the elements in the higher-order FPM from density sequences obtained using PWC and PWL probing densities. A map from the newly proposed class of generalized hat maps, which is a superset of the canonical maps, is then constructed such that the same elements in its higher-order FPM match those that were estimated.

Using a series of propositions and two numerical examples, we show that the proposed solution is able to successfully reconstruct maps from the larger class of canonical PWL semi-Markov maps (i.e., independent of branch monotonicity). In contrast to the original solution, it is shown that the invariant density *and* the power spectrum mode[1] characteristics (i.e., mode center frequencies and bandwidths) of maps constructed using the proposed solution are also consistent with those of any unknown PWL semi-Markov map, even if it is not canonical or possesses decreasing branch functions.

The remainder of this paper is set out as follows. In section II we present the necessary mathematical background. In section III the original solution of [15], [16] is detailed and several observations are made regarding its characteristics. The proposed solution is presented and characterized in section IV. The numerical examples are presented in section V and conclusions are drawn in section VI.

II. Mathematical background

Consider a one-dimensional dynamical system with nonlinear evolution rule $S : \Omega \to \Omega$, where $\Omega \triangleq [0, 1]$ is the state space. Let $P(\mathcal{Q}) = \{q_1, q_2, \ldots, q_{M+1}\}$ denote the partition points of the partition $\mathcal{Q} = \{Q_1, Q_2, \ldots, Q_M\}$ of Ω, where $Q_m = (q_m, q_{m+1})$ denotes the partition intervals and $0 = q_1 < q_2 < \ldots < q_{M+1} = 1$ with integer $M > 1$. We restrict our attention to maps S that belong to the class of piecewise monotonic maps $\Gamma(\mathcal{Q})$.

Definition II.1 (Class $\Gamma(\mathcal{Q})$)**.** *A surjective map $S : \Omega \to \Omega$ is piecewise monotonic [2] (i.e., $S \in \Gamma(\mathcal{Q})$) if there exists a partition $\mathcal{Q}$ of the unit interval such that (i) each branch function $S_m \triangleq S|_{(q_m, q_{m+1})}$ is a C^r function, for some $r \geq 1$, that can be extended to a C^r function over $[q_m, q_{m+1}]$, and (ii) $|S'_m| > 0$ for $m = 1, 2, \ldots, M$, where $S'_m \triangleq dS_m/dx$.*

We characterize $S \in \Gamma(\mathcal{Q})$ from a probabilistic perspective. Let X_1 denote a real random variable on the probability space $(\Omega, \mathcal{B}, P_1)$, which represents the initial state of the dynamical system. Here, $\mathcal{B}$ denotes the Borel σ-algebra on Ω and P_1 is a probability measure. Since S is surjective and measurable on $\mathcal{B}$ [5], the random variables $X_2, X_3, \ldots$, defined by $X_{t+1} = S(X_t)$ for $t \geq 1$, have corresponding probability measures $P_{t+1}(B) = P_t(S^{-1}(B))$, $B \in \mathcal{B}$.

Let P_1 be absolutely continuous with respect to the Borel measure μ. The Radon-Nikodym theorem then implies the existence of a function $f_1 \in \mathcal{D}$ associated with X_1, where $\mathcal{D} \triangleq \{f \in L^1(\Omega) : f \geq 0 \text{ and } ||f|| = 1\}$ denotes the set of all probability density functions (we also use the term *densities*) on Ω. The density f_1 satisfies $P_1(B) = \int_B f_1(x)d\mu(x)$ for all $B \in \mathcal{B}$. Since $S \in \Gamma(\Omega)$ is nonsingular with respect to μ [5], each measure P_t, where $t > 1$, is also absolutely continuous with respect to μ, and may be associated with a corresponding density $f_t \in \mathcal{D}$. The linear operator $\mathcal{P}_S : L^1(\Omega) \to L^1(\Omega)$, referred to as the Frobenius-Perron operator (FPO) of S, characterizes the evolution of densities under iteration of S according to the expression $f_{t+1} = \mathcal{P}_S f_t$ for $t \geq 1$. The FPO is uniquely defined by the expression

$$\int_B (\mathcal{P}_S f)(x)d\mu(x) = \int_{S^{-1}(B)} f(x)d\mu(x), \quad B \in \mathcal{B}. \quad (1)$$

[1]The power spectral density of an ergodic PWL semi-Markov map is, in general, a superposition of spectral modes.

The FPO of a map $S \in \Gamma(\Omega)$ has the explicit representation

$$(\mathcal{P}_S f_t)(x) = \sum_{m \in N(x)} \frac{f_t \circ S_m^{-1}(x)}{|S'_m \circ S_m^{-1}(x)|}, \tag{2}$$

where $N(x) \triangleq \{m : x \in S(Q_m)\}$ [2].

We next introduce a subset of the piecewise monotonic maps, referred to as the class of semi-Markov maps [8]. The structure of these maps accommodates matrix solutions to the IFPP. Semi-Markov maps are defined over a two-level hierarchy of partitions. Let $\mathcal{Q}$ denote a partition of the unit interval as before, and let the non-empty set of open intervals $\mathcal{R}^{(m)} \triangleq \{R_1^{(m)}, R_2^{(m)}, \ldots, R_{c(m)}^{(m)}\}$ constitute a partition of the closure of Q_m for each $m = 1, 2, \ldots, M$. We refer to the unit-interval partition $\mathcal{R} = \{R_j^{(m)}\}$, $m = 1, 2, \ldots, M$ and $j = 1, 2, \ldots, c(m)$, as a *sub-partition* of $\mathcal{Q}$. The $\mathcal{R}(\mathcal{Q})$-semi-Markov maps are then defined as follows (an example is plotted in Fig. 1).

Definition II.2 ($\mathcal{R}(\mathcal{Q})$-semi-Markov map)**.** *Let $\mathcal{Q}$ denote a partition of Ω, and let $\mathcal{R}$ denote a sub-partition of $\mathcal{Q}$. The map $S \in \Gamma(\mathcal{R})$ is $\mathcal{R}(\mathcal{Q})$-semi-Markov if the image of each interval $R \in \mathcal{R}$ under S is an interval of the partition $\mathcal{Q}$; i.e.,*

$$S(R) \in \mathcal{Q}, \quad R \in \mathcal{R}. \tag{3}$$

Consider an $\mathcal{R}(\mathcal{Q})$-semi-Markov map S with linear branch functions $S_{j,m} \triangleq S|_{R_j^{(m)}}$. Let $\mathcal{D}_{\mathcal{Q}}^{(\ell)} \subset \mathcal{D}$ denote the set of density functions that are defined piecewise over $\mathcal{Q}$, such that $f_t \in \mathcal{D}_{\mathcal{Q}}^{(\ell)}$ implies $f_t|_{Q_m}$ is a polynomial function with degree not exceeding $\ell \in \mathbb{N}_0$ for each $m = 1, 2, \ldots, M$. Substitution of $f_t \in \mathcal{D}_{\mathcal{Q}}^{(\ell)}$ into (2) reveals that the density $f_{t+1} = \mathcal{P}_S(f_t)$ is also a member of $\mathcal{D}_{\mathcal{Q}}^{(\ell)}$. Hence, the evolution of these densities can be characterized using matrix equations. Specifically, with a PWL $\mathcal{R}(\mathcal{Q})$-semi-Markov map S and $f_t \in \mathcal{D}_{\mathcal{Q}}^{(\ell)}$ we have

$$\mathbf{f}_{t+1} = \mathbf{P}_S^{(\ell)} \mathbf{f}_t, \tag{4}$$

where $\mathbf{f}_t$ is the order-ℓ vector representation of the density function f_t and $\mathbf{P}_S^{(\ell)}$ is the order-ℓ matrix representation of the FPO $\mathcal{P}_S$. We refer to $\mathbf{P}_S^{(\ell)}$ as the order-ℓ FPM of S. The vectors in (4) are defined as $\mathbf{f}_t \triangleq [\mathbf{f}_{t,\ell}, \mathbf{f}_{t,\ell-1}, \ldots, \mathbf{f}_{t,0}]^T$, where $\mathbf{f}_{t,\ell} \triangleq [f_{t,\ell,1}, f_{t,\ell,2}, \ldots, f_{t,\ell,M}]$ and

$$f_{t,\ell,m} = \frac{1}{\ell!} \frac{d^\ell}{dx^\ell} f_t|_{Q_m}(x)\bigg|_{x=0}. \tag{5}$$

We provide expressions for the order-zero and order-one FPMs, which characterize the evolution of PWC and PWL density functions, respectively. Let S be given by

$$S(x) = \sum_{m=1}^{M} \sum_{j=1}^{c(m)} S_{j,m}(x) \chi_{R_j^{(m)}}(x), \tag{6}$$

where

$$\chi_A(x) \triangleq \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \tag{7}$$

Furthermore, let each branch function $S_{j,m}$ be given by

$$S_{j,m}(x) = v_{j,m} x + w_{j,m}. \tag{8}$$

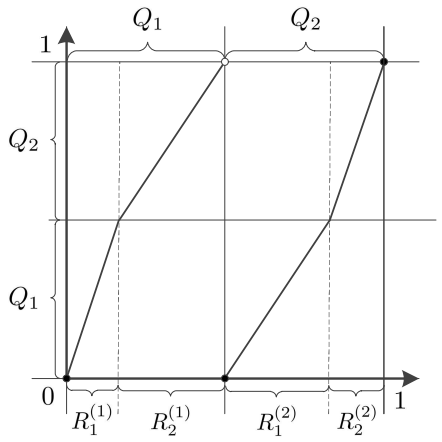

Fig. 1. An example of an $\mathcal{R}(\mathcal{Q})$-semi-Markov map.

Using (2), the order-one FPM $\mathbf{P}_S^{(1)}$ is then derived as

$$\mathbf{P}_S^{(1)} = \begin{pmatrix} \mathbf{Q}_S & \mathbf{0}_{M \times M} \\ \mathbf{R}_S & \mathbf{P}_S^{(0)} \end{pmatrix}, \tag{9}$$

where $\mathbf{Q}_S \triangleq [q_{k,m}]_{k,m=1}^M$, $\mathbf{R}_S \triangleq [r_{k,m}]_{k,m=1}^M$, and $\mathbf{0}_{M \times M}$ denotes the $M \times M$ matrix containing zeros. The elements of the submatrices $\mathbf{Q}_S$ and $\mathbf{R}_S$ are given by

$$q_{k,m} = \begin{cases} \displaystyle\sum_{j \in T(k,m)} s_{j,m} \frac{1}{v_{j,m}^2}, & \text{if } |T(k,m)| > 0, \\ 0, & \text{if } |T(k,m)| = 0, \end{cases} \tag{10}$$

and

$$r_{k,m} = \begin{cases} \displaystyle -\sum_{j \in T(k,m)} s_{j,m} \frac{w_{j,m}}{v_{j,m}^2}, & \text{if } |T(k,m)| > 0, \\ 0, & \text{if } |T(k,m)| = 0, \end{cases} \tag{11}$$

where

$$T(k,m) \triangleq \{j : S(R_j^{(m)}) = Q_k\}. \tag{12}$$

In these expressions, $s_{j,m} \triangleq \operatorname{sgn}(v_{j,m})$, where

$$\operatorname{sgn}(x) \triangleq \begin{cases} -1 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases} \tag{13}$$

The order-zero FPM $\mathbf{P}_S^{(0)} \triangleq [p_{k,m}]_{k,m=1}^M$ has elements

$$p_{k,m} = \begin{cases} \displaystyle\sum_{j \in T(k,m)} |v_{j,m}|^{-1}, & \text{if } |T(k,m)| > 0, \\ 0, & \text{if } |T(k,m)| = 0. \end{cases} \tag{14}$$

The Folklore theorem establishes conditions for both the existence and uniqueness of absolutely continuous invariant (ACI) measures for Markov maps that are piecewise C^2 [2]. Let $S \in \Gamma(\mathcal{R})$ denote a PWL $\mathcal{R}(\mathcal{Q})$-semi-Markov map that is aperiodic and piecewise expanding. By the Folklore theorem, S possesses a unique ACI measure P^*, and is therefore ergodic. Since P^* is absolutely continuous, there exists a corresponding invariant density function f_S^*. This function is a fixed point of the FPO associated with S, and can be chosen as a PWC function with discontinuities only at the partition points $P(\mathcal{Q})$, such that $K^{-1} \leq f_S^* \leq K$ for some $K > 0$.

Let S denote an $\mathcal{R}(\mathcal{Q})$-semi-Markov map that satisfies the conditions of the Folklore theorem. The order-zero FPM of S is similar to a stochastic matrix (i.e., a matrix with nonnegative entries and columns that sum to unity) and has a unity maximal eigenvalue λ_1, referred to as the *Perron eigenvalue* [5]. This eigenvalue is a simple root of the characteristic polynomial of the matrix. It is associated with a left eigenvector $\mathbf{q}^T$ (the left Perron eigenvector) that possesses elements coinciding with the lengths of the intervals in $\mathcal{Q}$, and a right eigenvector $\mathbf{f}_S^*$ (the right Perron eigenvector) that coincides with the unique PWC invariant density function f_S^*. The remaining eigenvalues of the order-zero FPM are strictly smaller than λ_1 in absolute value, which implies the matrix has a spectral radius of unity.

The *time correlation function* ϕ_S of S is defined as [20]

$$\phi_S(\tau) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} S^t(x)S^{t+|\tau|}(x),\ \tau = 0, \pm 1, \pm 2, \ldots, \tag{15}$$

for almost all $x \in \Omega$, where S^t denotes the t-fold composition of S with itself and $S^0(x) \triangleq x$. The corresponding *power spectral density* (PSD) Φ_S is given by

$$\Phi_S(\omega) = \sum_{\tau=-\infty}^{\infty} \phi_S(\tau)e^{-i\omega\tau}, \quad \omega \in [-\pi, \pi). \tag{16}$$

Equality of time and ensemble averages is a consequence of Birkhoff's ergodicity theorem [21]. An expression for ϕ_S may be derived by using this relationship and the theory developed in [4]. It follows that

$$\phi_S(\tau) = \mathbf{g}^T\big(\mathbf{P}_S^{(1)}\big)^{|\tau|}\mathbf{h}, \tag{17}$$

where $\mathbf{h} \triangleq [(\mathbf{f}^*)^T, \mathbf{0}_{1\times M}]^T$ and $(\mathbf{P}_S^{(1)})^0 \triangleq \mathbf{I}$. In this expression $\mathbf{g} \triangleq [\mathbf{g}_3, \mathbf{g}_2]^T$, where the elements of $\mathbf{g}_2 \triangleq [g_{2,1}, g_{2,2}, \ldots, g_{2,M}]$ and $\mathbf{g}_3 \triangleq [g_{3,1}, g_{3,2}, \ldots, g_{3,M}]$ are given by $g_{n,m} \triangleq (q_{m+1}^n - q_m^n)/n$.

The block structure of $\mathbf{P}_S^{(1)}$ in (9) implies that the eigenvalues of this matrix comprise the eigenvalues of $\mathbf{P}_S^{(0)}$ (denoted by $\lambda_1, \lambda_2, \ldots, \lambda_M$) and the eigenvalues of $\mathbf{Q}_S$ (denoted by $\lambda_{M+1}, \lambda_{M+2}, \ldots, \lambda_{2M}$). Now, assume that $\mathbf{P}_S^{(1)}$ is diagonalizable (the general case is considered in [4]), and let $\mathbf{u}_m^T$ and $\mathbf{v}_m$ denote the left and right eigenvectors of $\mathbf{P}_S^{(1)}$ that are associated with eigenvalue λ_m (we assume that the eigenvectors are scaled such that $\mathbf{u}_m^T\mathbf{v}_m = 1$ for all m). Equation (17) may be simplified to obtain

$$\phi_S(\tau) = \sum_{m=1}^{2M} c_m\lambda_m^{|\tau|}, \tag{18}$$

where $c_m = \mathbf{g}^T\mathbf{v}_m\mathbf{u}_m^T\mathbf{h}$. This expression implies that the PSD is a superposition of $2M$ spectral modes with center frequencies equal to $\arg(\lambda_m)$ and bandwidths inversely proportional to $|\lambda_m|$. Modes with $|\lambda_m| = 1$ appear as lines in the PSD, whereas the remaining modes are continuous. In particular,

$$\begin{aligned}\Phi_S(\omega) &= 2\pi \sum_{m:|\lambda_m|=1} c_m\delta(\omega - \arg(\lambda_m)) \\ &\quad + \sum_{m:0\le|\lambda_m|<1} \frac{c_m(1-\lambda_m^2)}{(e^{-j\omega}-\lambda_m)(e^{j\omega}-\lambda_m)},\end{aligned} \tag{19}$$

where $\delta(\cdot)$ denotes the Dirac delta function.

III. Existing work

In this section we first outline the original solution[2] for reconstructing unknown ergodic maps that are PWL and $\mathcal{R}(\mathcal{Q})$-semi-Markov, as proposed in [16]. Thereafter, we consider the characteristics of maps constructed using this solution.

A. Original solution

The original solution constructs an estimate $\hat{S}$ of the unknown map S in three steps, as outlined below. It is assumed that partitions $\mathcal{Q}$ and $\mathcal{R}$ are both unknown.

1) Estimation of the Markov partition: Let $\{x_k^*\}_{k=1}^{K_0}$ denote a set of independent state observations drawn from the invariant state density f_S^* of S (in practice, this set is constructed by allowing an initial ensemble of states drawn from the uniform distribution over Ω to evolve over a sufficient number of time instants under iteration of S). Since f_S^* is PWC over $\mathcal{Q}$, the problem of estimating the Markov partition $\mathcal{Q}$ is formulated as a search for a partition $\hat{\mathcal{Q}} = \{\hat{Q}_1, \hat{Q}_2, \ldots, \hat{Q}_{\hat{M}}\}$ that satisfies

$$f_S^*(x) \approx \sum_{m=1}^{\hat{M}} h_m\chi_{\hat{Q}_m}(x), \tag{20}$$

where h_m is given by

$$h_m = \frac{1}{K_0\mu(\hat{Q}_m)}\sum_{k=1}^{K_0}\chi_{\hat{Q}_m}(x_k^*). \tag{21}$$

The search for $\hat{\mathcal{Q}}$ proceeds as follows (refer to [16] for specifics). A partition $\mathcal{U}$ that is uniform over Ω is constructed, where the number of partition intervals is selected to minimize a penalized log-likelihood function. Next, a histogram of the observations $\{x_k^*\}_{k=1}^{K_0}$ is derived, such that the histogram bins coincide with the intervals of $\mathcal{U}$. Different candidates for $\hat{\mathcal{Q}}$ are constructed by merging neighboring bins of the histogram with absolute height differences that fall below a threshold, and by varying the threshold value. The candidate partition associated with the minimum integrated squared error between the densities coinciding with its histogram and the histogram over the uniform partition is then selected.

2) Estimation of the order-zero FPM: An estimate $\hat{\mathbf{P}}_S^{(0)}$ of the order-zero FPM $\mathbf{P}_S^{(0)}$ belonging to S is obtained from state density function estimates constructed after conducting N probing experiments. In each probing experiment $n = 1, 2, \ldots, N$, an ensemble $\{x_{1,k}^n\}_{k=1}^K$ of K states drawn independently from a probing density f_1^n evolves under iteration of S. The ensemble is measured at time instants $t = 2, \ldots, T$. Let these measurements be denoted by $\{x_{t,k}^n\}_{k=1}^K$.

Each probing density function is selected to be unique and PWC over the estimated Markov partition $\hat{\mathcal{Q}}$. If it was assumed that $\hat{\mathcal{Q}} = \mathcal{Q}$, the evolved density functions f_t^n, $t = 2, 3, \ldots, T$, would also be PWC over $\hat{\mathcal{Q}}$. This observation motivates the

[2]Whereas the original solution was extended in [16] to allow for the reconstruction of a subset of those continuous maps that possess a unique invariant density, in this paper we consider maps that need not be continuous. Solutions to more general formulations of IFPP-III (i.e., for stochastically perturbed systems and systems driven by external inputs) were also presented in [22]–[24]. A summary of these solutions may be found in [5].

construction of density function estimates $\hat{f}_t^n \approx f_t^n$ that are PWC over $\hat{\mathcal{Q}}$. Specifically, the estimates

$$\hat{f}_t^n(x) = \begin{cases} f_1^n(x) & \text{if } t = 1, \\ \sum_{m=1}^{\hat{M}} \hat{f}_{t,0,m} \chi_{\hat{Q}_m}(x) & \text{if } t = 2, 3, \ldots, T, \end{cases} \tag{22}$$

where

$$\hat{f}_{t,0,m} = \frac{1}{K\mu(\hat{Q}_m)} \sum_{k=1}^{K} \chi_{\hat{Q}_m}(x_{t,k}^n), \tag{23}$$

are constructed.

Let $\hat{\mathbf{f}}_t^n \triangleq [\hat{f}_{t,0,1}, \hat{f}_{t,0,2}, \ldots, \hat{f}_{t,0,\hat{M}}]^T$ denote the order-zero vector representation of the density $\hat{f}_t^n$. Furthermore, let

$$\hat{\mathbf{A}} \triangleq [\hat{\mathbf{F}}_1, \hat{\mathbf{F}}_2, \ldots, \hat{\mathbf{F}}_{T-1}] \tag{24}$$

and

$$\hat{\mathbf{B}} \triangleq [\hat{\mathbf{F}}_2, \hat{\mathbf{F}}_3, \ldots, \hat{\mathbf{F}}_T], \tag{25}$$

where $\hat{\mathbf{F}}_t \triangleq [\hat{\mathbf{f}}_t^1, \hat{\mathbf{f}}_t^2, \ldots, \hat{\mathbf{f}}_t^N]$. If it were assumed that $\hat{\mathcal{Q}} = \mathcal{Q}$ and each density estimate satisfies $\hat{f}_t^n = f_t^n$, (4) implies that

$$\hat{\mathbf{B}} = \mathbf{P}_S^{(0)} \hat{\mathbf{A}}. \tag{26}$$

Based on this observation, the estimate $\hat{\mathbf{P}}_S^{(0)}$ is ideally selected as a solution to the constrained optimization problem

$$\mathbf{P}_S^* = \arg\min_{\mathbf{P}} : ||\hat{\mathbf{B}} - \mathbf{P}\hat{\mathbf{A}}||_F, \tag{27}$$

subject to $\hat{\mathbf{q}}^T \mathbf{P} = \hat{\mathbf{q}}^T$ and to the elements of $\mathbf{P} \triangleq [p_{k,m}]_{k,m=1}^{\hat{M}}$ being nonnegative, where $\hat{\mathbf{q}} \triangleq [\mu(\hat{Q}_1), \mu(\hat{Q}_2), \ldots, \mu(\hat{Q}_{\hat{M}})]^T$ and $||\cdot||_F$ denotes the Frobenius norm.

3) Construction of a map with matching order-zero FPM: The final solution step is the construction of a PWL $\hat{\mathcal{R}}(\hat{\mathcal{Q}})$-semi-Markov map $\hat{S}$ with order-zero FPM $\mathbf{P}_{\hat{S}}^{(0)}$ that is equal to the estimate $\hat{\mathbf{P}}_S^{(0)}$ of the unknown map's order-zero FPM. In general, multiple PWL semi-Markov maps may be constructed to possess the same order-zero FPM. The original solution constructs maps with a particular branch structure. We describe these maps as being *canonical* or having a *canonical structure*.

Definition III.1 (Canonical semi-Markov map)**.** *Suppose that $S : \Omega \to \Omega$ is a PWL $\mathcal{R}(\mathcal{Q})$-semi-Markov map and let*

$$R_j^{(m)} \overset{S}{\mapsto} Q_{U(j,m)}, \tag{28}$$

where $U \in \{1, 2, \ldots, M\}$. Map S is canonical iff

$$j_1 < j_2 \implies U(j_1, m) < U(j_2, m) \tag{29}$$

for all $m = 1, 2, \ldots, M$.

Figure 2(a) illustrates the structure of a canonical map by representing the region spanned by each branch function's domain and codomain with a shaded rectangle. The condition of (29) imposes a strictly ascending order, over the unit interval, on the codomains of the intervals in each of the partitions $\mathcal{R}^{(m)} = \{R_1^{(m)}, R_2^{(m)}, \ldots, R_{c(m)}^{(m)}\}$. Non-canonical maps do not possess this property; examples of such maps are provided in Fig. 2(b) and Fig. 2(c).

Certain characteristics of a canonical map's branch functions may be discerned from the map's order-zero FPM. Let $S : \Omega \to \Omega$ denote a PWL and canonical $\mathcal{R}(\mathcal{Q})$-semi-Markov map that satisfies (28). Equation (29) implies that S maps *at most one* interval of each partition $\mathcal{R}^{(m)}$ to each interval Q_k. Equation (14) then implies that S maps *one* interval of $\mathcal{R}^{(m)}$ to Q_k if the k^{th} element in column m of the map's order-zero FPM is nonzero; otherwise, S maps no interval of $\mathcal{R}^{(m)}$ to Q_k. Furthermore, S maps the intervals of each $\mathcal{R}^{(m)}$ to the intervals of $\mathcal{Q}$ such that their codomains are strictly ascending. Let $z_{1,m} < z_{2,m} < \ldots < z_{w(m),m}$ denote the indices of the $w(m)$ nonzero entries in column m of the map's order-zero FPM. It follows that $U(j, m) = z_{j,m}$.

The original solution uses the observations above to construct a canonical map $\hat{S}$ from the estimate $\hat{\mathbf{P}}_S^{(0)}$. A sub-partition $\hat{\mathcal{R}}$ of the estimated Markov partition $\hat{\mathcal{Q}}$ is constructed. This sub-partition consists of the intervals of partitions $\hat{\mathcal{R}}^{(m)} = \{\hat{R}_1^{(m)}, \hat{R}_2^{(m)}, \ldots, \hat{R}_{\hat{c}(m)}^{(m)}\}$, $m = 1, 2, \ldots, \hat{M}$. Each $\hat{\mathcal{R}}^{(m)}$ is a partition of the closure of $\hat{Q}_m$, and is defined with an interval count $\hat{c}(m)$ equal to the number of nonzero entries in column m of $\hat{\mathbf{P}}_S^{(0)}$. The intervals of $\hat{\mathcal{R}}$ are mapped by $\hat{S}$ according to

$$\hat{R}_j^{(m)} \overset{\hat{S}}{\mapsto} \hat{Q}_{\hat{U}(j,m)}, \tag{30}$$

where $\hat{U}(j, m) = \hat{z}_{j,m}$ and $\hat{z}_{j,m}$ denotes the index of the j^{th} nonzero element in column m of $\hat{\mathbf{P}}_S^{(0)}$. The length of interval $\hat{R}_j^{(m)}$ is determined from the corresponding element of $\hat{\mathbf{P}}_S^{(0)}$ and the length of the interval's codomain. Since the order-zero FPM of $\hat{S}$ is to match $\hat{\mathbf{P}}_S^{(0)} = [\hat{p}_{k,m}]_{k,m=1}^{\hat{M}}$, we have

$$\hat{p}_{\hat{U}(j,m),m} = |\hat{S}'_{j,m}|^{-1} \tag{31}$$

and

$$\mu(\hat{R}_j^{(m)}) = \mu(\hat{Q}_{\hat{U}(j,m)}) \hat{p}_{\hat{U}(j,m),m}. \tag{32}$$

The original solution defines each of the branch functions $\hat{S}_{j,m} = \hat{S}|_{\hat{R}_j^{(m)}}$ to be increasing. The branch functions are selected as

$$\hat{S}_{j,m}(x) = \hat{p}_{\hat{U}(j,m),m}^{-1}(x - \hat{r}_{j,m}) + \hat{q}_{\hat{U}(j,m)}, \tag{33}$$

where $\hat{Q}_m = (\hat{q}_m, \hat{q}_{m+1})$, $\hat{R}_j^{(m)} = (\hat{r}_{j,m}, \hat{r}_{j+1,m})$ and

$$\hat{r}_j^{(m)} = \begin{cases} \hat{q}_m + \sum_{k=1}^{\hat{U}(j,m)-1} \mu(\hat{Q}_k) \hat{p}_{k,m}, & \text{if } j = 1, 2 \ldots, \hat{c}(m), \\ \hat{q}_{m+1}, & \text{if } j = \hat{c}(m) + 1. \end{cases} \tag{34}$$

B. Map characteristics

Characteristics of maps constructed with the original solution are presented in the following propositions. We assume here that the unknown map S is PWL and $\mathcal{R}(\mathcal{Q})$-semi-Markov.

Equation (14) reveals that each nonzero element of the order-zero FPM belonging to an arbitrary map S is a function of the *absolute value(s)* of one or more branch slopes. When attempting to reconstruct S from an estimate of this matrix alone, neither its branch count nor the order in which it maps the intervals of each $\mathcal{R}^{(m)}$ to the intervals of $\mathcal{Q}$ (i.e., the *interval mapping order*) can be discerned. Also, it is not possible to

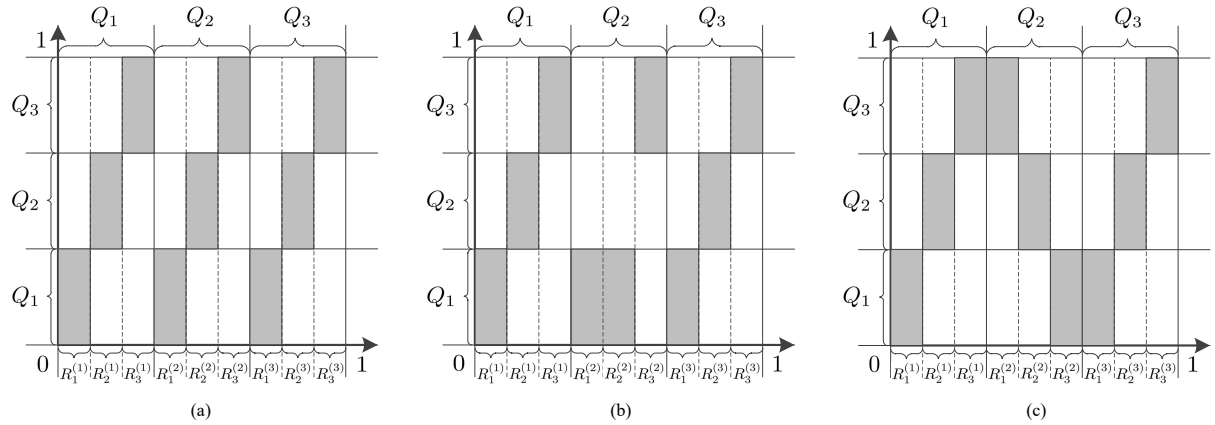

Fig. 2. Branch structure of $\mathcal{R}(\mathcal{Q})$-semi-Markov maps that are (a) canonical and (b, c) not canonical, where the region spanned by each branch domain and codomain is represented by a shaded rectangle.

discern each branch's monotonicity. This holds regardless of estimation accuracy. Thus, additional assumptions regarding S are required to construct a unique map from the order-zero FPM estimate. The original solution assumes S is canonical.

Proposition III.1. *Assume S is canonical and has only increasing branches. If the Markov partition and the order-zero FPM of S are estimated with perfect accuracy (i.e., $\hat{\mathcal{Q}} = \mathcal{Q}$ and $\hat{\mathbf{P}}_S^{(0)} = \mathbf{P}_S^{(0)}$), the original solution reconstructs S perfectly (i.e., $\hat{S} = S$). Conversely, if S is not canonical or possesses any decreasing branches, S cannot be reconstructed successfully (i.e., $\hat{S} \neq S$), regardless of estimation accuracy.*

Proof. If S is canonical, the number of branches of S over each $\mathcal{R}^{(m)}$ is equal to the number of nonzero entries in column m of its order-zero FPM. The order of the branches over each $\mathcal{R}^{(m)}$ is also uniquely determined (i.e., equal to the order of the nonzero entries of the column). Now, suppose that S is indeed canonical, and that the Markov partition and order-zero FPM estimates are perfect. It follows that the original solution correctly discerns the number of branches of S over each $\mathcal{R}^{(m)}$, their domains and codomains. By further assuming that all branches are increasing, it is able to construct the map uniquely from the order-zero FPM estimate. However, by assuming canonicity and increasing branches, the original solution cannot successfully reconstruct S if any of these assumptions are invalid, regardless of estimation accuracy. □

Proposition III.2. *If the original solution estimates the Markov partition and the order-zero FPM of S with perfect accuracy (i.e., $\hat{\mathcal{Q}} = \mathcal{Q}$ and $\hat{\mathbf{P}}_S^{(0)} = \mathbf{P}_S^{(0)}$), the invariant density $f_{\hat{S}}^*$ of $\hat{S}$ equals the invariant density f_S^* of S, regardless of whether S is canonical or the monotonicity of its branches.*

Proof. The invariant density of S coincides with the right Perron eigenvector of the map's order-zero FPM, and is PWC over the Markov partition. Thus, if the assumptions of Proposition III.2 hold, the invariant density functions are also equal. In section V, we present numerical examples that demonstrate the original solution's close approximation of invariant densities of several non-canonical maps and maps with decreasing branches, despite the estimates of the Markov partition and order-zero FPM not being equal to those of the unknown map (i.e., imperfect estimation accuracy). □

Proposition III.3. *If S possesses any decreasing branches, a subset of the PSD mode characteristics of map $\hat{S}$ (i.e., the center frequencies and bandwidths of the $2M$ spectral modes that constitute the map's PSD) are generally unequal to that of S, regardless of estimation accuracy.*

Proof. Equations (18) and (19) imply that, in the PSD of S, each spectral mode's center frequency and bandwidth coincide with the argument and modulus, respectively, of a particular eigenvalue of the map's order-one FPM. The eigenvalues of this matrix are comprised of the eigenvalues of its top-left submatrix $\mathbf{Q}_S$ and the eigenvalues of the order-zero FPM, and each nonzero element of $\mathbf{Q}_S$ is a function of the monotonicity of one or more branches of S (see (10)). Due to the original solution's assumption of S possessing no decreasing branches, submatrix $\mathbf{Q}_{\hat{S}}$ of the order-one FPM of $\hat{S}$ generally does not match its counterpart in the order-one FPM of S if the latter map does possess any decreasing branches. Thus, the eigenvalues of these matrices are generally not equal. □

Proposition III.3 implies that the PSDs of $\hat{S}$ and S are generally not equal if S possesses any decreasing branches.

IV. Proposed solution

We first describe the proposed approach towards solving IFPP-III. Thereafter, we outline the steps of the proposed solution and characterize the maps it constructs.

A. Solution approach

The proposed solution assumes that the unknown map S is PWL and $\mathcal{R}(\mathcal{Q})$-semi-Markov, but no assumptions are made regarding the monotonicity of the unknown map's branches. Similar to the original solution [16], it constructs a map $\hat{S}$ according to a matrix estimate of the unknown map's FPO. However, $\hat{S}$ is constructed such that *both* its order-zero FPM $\mathbf{P}_{\hat{S}}^{(0)}$ and submatrix $\mathbf{Q}_{\hat{S}}$ of its order-one FPM are (if perfect estimation accuracy is assumed) equal to their counterparts of the unknown map. By selecting $\hat{S}$ from a new class of

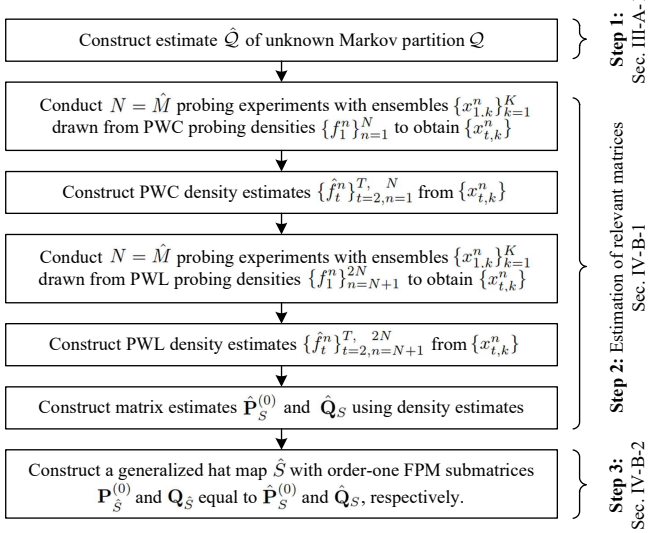

Fig. 3. Flowchart of the proposed solution.

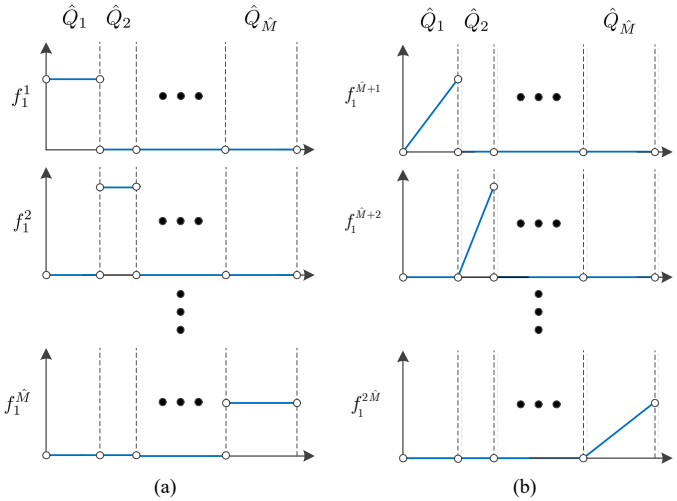

Fig. 4. Probing densities used in the (a) first round and (b) second round of probing experiments.

generalized hat maps (a subset of this class was introduced in [25]), this is achieved whether S is canonical or not.

B. Solution steps

The three steps of the proposed solution are similar to those of the original solution, and are outlined in the flowchart of Fig. 3. Two rounds of probing experiments are performed in the second step. These rounds of probing experiments use probing densities that are PWC and PWL over the estimated Markov partition, respectively. Using the density estimates, the order-zero FPM and submatrix $\mathbf{Q}_S$ of the order-one FPM belonging to the unknown map are estimated, thereby producing $\hat{\mathbf{P}}_S^{(0)}$ and $\hat{\mathbf{Q}}_S$. In the final step, a generalized hat map $\hat{S}$ is constructed with both $\mathbf{P}_{\hat{S}}^{(0)} = \hat{\mathbf{P}}_S^{(0)}$ and $\mathbf{Q}_{\hat{S}} = \hat{\mathbf{Q}}_S$.

1) Estimation of the relevant matrices: A first round of probing experiments $n = 1, 2, \ldots, \hat{M}$ is conducted, where $\hat{M}$ denotes the number of intervals in the estimated Markov partition $\hat{\mathcal{Q}}$. Probing densities $f_1^1, f_1^2, \ldots, f_1^{\hat{M}}$ are selected to be PWC over $\hat{\mathcal{Q}}$ according to the expression

$$f_1^n(x) = \sum_{m=1}^{\hat{M}} f_{1,0,m}^n \chi_{\hat{Q}_m}(x), \tag{35}$$

where

$$f_{1,0,m}^n = [1/\mu(\hat{Q}_m)]\delta_{n,m} \tag{36}$$

and $\delta_{n,m}$ is the Kronecker delta function. These probing densities are plotted in Fig. 4(a).

The density function estimates $\hat{f}_t^1, \hat{f}_t^2, \ldots, \hat{f}_t^{\hat{M}}$, where $t = 1, 2, \ldots, T$, are constructed to be PWC over $\hat{\mathcal{Q}}$ and as specified in (22) and (23). The matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$, which possess as columns the order-zero vector representations of the density function estimates, are then constructed as specified in (24) and (25), respectively.

A second round of probing experiments $n = \hat{M} + 1$, $\hat{M} + 2, \ldots, 2\hat{M}$ is performed next. The probing densities of these experiments, denoted by $f_1^{\hat{M}+1}, f_1^{\hat{M}+2}, \ldots, f_1^{2\hat{M}}$, are selected to be PWL over $\hat{\mathcal{Q}}$ according to the expression

$$f_1^n(x) = \sum_{m=1}^{\hat{M}} (f_{1,1,m}^n x + f_{1,0,m}^n)\chi_{\hat{Q}_m}(x), \tag{37}$$

where

$$f_{1,0,m}^n = -[2\hat{q}_m/\mu(\hat{Q}_m)^2]\delta_{n,m} \tag{38}$$

and

$$f_{1,1,m}^n = [2/\mu(\hat{Q}_m)^2]\delta_{n,m}. \tag{39}$$

These probing densities are plotted in Fig. 4(b).

Densities that are PWL over $\mathcal{Q}$ remain PWL over the same partition under iteration of an $\mathcal{R}(\mathcal{Q})$-semi-Markov map (refer to section II). Thus, after conducting the probing experiments, the density function estimates $\hat{f}_t^{\hat{M}+1}, \hat{f}_t^{\hat{M}+2}, \ldots, \hat{f}_t^{2\hat{M}}$, where $t = 1, 2, \ldots, T$, are constructed to be PWL over the estimated partition $\hat{\mathcal{Q}}$ according to the expression

$$\hat{f}_t^n(x) = \begin{cases} f_1^n(x) & \text{if } t = 1, \\ \sum_{m=1}^{\hat{M}} (\hat{f}_{t,1,m}^n x + \hat{f}_{t,0,m}^n)\chi_{\hat{Q}_m}(x) & \text{if } t = 2, \ldots, T. \end{cases} \tag{40}$$

The coefficients $\hat{f}_{t,0,m}^n$ and $\hat{f}_{t,1,m}^n$ are estimated by applying the method of moments to $f_t^n|_{\hat{Q}_m}$. Let $c_{t,p,m}^n$ denote the p^{th} moment of $f_t^n|_{\hat{Q}_m}$ centered around the midpoint of $\hat{Q}_m$; i.e.,

$$c_{t,p,m}^n \triangleq \int_{\hat{q}_m}^{\hat{q}_{m+1}} \left(x - \frac{\hat{q}_m + \hat{q}_{m+1}}{2}\right)^p f_t^n(x)dx. \tag{41}$$

The corresponding empirical moments are defined as

$$\hat{c}_{t,p,m}^n \triangleq \frac{1}{K}\sum_{k=1}^{K} \left(x_{t,k}^n - \frac{\hat{q}_m + \hat{q}_{m+1}}{2}\right)^p \chi_{\hat{Q}_m}(x_{t,k}^n). \tag{42}$$

Replacing f_t^n in (41) with $\hat{f}_t^n$, equating the resulting expression to $\hat{c}_{t,p,m}^n$ with $p = 0$ and $p = 1$, and simultaneously solving these equations for the unknown coefficients yield

$$\hat{f}_{t,0,m}^n = \frac{1}{\mu(\hat{Q}_m)}\left[\hat{c}_{t,0,m}^n - \frac{6(\hat{q}_{m+1} + \hat{q}_m)}{\mu(\hat{Q}_m)^2}\hat{c}_{t,1,m}^n\right] \tag{43}$$

and

$$\hat{f}^n_{t,1,m} = 12\hat{c}^n_{t,1,m}/\mu(\hat{Q}_m)^3. \tag{44}$$

We define the matrix $\hat{\mathbf{F}}_{t,1} \triangleq [\hat{\mathbf{f}}^{\hat{M}}_{t,1}, \hat{\mathbf{f}}^{\hat{M}+1}_{t,1}, \ldots, \hat{\mathbf{f}}^{2\hat{M}}_{t,1}]$, where $\hat{\mathbf{f}}^n_{t,1} \triangleq [\hat{f}^n_{t,1,1}, \hat{f}^n_{t,1,2}, \ldots, \hat{f}^n_{t,1,\hat{M}}]^T$, and let

$$\hat{\mathbf{C}} \triangleq [\hat{\mathbf{F}}_{1,1}, \hat{\mathbf{F}}_{2,1}, \ldots, \hat{\mathbf{F}}_{T-1,1}] \tag{45}$$

and

$$\hat{\mathbf{D}} \triangleq [\hat{\mathbf{F}}_{2,1}, \hat{\mathbf{F}}_{3,1}, \ldots, \hat{\mathbf{F}}_{T,1}]. \tag{46}$$

Having estimated the densities, we proceed by estimating $\mathbf{P}^{(0)}_S$ and $\mathbf{Q}_S$ from $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}$ and $\hat{\mathbf{D}}$. We observe that every PWL and $\mathcal{R}(\mathcal{Q})$-semi-Markov map S has branch functions $S_{j,m}$ with slopes $v_{j,m}$ that satisfy

$$\sum_{j\in T(k,m)} \frac{s_{j,m}}{v^2_{j,m}} \leq \sum_{j\in T(k,m)} \frac{1}{v^2_{j,m}} \leq \Big[\sum_{j\in T(k,m)} \frac{1}{|v_{j,m}|}\Big]^2. \tag{47}$$

Together with (10) and (14), this inequality implies that the elements of matrices $\mathbf{Q}_S = [q_{k,m}]^M_{k,m=1}$ and $\mathbf{P}^{(0)}_S = [p_{k,m}]^M_{k,m=1}$ must satisfy $|q_{k,m}| \leq p^2_{k,m}$.

Now, if it were assumed that $\hat{\mathcal{Q}} = \mathcal{Q}$ and all density estimates are perfect (i.e., $\hat{f}^n_t = f^n_t$), we have $\hat{\mathbf{B}} = \mathbf{P}^{(0)}_S \hat{\mathbf{A}}$ and $\hat{\mathbf{D}} = \mathbf{Q}_S \hat{\mathbf{C}}$. This leads to the approach of estimating $\mathbf{P}^{(0)}_S$ and $\mathbf{Q}_S$ by solving the *joint* constrained optimization problem

$$\mathbf{P}^*_S, \mathbf{Q}^*_S = \arg\min_{\mathbf{P},\mathbf{Q}} : ||\hat{\mathbf{W}} - \mathbf{D}_{\mathbf{P},\mathbf{Q}}\hat{\mathbf{V}}||_F, \tag{48}$$

subject to (i) $\hat{\mathbf{q}}^T\mathbf{P} = \hat{\mathbf{q}}^T$, (ii) $\mathbf{P} \triangleq [p_{k,m}]^{\hat{M}}_{k,m=1}$ being nonnegative, and (iii) $\mathbf{Q} \triangleq [q_{k,m}]^{\hat{M}}_{k,m=1}$ satisfying $|q_{k,m}| \leq p^2_{k,m}$. The matrices in (48) are defined as $\hat{\mathbf{V}} \triangleq [\hat{\mathbf{C}}; \hat{\mathbf{A}}]$ and $\hat{\mathbf{W}} \triangleq [\hat{\mathbf{D}}; \hat{\mathbf{B}}]$, where ';' denotes vertical matrix concatenation, and

$$\mathbf{D}_{\mathbf{P},\mathbf{Q}} \triangleq \begin{pmatrix} \mathbf{Q} & \mathbf{0}_{\hat{M}\times\hat{M}} \\ \mathbf{0}_{\hat{M}\times\hat{M}} & \mathbf{P} \end{pmatrix}. \tag{49}$$

We prove in Proposition A.1 of Appendix A that the joint problem is not convex. This motivates the formulation and solution of two *separate* optimization problems. The first problem is identical to that defined in section III-A2. Specifically, the problem

$$\mathbf{P}^*_S = \arg\min_{\mathbf{P}} : ||\hat{\mathbf{B}} - \mathbf{P}\hat{\mathbf{A}}||_F, \tag{50}$$

subject to

$$\hat{\mathbf{q}}^T\mathbf{P} = \hat{\mathbf{q}}^T \tag{51}$$

and $\mathbf{P}$ being nonnegative, is solved to obtain an estimate $\hat{\mathbf{P}}^{(0)}_S = [\hat{p}_{k,m}]^{\hat{M}}_{k,m=1}$. The second problem

$$\mathbf{Q}^*_S = \arg\min_{\mathbf{Q}} : ||\hat{\mathbf{D}} - \mathbf{Q}\hat{\mathbf{C}}||_F, \tag{52}$$

subject to $\mathbf{Q} \triangleq [q_{k,m}]^{\hat{M}}_{k,m=1}$ satisfying

$$|q_{k,m}| \leq \hat{p}^2_{k,m}, \tag{53}$$

is then solved to obtain the estimate $\hat{\mathbf{Q}}_S$. We prove in Propositions A.2 and A.3 of Appendix A that the separate problems are convex. Hence, standard methods for solving convex problems may be used to approximate their optimal solutions. However, by considering separate problems, the resulting estimates are generally a suboptimal solution to the joint problem, regardless of the method used.

Consider the estimates

$$\hat{\mathbf{P}}^{(0)}_S = \hat{\mathbf{B}}\hat{\mathbf{A}}^\dagger \tag{54}$$

and

$$\hat{\mathbf{Q}}_S = \hat{\mathbf{D}}\hat{\mathbf{C}}^\dagger, \tag{55}$$

where $\mathbf{Z}^\dagger$ denotes the Moore-Penrose pseudoinverse of $\mathbf{Z}$; i.e., $\mathbf{Z}^\dagger \triangleq \mathbf{Z}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}$. These estimates minimize the Frobenius norms of (50) and (52) [26]. Now, suppose that all density functions are estimated with perfect accuracy. Since $\hat{\mathbf{A}}$ and $\hat{\mathbf{C}}$ have rank equal to $\hat{M}$ (this is a consequence of and the primary motivation for selecting the probing densities of Fig. 4), it follows that $\hat{\mathbf{P}}^{(0)}_S = \mathbf{P}^{(0)}_S$ and $\hat{\mathbf{Q}}_S = \mathbf{Q}_S$. Thus, with perfect density estimates, (54) and (55) are the optimal solutions to both the separate and joint constrained optimization problems.

Motivated by these observations, we solve the constrained optimization problem of (50) by first computing $\overline{\mathbf{P}}_S = \hat{\mathbf{B}}\hat{\mathbf{A}}^\dagger$, and then manipulating this matrix to satisfy the constraints. All negative elements of $\overline{\mathbf{P}}_S$ are first replaced with zeros. Let the columns of the resulting matrix be denoted by $\overline{\mathbf{p}}_1, \overline{\mathbf{p}}_2, \ldots, \overline{\mathbf{p}}_{\hat{M}}$. The approximate solution $\hat{\mathbf{P}}^{(0)}_S \triangleq [\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2, \ldots, \hat{\mathbf{p}}_{\hat{M}}]$ is obtained by normalizing the columns of the resulting matrix as

$$\hat{\mathbf{p}}_m = \frac{\mu(\hat{Q}_m)}{\hat{\mathbf{q}}^T\overline{\mathbf{p}}_m}\overline{\mathbf{p}}_m, \quad m = 1, 2, \ldots, \hat{M}. \tag{56}$$

he constrained optimization problem of (52) is solved by computing $\overline{\mathbf{Q}}_S = \hat{\mathbf{D}}\hat{\mathbf{C}}^\dagger$. The estimate $\hat{\mathbf{Q}}_S \triangleq [\hat{q}_{k,m}]^{\hat{M}}_{k,m=1}$ is then selected as

$$\hat{q}_{k,m} = \begin{cases} \overline{q}_{k,m} & \text{if } -\hat{p}^2_{k,m} \leq \overline{q}_{k,m} \leq \hat{p}^2_{k,m}, \\ -\hat{p}^2_{k,m} & \text{if } \overline{q}_{k,m} < -\hat{p}^2_{k,m}, \\ \hat{p}^2_{k,m} & \text{if } \overline{q}_{k,m} > \hat{p}^2_{k,m}, \end{cases} \tag{57}$$

to satisfy the constraints, where $\overline{\mathbf{Q}}_S = [\overline{q}_{k,m}]^{\hat{M}}_{k,m=1}$.

The proposed solutions to the separate optimization problems are suboptimal. Together with imperfect density estimates, this leads to imperfect matrix estimates. An imperfect estimate of $\mathbf{P}^{(0)}_S$ introduces errors into the reconstructed map (the domains and slopes of the reconstructed branches deviate from those of the unknown map). In addition, an imperfect estimate of $\mathbf{Q}_S$ introduces unnecessary branches into the reconstructed map. Both of these errors are observed in example 1 of section V-A. However, given the preceding observations, we expect the overall accuracy of the reconstructed map to improve with greater density estimation accuracy (i.e., by assigning more states per ensemble in the probing experiments). The numerical results of example 1 support this conjecture.

2) Construction of a generalized hat map: A map from the class of generalized hat maps is constructed in the final step. Conceptually, any map $\hat{S}$ from this class may be derived from a corresponding PWL and canonical semi-Markov map $\tilde{S}$. We define the generalized hat maps using this approach.

Let $\tilde{S}$ denote a PWL and canonical $\tilde{\mathcal{R}}(\tilde{\mathcal{Q}})$-semi-Markov map, and let $\tilde{\mathcal{R}}^{(m)} \triangleq \{\tilde{R}^{(m)}_1, \tilde{R}^{(m)}_2, \ldots, \tilde{R}^{(m)}_{\tilde{c}(m)}\}$ denote the partition of $\tilde{Q}_m$ in the sub-partition $\tilde{\mathcal{R}}$. Suppose that intervals

$$\tilde{A}^{(m)}_j \triangleq \big(\inf(\tilde{R}^{(m)}_j), \inf(\tilde{R}^{(m)}_j) + \tilde{\beta}_{j,m}\mu(\tilde{R}^{(m)}_j)\big), \tag{58}$$

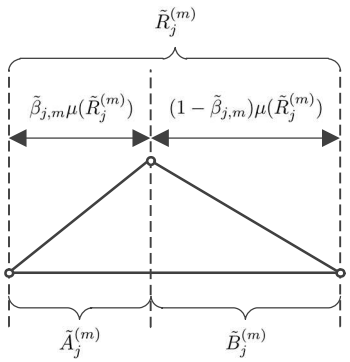

Fig. 5. Bifurcation of the interval $\tilde{R}_j^{(m)}$ and definition of new branch functions during the construction of a generalized hat map.

$$\tilde{B}_j^{(m)} \triangleq \left(\inf(\tilde{R}_j^{(m)}) + \tilde{\beta}_{j,m}\mu(\tilde{R}_j^{(m)}), \sup(\tilde{R}_j^{(m)})\right) \tag{59}$$

are defined from some predetermined parameter $\tilde{\beta}_{j,m} \in [0,1]$. These intervals form a partition of $\tilde{R}_j^{(m)}$, as shown in Fig. 5; in the case where $\tilde{\beta}_{j,m} = 0$ or $\tilde{\beta}_{j,m} = 1$, interval $\tilde{A}_j^{(m)}$ or $\tilde{B}_j^{(m)}$ is defined to be the empty interval, respectively.

We derive the generalized hat map $\hat{S}$ from its counterpart $\tilde{S}$ and a collection of specified parameters $\tilde{\beta}_{j,m}$, where $m = 1, 2, \ldots, \tilde{M}$ and $j = 1, 2, \ldots, \tilde{c}(m)$. Let $\hat{S}$ be $\hat{\mathcal{R}}(\hat{\mathcal{Q}})$-semi-Markov and PWL over $\hat{\mathcal{R}}$, and inherit its Markov partition $\hat{\mathcal{Q}}$ from $\tilde{S}$. We define the partition $\hat{\mathcal{R}}^{(m)}$ of $\hat{Q}_m$ as

$$\hat{\mathcal{R}}^{(m)} \triangleq \{\tilde{A}_j^{(m)}\}_{j=1}^{\tilde{c}(m)} \cup \{\tilde{B}_j^{(m)}\}_{j=1}^{\tilde{c}(m)}, \tag{60}$$

and specify that $\hat{S}$ maps all nonempty intervals $\tilde{A}_j^{(m)}$ and $\tilde{B}_j^{(m)}$ to the same interval that $\tilde{R}_j^{(m)}$ is mapped to by $\tilde{S}$; i.e.,

$$\tilde{A}_j^{(m)} \overset{\hat{S}}{\mapsto} \tilde{S}(\tilde{R}_j^{(m)}), \quad \text{if } \tilde{\beta}_{j,m} > 0, \tag{61}$$

$$\tilde{B}_j^{(m)} \overset{\hat{S}}{\mapsto} \tilde{S}(\tilde{R}_j^{(m)}), \quad \text{if } \tilde{\beta}_{j,m} < 1. \tag{62}$$

We specify that branches of $\hat{S}$ defined over $\tilde{A}_j^{(m)}$ and $\tilde{B}_j^{(m)}$ are increasing and decreasing, respectively (refer to Fig. 5). This implies that the single branch of $\hat{S}$ defined over $\tilde{R}_j^{(m)}$ is decreasing or increasing if $\tilde{\beta}_{j,m} = 0$ or $\tilde{\beta}_{j,m} = 1$, respectively.

Any map $\hat{S}$ derived from a PWL and canonical $\tilde{\mathcal{R}}(\tilde{\mathcal{Q}})$-semi-Markov map $\tilde{S}$ and a collection of parameters $\tilde{\beta}_{j,m}$ in the manner outlined above is referred to as a generalized hat map. The name of this map class originates from the shape of consecutive pairs of map branches, which resembles a hat (in the case where $\beta_{j,m} \neq 1/2$, the hat is not symmetric; thus, we use the term *generalized*). Also, by selecting appropriate values for parameters $\tilde{\beta}_{j,m} \in \{0,1\}$ according to the monotonicity of the branches of any canonical map $\tilde{S}$, it is always possible to construct a generalized hat map that satisfies $\hat{S} = \tilde{S}$. Thus, the canonical maps are a subset of the generalized hat maps.

The following propositions describe two properties of generalized hat maps that are relevant to the final step of the proposed solution. These propositions are proved in Appendix B and Appendix C, respectively.

Proposition IV.1. *The generalized hat map $\hat{S}$ inherits the order-zero FPM $\mathbf{P}_{\tilde{S}}^{(0)} \triangleq [p_{k,m}^{(\tilde{S})}]_{k,m=1}^{\tilde{M}}$ of its counterpart $\tilde{S}$.*

Proposition IV.2. *Let $\mathbf{Q}_{\hat{S}} \triangleq [q_{k,m}^{(\hat{S})}]_{k,m=1}^{\tilde{M}}$ denote the top-left submatrix of the order-one FPM belonging to map $\hat{S}$. The zero-valued elements of this matrix and $\mathbf{P}_{\tilde{S}}^{(0)}$ coincide, whereas their nonzero elements satisfy*

$$q_{k,m}^{(\hat{S})} = (2\tilde{\beta}_{\tilde{T}(k,m),m} - 1)(p_{k,m}^{(\tilde{S})})^2, \tag{63}$$

where $\tilde{T}(k,m)$ denotes the index j satisfying $\tilde{S}(\tilde{R}_j^{(m)}) = \tilde{Q}_k$.

These properties are used to construct $\hat{S}$ such that its order-zero FPM and the top-left submatrix $\mathbf{Q}_{\hat{S}}$ of its order-one FPM are equal to the estimates $\hat{\mathbf{P}}_S^{(0)} \triangleq [\hat{p}_{k,m}]_{k,m=1}^{\hat{M}}$ and $\hat{\mathbf{Q}}_S \triangleq [\hat{q}_{k,m}]_{k,m=1}^{\hat{M}}$ obtained in the previous step. We first construct a PWL and canonical semi-Markov map $\tilde{S}$ using the original solution, as outlined in section III-A3. This ensures that the order-zero FPM of $\tilde{S}$ matches the estimate $\hat{\mathbf{P}}_S^{(0)}$. A generalized hat map $\hat{S}$ is then constructed from $\tilde{S}$, thereby ensuring that the order-zero FPM of $\hat{S}$ also equals $\hat{\mathbf{P}}_S^{(0)}$ (Proposition IV.1). The parameters $\tilde{\beta}_{j,m}$ used during the construction of $\hat{S}$ are selected to ensure that $\mathbf{Q}_{\hat{S}}$ matches the estimate $\hat{\mathbf{Q}}_S$. Proposition IV.2 and (63) imply that this may be achieved by selecting

$$\tilde{\beta}_{j,m} = \frac{1}{2}\left(\frac{\hat{q}_{\tilde{U}(j,m),m}}{\hat{p}^2_{\tilde{U}(j,m),m}} + 1\right) \tag{64}$$

for each $m = 1, 2, \ldots, \tilde{M}$ and $j = 1, 2, \ldots, \tilde{c}(m)$, where $\tilde{U}(j,m)$ is the index k satisfying $\tilde{S}(\tilde{R}_j^{(m)}) = \tilde{Q}_k$.

C. Map characteristics

In the following propositions, we describe several characteristics of maps $\hat{S}$ constructed using the proposed solution.

Proposition IV.3. *The order-zero FPM of $\hat{S}$ can be selected to equal any order-zero FPM estimate $\hat{\mathbf{P}}_S^{(0)}$ that satisfies* (51) *and has nonnegative elements. Simultaneously, the top-left submatrix $\mathbf{Q}_{\hat{S}}$ of the order-one FPM belonging to $\hat{S}$ can be selected to equal any estimate $\hat{\mathbf{Q}}_S$ that satisfies* (53)*.*

Proof. The statement regarding the order-zero FPM follows from $\hat{S}$ inheriting the order-zero FPM of $\tilde{S}$ (Proposition IV.1), and from the order-zero FPM of map $\tilde{S}$ (constructed using the original solution) being equal to any estimate satisfying the constraints [16]. The statement regarding submatrix $\mathbf{Q}_{\hat{S}}$ follows from Proposition IV.2. Specifically, (63) implies that any estimate satisfying the constraint of (53) can be realized by selecting an appropriate value for each parameter $\tilde{\beta}_{\tilde{T}(k,m),m}$ from $[0,1]$ using (64), and then constructing the generalized hat map $\hat{S}$, using these parameters, from $\tilde{S}$. □

Proposition IV.4. *Suppose that $\hat{\mathcal{Q}} = \mathcal{Q}$, and that the matrix estimates satisfy $\hat{\mathbf{P}}_S^{(0)} = \mathbf{P}_S^{(0)}$ and $\hat{\mathbf{Q}}_S = \mathbf{Q}_S$. If S is canonical, it can be successfully reconstructed using the proposed solution (i.e., $\hat{S} = S$) regardless of branch monotonicity.*

Proof. Suppose that the conditions stated in the proposition hold. Using the original solution, an $\tilde{\mathcal{R}}(\tilde{\mathcal{Q}})$-semi-Markov map $\tilde{S}$ is constructed with $\mathbf{P}_{\tilde{S}}^{(0)} = \mathbf{P}_S^{(0)}$. The equality of the order-zero FPMs and the canonicity of S imply that $\tilde{\mathcal{R}} = \mathcal{R}$ and that

each $R_j^{(m)}$ has the same codomain under $\tilde{S}$ and S. Since S is canonical, (10) and (14) imply that the elements of $\mathbf{P}_S^{(0)}$ and $\mathbf{Q}_S$ satisfy $q_{k,m} = s_{T(k,m),m} p_{k,m}^2$. A generalized hat map $\hat{S}$ is then constructed from $\tilde{S}$ using the proposed solution, where the parameters $\tilde{\beta}_{j,m} = \chi_{[0,\infty)}(s_{j,m})$ are selected using (64). The map $\hat{S}$ is therefore $\mathcal{R}(\mathcal{Q})$-semi-Markov, and it is designed such that its branch over $R_j^{(m)}$ is decreasing if $\tilde{\beta}_{j,m} = 0$ and increasing if $\tilde{\beta}_{j,m} = 1$. Thus, corresponding branches from S and $\hat{S}$ have the same monotonicity. As these branches also have the same domains and codomains, we have $\hat{S} = S$. □

Proposition IV.5. *Suppose that $\hat{\mathcal{Q}} = \mathcal{Q}$, and that the matrix estimates satisfy $\hat{\mathbf{P}}_S^{(0)} = \mathbf{P}_S^{(0)}$ and $\hat{\mathbf{Q}}_S = \mathbf{Q}_S$. Let $\hat{S}$ be constructed using the proposed solution. The invariant density and PSD mode characteristics of $\hat{S}$ are then equal to those of the unknown PWL and semi-Markov map S, regardless of the canonicity of S or the monotonicity of its branches.*

Proof. Suppose that the conditions stated in the proposition hold. Proposition IV.3 then implies that the order-zero FPM and the top-left submatrix $\mathbf{Q}_{\hat{S}}$ of the order-one FPM belonging to $\hat{S}$ are equal to their counterparts of map S. Since the order-zero FPMs share the same (normalized) right Perron eigenvector, the order-zero vector representations of their invariant densities are identical. The equality of the maps' Markov partitions then imply that their invariant densities are equal. Furthermore, recall that the PSD mode characteristics of $\hat{S}$ coincide with the eigenvalues of its order-zero FPM and the eigenvalues of submatrix $\mathbf{Q}_{\hat{S}}$ of its order-one FPM. As both these matrices equal their counterparts of S, their eigenvalues and thus the maps' PSD mode characteristics match. □

V. Numerical examples

We present two examples in which the original and proposed solutions were applied with the objective of reconstructing an unknown PWL and semi-Markov map S. These examples illustrate certain characteristics of the maps constructed using both solutions and serve to quantify the reconstruction accuracy. The first example involves a canonical map S with several decreasing branches, and the second example a map S that is neither canonical nor a generalized hat map.

All simulations were carried on a laptop with an Intel i7-9850H processor using the MATLAB numeric computing environment. In each example, the original and proposed solutions were carried out as described in sections III-A and IV-B, with the exception of the step for estimating the Markov partition. This step is identical in both solutions and its accuracy was evaluated in [16]. Hence, we set the estimated Markov partition equal to that of the unknown map in each example. The original solution involved $\hat{M}$ probing experiments and the probing densities of Fig. 4(a), whereas the proposed solution involved $2\hat{M}$ probing experiments and the probing densities of Fig. 4(a) and 4(b). Each probing experiment produced a sequence of $T = 2$ densities.

Each solution was applied in multiple rounds for each example. A different value was selected for the number of states per ensemble K in each round, with K ranging between $K = 1.25 \times 10^3$ and $K = 80 \times 10^3$. To estimate the mean reconstruction error, each solution was repeated 10^3 times per round, with the initial ensemble of each probing experiment drawn randomly from the probing density at the start of each repetition. We consider the L^1 error $\varepsilon_S \triangleq ||S - \hat{S}||_1$ between the constructed and the unknown map, as well as the L^1 errors between their invariant densities and PSDs (denoted by $\varepsilon_{f_S^*} \triangleq ||f_S^* - f_{\hat{S}}^*||_1$ and $\varepsilon_{\Phi_S} \triangleq ||\Phi_S - \Phi_{\hat{S}}||_1$). The mean of these errors was computed over the repetitions in each round.

TABLE I
Branch functions of the map considered in example 1.

Branch	Domain	Codomain	$v_{j,m}$	$w_{j,m}$
$S_{1,1}$	$R_1^{(1)} = (0, 0.04)$	$Q_1 = (0, 0.40)$	10.00	0
$S_{2,1}$	$R_2^{(1)} = (0.04, 0.12)$	$Q_2 = (0.40, 0.50)$	1.25	0.35
$S_{3,1}$	$R_3^{(1)} = (0.12, 0.24)$	$Q_3 = (0.50, 0.80)$	-2.50	1.10
$S_{4,1}$	$R_4^{(1)} = (0.24, 0.40)$	$Q_4 = (0.80, 1.00)$	1.25	0.50
$S_{1,2}$	$R_1^{(2)} = (0.40, 0.42)$	$Q_1 = (0, 0.40)$	20.00	-8.00
$S_{2,2}$	$R_2^{(2)} = (0.42, 0.45)$	$Q_2 = (0.40, 0.50)$	-3.33	1.90
$S_{3,2}$	$R_3^{(2)} = (0.45, 0.47)$	$Q_3 = (0.50, 0.80)$	15.00	-6.25
$S_{4,2}$	$R_4^{(2)} = (0.47, 0.50)$	$Q_4 = (0.80, 1.00)$	6.67	-2.33
$S_{1,3}$	$R_1^{(3)} = (0.50, 0.68)$	$Q_1 = (0, 0.40)$	-2.22	1.51
$S_{2,3}$	$R_2^{(3)} = (0.68, 0.74)$	$Q_2 = (0.40, 0.50)$	1.67	-0.73
$S_{3,3}$	$R_3^{(3)} = (0.74, 0.77)$	$Q_3 = (0.50, 0.80)$	10.00	-6.90
$S_{4,3}$	$R_4^{(3)} = (0.77, 0.80)$	$Q_4 = (0.80, 1.00)$	6.67	-4.33
$S_{1,4}$	$R_1^{(4)} = (0.80, 0.84)$	$Q_1 = (0, 0.40)$	10.00	-8.00
$S_{2,4}$	$R_2^{(4)} = (0.84, 0.88)$	$Q_2 = (0.40, 0.50)$	2.50	-1.70
$S_{3,4}$	$R_3^{(4)} = (0.88, 0.96)$	$Q_3 = (0.50, 0.80)$	3.75	-2.80
$S_{4,4}$	$R_4^{(4)} = (0.96, 1.00)$	$Q_4 = (0.80, 1.00)$	-5.00	5.80

A. Example 1

The canonical $\mathcal{R}(\mathcal{Q})$-semi-Markov map plotted in Fig. 6(a) was selected as the unknown map S (a closely related map was considered in [16]). The map and its branch functions are defined by (6) and (8), respectively. Table I lists the domain, codomain and parameters of each branch function. The Markov partition of S consists of intervals $Q_1 = (0, 0.40)$, $Q_2 = (0.40, 0.50)$, $Q_3 = (0.50, 0.80)$ and $Q_4 = (0.80, 1.00)$.

Figures 6(b) and 6(c) are plots of the maps constructed using the proposed and original solutions with $K = 10^4$ states per ensemble. Since S is canonical, Proposition IV.4 does not exclude it from being successfully reconstructed using the proposed solution. A comparison of Figs. 6(a) and 6(b) reveals that the constructed map closely approximates the unknown map, despite the matrix estimates not being equal to the submatrices of the unknown map's order-one FPM. In particular, the four decreasing branches of the unknown map were successfully reconstructed as decreasing branches by the proposed solution (these branches are colored magenta in Fig. 6), whereas the remaining branches were correctly reconstructed with increasing monotonicity. This implies that the decreasing branches' monotonicity was successfully identified and distinguished from that of the increasing branches by the proposed solution (recall that this is made possible by estimating submatrix $\mathbf{Q}_S$ of the unknown map's order-one FPM, in addition to the order-zero FPM). Figure 6(c) reveals that the original solution incorrectly reconstructs the decreasing branches as increasing branches, as it cannot discern branch

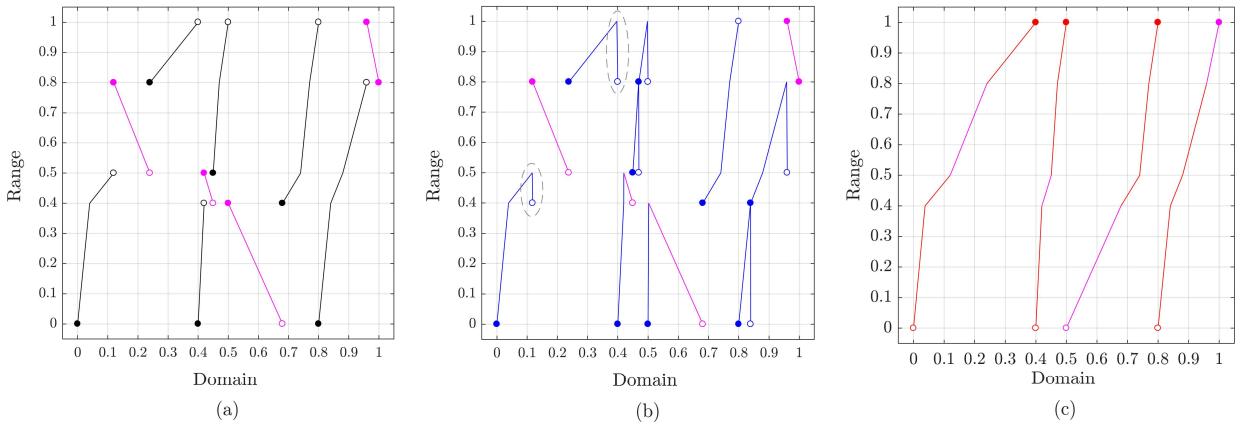

Fig. 6. The maps associated with example 1. The unknown map is plotted in (a), whereas the maps constructed using the proposed and original solutions are plotted in (b) and (c), respectively.

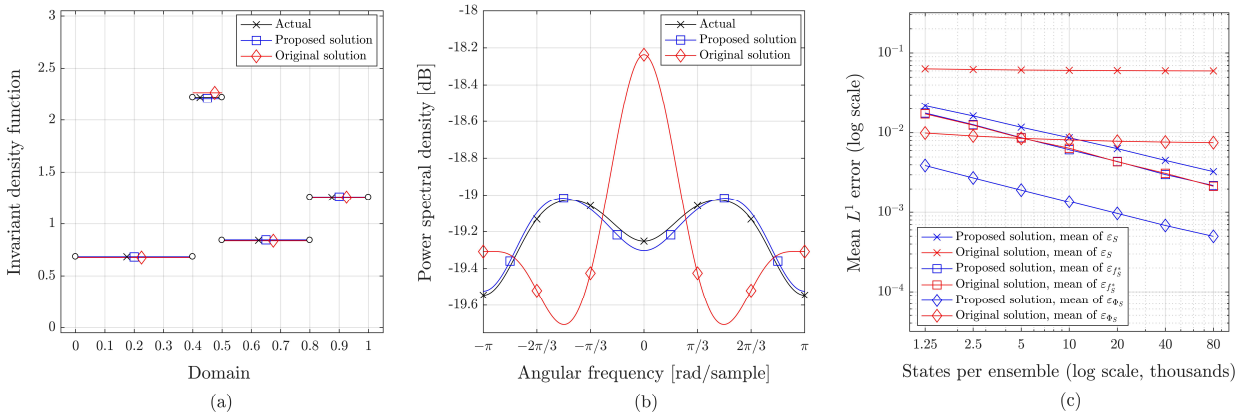

Fig. 7. Plots of the (a) invariant densities and (b) PSDs of the unknown map and the maps constructed using the proposed and original solutions in example 1. The mean L^1 errors of the maps, invariant densities and PSDs are plotted as a function of the number of states per ensemble in (c).

monotonicity from the order-zero FPM estimate only. This confirms the prediction of Proposition III.1.

Imperfectly estimating the order-zero FPM leads to an imperfect match between the branch domains and slopes of both solutions' reconstructed maps and their counterparts of the unknown map; whereas this imperfect match is not clearly visible in Fig. 6(b), it contributes towards the nonzero L^1 error ε_S, plotted as a function of the number of states per ensemble K in Fig. 7(c). The map constructed using the proposed solution also has unnecessary branches (a subset of these branches is indicated using dashed ellipses in Fig. 6(b)), and certain branch pairs possess the characteristic hat-like shape; this is due to the estimate $\hat{\mathbf{Q}}_S$ not being equal to $\mathbf{Q}_S$, which causes unnecessary bifurcation of some intervals of the sub-partition in the final solution step. However, the support of these unnecessary branches has negligible length. The original solution does not produce unnecessary branches, as it does not use an estimate of $\mathbf{Q}_S$. However, by incorrectly estimating the monotonicity of some branches, its error ϵ_S is larger than that of the original solution (see Fig. 7(c)).

The invariant density function of S is plotted in Fig. 7(a). This density function is PWC over $\mathcal{Q}$ and is given by $f^*_S(x) = 0.68\chi_{Q_1} + 2.22\chi_{Q_2} + 0.84\chi_{Q_3} + 1.26\chi_{Q_4}$, where $\chi_{Q_m} \triangleq \chi_{Q_m}(x)$. The invariant densities of the maps constructed using both solutions are also plotted in Fig. 7(a). The figure reveals that these densities closely approximate the invariant density of S, despite imperfect estimation of the order-zero FPM (as required by Propositions III.2 and IV.5). The L^1 error $\varepsilon_{f^*_S}$, plotted as a function of the number of states per ensemble K in Fig. 7(c), suggests the accuracy of the two solutions are identical in this regard. The invariant densities coincide with the right Perron eigenvector of the order-zero FPM estimate; as both solutions estimate this matrix in the same manner, the densities are expected to be equally accurate.

TABLE II
POWER SPECTRAL DENSITY PARAMETERS FOR EXAMPLE 1.

Mode m	Coefficient c_m	Eigenvalue λ_m
1	2.97×10^{-1}	1.00
2	$(1.41 - 1.38j) \times 10^{-1}$	$(-2.00 + 1.73j) \times 10^{-1}$
3	$(1.41 + 1.38j) \times 10^{-1}$	$(-2.00 - 1.73j) \times 10^{-1}$
4	8.22×10^{-3}	1.00×10^{-1}
5	9.18×10^{-3}	-2.75×10^{-1}
6	2.61×10^{-4}	-1.23×10^{-1}
7	$(1.86 + 1.35j) \times 10^{-2}$	$(1.44 + 1.18j) \times 10^{-1}$
8	$(1.86 - 1.35j) \times 10^{-2}$	$(1.44 - 1.18j) \times 10^{-1}$

The PSD of S is given by (19), and its parameters are listed in Table II. The continuous component of the map's PSD

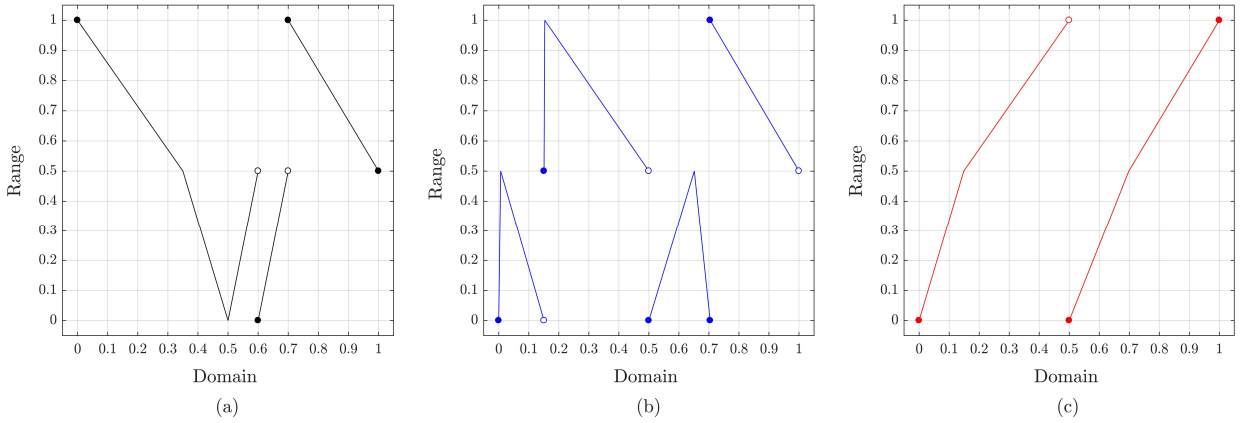

Fig. 8. The maps associated with example 2. The unknown map is plotted in (a), whereas the maps constructed using the proposed and original solutions are plotted in (b) and (c), respectively.

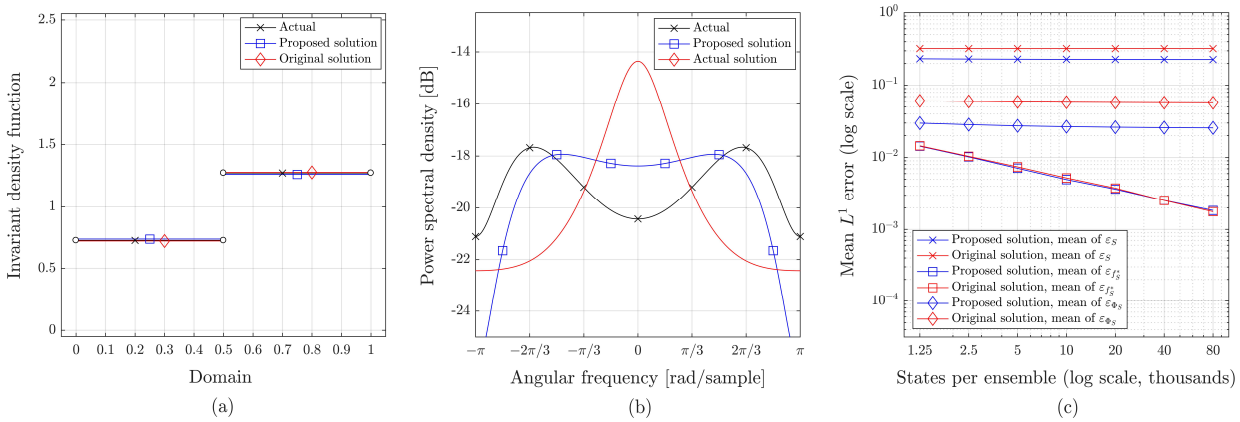

Fig. 9. Plots of the (a) invariant densities and (b) PSDs of the unknown map and the maps constructed using the proposed and original solutions in example 2. The mean L^1 errors of the maps, invariant densities and PSDs are plotted as a function of the number of states per ensemble in (c).

is plotted in Fig. 7(b). The figure also contains plots of the continuous PSD components belonging to maps constructed using the proposed and original solutions. Whereas Proposition IV.4 implies that the PSD associated with the proposed solution equals that of S if perfect matrix estimates are used, the figure reveals that the PSDs remain approximately equal despite these estimates being imperfect. In general, canonicity of S is required for these PSDs to be approximately equal. If S is not canonical, submatrix $\mathbf{R}_S$ of the maps' order-one FPMs will generally not be equal, and the PSDs will deviate. In contrast, the PSD associated with the original solution deviates from that of S (Proposition III.3 is not satisfied).

The mean L^1 errors associated with the constructed maps, their invariant densities and PSDs are plotted as a function of the number of states per ensemble K in Fig. 7(c). The figure reveals that all observed errors associated with the proposed solution are decreasing, whereas the error of the invariant density $\varepsilon_{f^*_S}$ alone decreases in the case of the original solution. Hence, despite the imperfections of the map constructed using the proposed solution, its measured L^1 error ε_S decreases as the accuracy of the density estimates improve. Thus, the proposed solution successfully reconstructs the map by not having to assume all its branches are increasing. In contrast, the original solution makes this assumption and fails to reconstruct the map regardless of the accuracy of the density estimates. However, as no assumption has to be made regarding canonicity or branch monotonicity to estimate the order-zero FPM (and hence the invariant density), *both* solutions' L^1 error $\varepsilon_{f^*_S}$ also decreases as a function of K.

B. Example 2

The $\mathcal{R}(\mathcal{Q})$-semi-Markov map plotted in Fig. 8(a) was selected as the unknown map S. This map is neither canonical nor a member of the generalized hat maps. The map and its branch functions are defined by (6) and (8), respectively. Table III lists the domain, codomain and parameters of each branch function. The map's Markov partition $\mathcal{Q}$ has two intervals $Q_1 = (0, 0.50)$ and $Q_2 = (0.50, 1.00)$.

Figures 8(b) and 8(c) are plots of the maps constructed using the proposed and original solutions with $K = 10^4$ states per ensemble. The figures reveal that neither map approximates S. This is consistent with Proposition III.1, which disqualifies the original solution from successfully reconstructing non-canonical maps. As S is not a generalized hat map, the proposed solution also cannot successfully reconstruct it (the conditions of Proposition IV.4 are not met).

The invariant density function of S is plotted in Fig. 9(a). This density function is PWC over $\mathcal{Q}$ and is given by $f^*_S(x) =$

TABLE III
BRANCH FUNCTIONS OF THE MAP CONSIDERED IN EXAMPLE 2.

Branch	Domain	Codomain	$v_{j,m}$	$w_{j,m}$
$S_{1,1}$	$R_1^{(1)} = (0, 0.35)$	$Q_2 = (0.50, 1.00)$	-1.43	1.00
$S_{1,2}$	$R_2^{(1)} = (0.35, 0.5)$	$Q_1 = (0, 0.50)$	-3.33	1.67
$S_{1,2}$	$R_1^{(2)} = (0.50, 0.60)$	$Q_1 = (0, 0.50)$	5.00	-2.5
$S_{2,2}$	$R_2^{(2)} = (0.60, 0.70)$	$Q_1 = (0, 0.50)$	5.00	-3.00
$S_{3,2}$	$R_3^{(2)} = (0.70, 1.00)$	$Q_2 = (0.50, 1.00)$	-1.67	2.17

TABLE IV
POWER SPECTRAL DENSITY PARAMETERS FOR EXAMPLE 2.

Mode m	Coefficient c_m	Eigenvalue λ_m
1	3.23×10^{-1}	1.00
2	-1.98×10^{-1}	-1.00×10^{-1}
3	$(1.38 - 1.31j) \times 10^{-1}$	$(-2.25 + 1.45j) \times 10^{-1}$
4	$(1.38 + 1.31j) \times 10^{-1}$	$(-2.25 - 1.45j) \times 10^{-1}$

$0.73\chi_{Q_1}(x) + 1.27\chi_{Q_2}(x)$. The invariant density functions of the maps constructed using both solutions are also plotted in Fig. 9(a). These densities closely approximate the invariant density of S, despite the use of imperfect matrix estimates.

The PSD of S is given by (19), and its parameters are listed in Table IV. The continuous component of the map's PSD is plotted in Fig. 9(b). The figure also contains plots of the continuous component of the PSDs belonging to maps constructed using both solutions. Neither solution constructs a map with PSD approximating that of S. In general, canonicity of S is necessary for any of the PSDs to be approximately equal. However, canonicity is not required for the PSD mode characteristics to be consistent, as demonstrated next.

The mean L^1 errors associated with the constructed maps, their invariant densities and PSDs are plotted as a function of the number of states per ensemble K in Fig. 9(c). The figure reveals that the error of the invariant densities $\varepsilon_{f_{\hat{S}}^*}$ alone decreases for both solutions. In particular, Fig. 9(c) confirms a nondecreasing L^1 error ε_S for both solutions. This implies that the inability to reconstruct the map does not stem from estimation inaccuracy, but from the canonicity assumption.

Proposition IV.5 implies that the PSD mode characteristics of maps constructed using the proposed solution coincide with those of S if the matrix estimates equal the submatrices of the unknown map's order-one FPM. However, this does not hold for the original solution. The PSD mode characteristics of maps constructed using both solutions are investigated with scatter plots of the order-one FPM eigenvalues.

Figures 10(a) and 10(b) are scatter plots of the eigenvalues belonging to the order-one FPMs of 100 distinct maps $\hat{S}$ constructed using the proposed solution, for choices of $K = 2.5 \times 10^3$ and $K = 8 \times 10^4$ states per ensemble, respectively. The figures reveal that the eigenvalues associated with the constructed maps are clustered around the eigenvalues associated with S (indicated using crosses), and that an increase in the number of states per ensemble causes the boundaries of these clusters to constrict. This implies that the PSD mode characteristics of maps constructed using the proposed solution approximate those of the unknown map more accurately as density estimation accuracy improves. Figures 11(a) and 11(b) are scatter plots of the eigenvalues belonging to the order-one FPMs of 100 distinct maps constructed using the original solution. These figures reveal that only a subset of the eigenvalues associated with the constructed maps cluster around the eigenvalues of S, regardless of estimation accuracy.

C. Mean reconstruction time

The mean time t_m required to reconstruct a map using each of the solutions is plotted in Fig. 12 as a function of the number of states per ensemble K for both examples. The figure reveals that the proposed solution requires more time to reconstruct the map than the original solution in both examples, with a maximum of $t_m = 33.5$ seconds required to reconstruct the most sophisticated map (i.e., example 1 with 80×10^3 states per ensemble). Further investigation revealed that the probing experiments, which include the generation of the initial state ensemble, require 95.3% to 99.8% of the overall reconstruction time. The longer execution time of the proposed solution can thus be ascribed to it carrying out twice the number of probing experiments as compared to the original solution.

The power law curve was fitted to the measured reconstruction times of each solution and example, such that

$$t_m(K) \approx aK^b. \tag{65}$$

The fitted curves are plotted as solid lines in Fig. 12. The figure suggests a close fit between the measurements and the curves. The parameters a and b of each curve are listed in Table V. The table reveals that the exponent b is approximately equal to unity in each case, which implies a (near) linear relationship between the measured reconstruction time and the number of states per ensemble.

VI. CONCLUSIONS

This paper proposed a novel solution to the inverse Frobenius-Perron problem of reconstructing an unknown PWL semi-Markov map from causal sequences of state density functions generated by the map. The proposed solution enlarges the class of maps reconstructible by the original solution of [16] from the canonical maps with increasing branch functions only to canonical maps with branch functions of arbitrary monotonicity. Also, in contrast to the original solution, the PSD mode characteristics of maps constructed using the proposed solution are consistent with those of any unknown PWL semi-Markov map; this holds regardless of the canonicity of the unknown map or the monotonicity of its branch functions. Whereas the propositions associated with these statements require that the Markov partition and relevant matrix estimates are equal to their counterparts of the unknown map, we demonstrated close approximation of the maps and their statistical characteristics using numerical examples that involved imperfect estimation accuracy. Specifically, we demonstrated that maps constructed using the proposed solution closely approximate a canonical map with a nonempty subset of decreasing branches. Two examples were presented

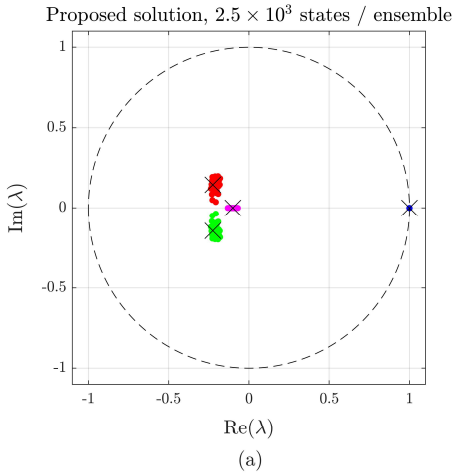

Fig. 10. Scatter plots of eigenvalues belonging to the order-one FPMs of maps constructed using the proposed solution in example 2, for (a) $K = 2.5 \times 10^3$ and (b) $K = 8 \times 10^4$ states per ensemble. The eigenvalues of the unknown map are indicated using crosses.

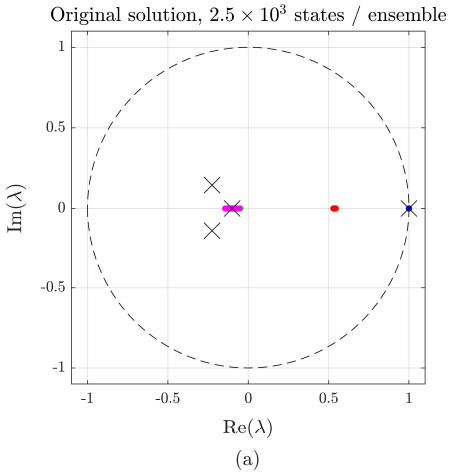

Fig. 11. Scatter plots of eigenvalues belonging to the order-one FPMs of maps constructed using the original solution in example 2, for (a) $K = 2.5 \times 10^3$ and (b) $K = 8 \times 10^4$ states per ensemble. The eigenvalues of the unknown map are indicated using crosses.

that demonstrate the close approximation of the invariant density and PSD mode characteristics of PWL semi-Markov maps. Thus, the proposed solution is a practical means for modelling a larger class of semi-Markov maps.

APPENDIX A
CONVEXITY OF THE CONSTRAINED OPTIMIZATION PROBLEMS

We first prove that the joint constrained optimization problem introduced in section IV-B1 is not convex. Thereafter, we prove that the separate constrained optimization problems introduced in the same section are both convex.

Proposition A.1. *Consider the joint optimization problem*

$$\mathbf{P}_S^*, \mathbf{Q}_S^* = \arg\min_{\mathbf{P},\mathbf{Q}} : ||\hat{\mathbf{W}} - \mathbf{D}_{\mathbf{P},\mathbf{Q}}\hat{\mathbf{V}}||_F, \tag{66}$$

subject to

$$\hat{\mathbf{q}}^T\mathbf{P} = \hat{\mathbf{q}}^T, \tag{67}$$

$\mathbf{P} \triangleq [p_{k,m}]_{k,m=1}^{\hat{M}}$ *being nonnegative, and* $\mathbf{Q} \triangleq [q_{k,m}]_{k,m=1}^{\hat{M}}$ *satisfying*

$$|q_{k,m}| \leq p_{k,m}^2. \tag{68}$$

The matrices in (66) *are defined in section IV-B1. This problem is not convex.*

Proof. We prove that the feasible set associated with (68) is not convex. Assume that $(\mathbf{P}_1, \mathbf{Q}_1)$ and $(\mathbf{P}_2, \mathbf{Q}_2)$ are both elements of the feasible set. If the problem were convex, then

$$(\alpha\mathbf{P}_1 + (1-\alpha)\mathbf{P}_2, \alpha\mathbf{Q}_1 + (1-\alpha)\mathbf{Q}_2) \tag{69}$$

must be an element of the feasible set for all $\alpha \in [0, 1]$ and hence satisfy (68); i.e.,

$$|\alpha q_{k,m}^{(1)} + (1-\alpha)q_{k,m}^{(2)}| \leq [\alpha p_{k,m}^{(1)} + (1-\alpha)p_{k,m}^{(2)}]^2 \tag{70}$$

for all $k, m = 1, 2, \ldots, \hat{M}$, where $\mathbf{P}_j \triangleq [p_{k,m}^{(j)}]_{k,m=1}^{\hat{M}}$ and $\mathbf{Q}_j \triangleq [q_{k,m}^{(j)}]_{k,m=1}^{\hat{M}}$.

As $(\mathbf{P}_1, \mathbf{Q}_1)$ and $(\mathbf{P}_2, \mathbf{Q}_2)$ are elements of the feasible set, they satisfy (68) and hence

$$-(p_{k,m}^{(1)})^2 \leq q_{k,m}^{(1)} \leq (p_{k,m}^{(1)})^2 \tag{71}$$

and

$$-(p_{k,m}^{(2)})^2 \leq q_{k,m}^{(2)} \leq (p_{k,m}^{(2)})^2. \tag{72}$$

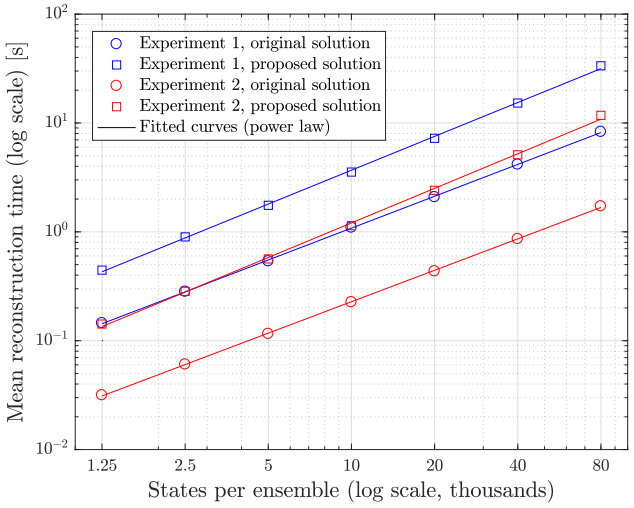

Fig. 12. Mean time required to reconstruct the maps of examples 1 and 2 using each solution.

Multiplying (71) with α and (72) with $1-\alpha$ and adding these expressions yield

$$|\alpha q_{k,m}^{(1)} + (1-\alpha) q_{k,m}^{(2)}| \leq \alpha (p_{k,m}^{(1)})^2 + (1-\alpha)(p_{k,m}^{(2)})^2. \quad (73)$$

Thus, if there exists any $\alpha \in [0,1]$ such that

$$\alpha (p_{k,m}^{(1)})^2 + (1-\alpha)(p_{k,m}^{(2)})^2 > [\alpha p_{k,m}^{(1)} + (1-\alpha) p_{k,m}^{(2)}]^2, \quad (74)$$

the feasible set associated with (68) is not convex. As $h(\cdot) = (\cdot)^2$ is a convex function, (74) holds for all $\alpha \in (0,1)$. Hence, the joint optimization problem is not convex. □

Proposition A.2. *Consider the optimization problem*

$$\mathbf{P}_S^* = \arg\min_{\mathbf{P}} : ||\hat{\mathbf{B}} - \mathbf{P}\hat{\mathbf{A}}||_F, \quad (75)$$

subject to

$$\hat{\mathbf{q}}^T \mathbf{P} = \hat{\mathbf{q}}^T, \quad (76)$$

where $\hat{\mathbf{q}} = [\mu(\hat{Q}_1), \mu(\hat{Q}_2), \ldots, \mu(\hat{Q}_{\hat{M}})]^T$*, and subject to the elements of* $\mathbf{P} \triangleq [p_{k,m}]_{k,m=1}^{\hat{M}}$ *satisfying*

$$p_{k,m} \geq 0, \quad k,m = 1,2,\ldots,M. \quad (77)$$

This problem is convex.

Proof. As norms are convex functions, the objective function is convex. We proceed by proving that the feasible set is a convex set. Suppose that $\mathbf{P}_1$ and $\mathbf{P}_2$ are both elements of the feasible set. We prove that $\alpha\mathbf{P}_1 + (1-\alpha)\mathbf{P}_2$, for any $\alpha \in [0,1]$, is also an element of the feasible set; i.e.,

$$\hat{\mathbf{q}}^T(\alpha\mathbf{P}_1 + (1-\alpha)\mathbf{P}_2) = \hat{\mathbf{q}}^T \quad (78)$$

and

$$\alpha p_{k,m}^{(1)} + (1-\alpha) p_{k,m}^{(2)} \geq 0, \quad k,m = 1,2,\ldots,\hat{M}, \quad (79)$$

where $\mathbf{P}_1 = [p_{k,m}^{(1)}]_{k,m=1}^{\hat{M}}$ and $\mathbf{P}_2 = [p_{k,m}^{(2)}]_{k,m=1}^{\hat{M}}$.

The left-hand side of (78) is manipulated as

$$\begin{aligned} \hat{\mathbf{q}}^T(\alpha\mathbf{P}_1 + (1-\alpha)\mathbf{P}_2) &= \alpha\hat{\mathbf{q}}^T\mathbf{P}_1 + (1-\alpha)\hat{\mathbf{q}}^T\mathbf{P}_2 \\ &= \alpha\hat{\mathbf{q}}^T + (1-\alpha)\hat{\mathbf{q}}^T \\ &= \hat{\mathbf{q}}^T, \end{aligned} \quad (80)$$

TABLE V
PARAMETERS OF THE POWER LAW CURVE FITTED TO THE MEASURED RECONSTRUCTION TIMES.

Curve	Parameter a	Parameter b
Original solution, experiment 1	1.395×10^{-4}	9.724×10^{-1}
Proposed solution, experiment 1	2.721×10^{-4}	1.033
Original solution, experiment 2	3.321×10^{-5}	9.592×10^{-1}
Proposed solution, experiment 2	7.295×10^{-5}	1.055

where we used the fact that, as elements of the feasible set, both $\mathbf{P}_1$ and $\mathbf{P}_2$ satisfy (76). Thus, the feasible set associated with (76) is convex.

The elements of both $\mathbf{P}_1$ and $\mathbf{P}_2$ are nonnegative due to the matrices' membership of the feasible set (see (77)). As $\alpha \in [0,1]$, the left-hand side of (79) is nonnegative. Thus, the feasible set associated with (77) is convex.

As the intersection of convex sets is convex, the problem's feasible set is also convex. Hence, the problem is convex. □

Proposition A.3. *Consider the optimization problem*

$$\mathbf{Q}_S^* = \arg\min_{\mathbf{Q}} : ||\hat{\mathbf{D}} - \mathbf{Q}\hat{\mathbf{C}}||_F, \quad (81)$$

subject to

$$-\hat{p}_{k,m}^2 \leq q_{k,m} \leq \hat{p}_{k,m}^2, \quad (82)$$

where $\mathbf{Q} \triangleq [q_{k,m}]_{k,m=1}^{\hat{M}}$*. This problem is convex.*

Proof. As norms are convex functions, the objective function is convex. We proceed by proving that the feasible set is a convex set. Suppose that $\mathbf{Q}_1 = [q_{k,m}^{(1)}]_{k,m=1}^{\hat{M}}$ and $\mathbf{Q}_2 = [q_{k,m}^{(2)}]_{k,m=1}^{\hat{M}}$ are both elements of the feasible set. We prove that $\alpha\mathbf{Q}_1 + (1-\alpha)\mathbf{Q}_2$, for any $\alpha \in [0,1]$, is also an element of the feasible set; i.e.,

$$-\hat{p}_{k,m}^2 \leq \alpha q_{k,m}^{(1)} + (1-\alpha) q_{k,m}^{(2)} \leq \hat{p}_{k,m}^2, \quad k,m = 1,2,\ldots,\hat{M}. \quad (83)$$

As $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are both members of the feasible set, (82) implies that their elements satisfy

$$-\hat{p}_{k,m}^2 \leq q_{k,m}^{(1)} \leq \hat{p}_{k,m}^2, \quad k,m = 1,2,\ldots,\hat{M}, \quad (84)$$

and

$$-\hat{p}_{k,m}^2 \leq q_{k,m}^{(2)} \leq \hat{p}_{k,m}^2, \quad k,m = 1,2,\ldots,\hat{M}. \quad (85)$$

Multiplying (84) by α and (85) by $1-\alpha$ and adding these expressions yields (83). This implies that the problem's feasible set is convex, and that the problem itself is convex. □

APPENDIX B
PROOF OF PROPOSITION IV.1

Let $\mathbf{P}_{\tilde{S}}^{(0)} \triangleq [p_{k,m}^{(\tilde{S})}]_{k,m=1}^{\tilde{M}}$ and $\mathbf{P}_{\hat{S}}^{(0)} \triangleq [p_{k,m}^{(\hat{S})}]_{k,m=1}^{\tilde{M}}$. When constructing $\hat{S}$, each branch $\tilde{S}_{j,m}$ is (i) left unchanged if $\tilde{\beta}_{j,m} = 1$, or (ii) modified to have a negative slope with the same codomain as $\tilde{S}_{j,m}$ if $\tilde{\beta}_{j,m} = 0$, or (iii) divided into two branches, each having the same codomain as $\tilde{S}_{j,m}$, if $0 < \tilde{\beta}_{j,m} < 1$. Therefore, if no branch of $\tilde{S}$ over partition $\tilde{\mathcal{R}}^{(m)}$ has codomain $\tilde{Q}_k$, $\hat{S}$ will have no branch over $\hat{\mathcal{R}}^{(m)}$ with codomain $\tilde{Q}_k$. Thus, $p_{k,m}^{(\tilde{S})} = 0 \implies p_{k,m}^{(\hat{S})} = 0$.

Now, suppose $p_{k,m}^{(\tilde{S})} \neq 0$. The canonicity of $\tilde{S}$ implies it has one branch $\tilde{S}_{\tilde{T}(k,m),m}$ over $\tilde{\mathcal{R}}^{(m)}$ with codomain equal to $\tilde{Q}_k$. If $\tilde{\beta}_{\tilde{T}(k,m),m} \in \{0,1\}$, map $\hat{S}$ will have a single branch over $\hat{\mathcal{R}}^{(m)}$ with codomain $\tilde{Q}_k$. This branch has a slope equal to $\pm\tilde{S}'_{\tilde{T}(k,m),m}$. Thus, $p_{k,m}^{(\hat{S})} = |\pm\tilde{S}'_{\tilde{T}(k,m),m}|^{-1} = p_{k,m}^{(\tilde{S})}$. If $0 < \tilde{\beta}_{\tilde{T}(k,m),m} < 1$, map $\hat{S}$ will have two branches over $\hat{\mathcal{R}}^{(m)}$ with codomain $\tilde{Q}_k$. These branches have slopes equal to

$$\hat{S}'|_{\tilde{A}^{(m)}_{\tilde{T}(k,m)}} = \mu(\tilde{Q}_k)/\big[\tilde{\beta}_{\tilde{T}(k,m),m}\mu(\tilde{R}^{(m)}_{\tilde{T}(k,m)})\big] \qquad (86)$$

and

$$\hat{S}'|_{\tilde{B}^{(m)}_{\tilde{T}(k,m)}} = -\mu(\tilde{Q}_k)/\big[(1-\tilde{\beta}_{\tilde{T}(k,m),m})\mu(\tilde{R}^{(m)}_{\tilde{T}(k,m)})\big]. \quad (87)$$

Thus, $p_{k,m}^{(\hat{S})} = |\hat{S}'|_{\tilde{A}^{(m)}_{\tilde{T}(k,m)}}|^{-1} + |\hat{S}'|_{\tilde{B}^{(m)}_{\tilde{T}(k,m)}}|^{-1} = p_{k,m}^{(\tilde{S})}$.

APPENDIX C
PROOF OF PROPOSITION IV.2

Let $\mathbf{Q}_{\tilde{S}} \triangleq [q_{k,m}^{(\tilde{S})}]_{k,m=1}^{\tilde{M}}$ and $\mathbf{Q}_{\hat{S}} \triangleq [q_{k,m}^{(\hat{S})}]_{k,m=1}^{\tilde{M}}$. Using the argument of Appendix B, it follows that $q_{k,m}^{(\tilde{S})} = 0$ implies $q_{k,m}^{(\hat{S})} = 0$. Now, suppose $q_{k,m}^{(\tilde{S})} \neq 0$. If $\tilde{\beta}_{\tilde{T}(k,m),m} \in \{0,1\}$, $\hat{S}$ will have a single branch over $\hat{\mathcal{R}}^{(m)}$ with codomain $\tilde{Q}_k$. This branch has a slope equal $\pm\tilde{S}'_{\tilde{T}(k,m),m}$. Thus, $q_{k,m}^{(\hat{S})} = \pm|\tilde{S}'_{\tilde{T}(k,m),m}|^{-2} = (2\tilde{\beta}_{\tilde{T}(k,m),m} - 1)(p_{k,m}^{(\tilde{S})})^2$. If $0 < \tilde{\beta}_{\tilde{T}(k,m),m} < 1$, map $\hat{S}$ will have two branches over $\hat{\mathcal{R}}^{(m)}$ with codomains $\tilde{Q}_k$. These branches have slopes given by (86) and (87). Thus, $q_{k,m}^{(\hat{S})} = |\hat{S}'|_{\tilde{A}^{(m)}_{\tilde{T}(k,m)}}|^{-2} - |\hat{S}'|_{\tilde{B}^{(m)}_{\tilde{T}(k,m)}}|^{-2} = (2\tilde{\beta}_{\tilde{T}(k,m),m} - 1)(p_{k,m}^{(\tilde{S})})^2$. Equation (63) then follows.

REFERENCES

[1] A. Lasota and M. C. Mackey, *Chaos, Fractals and Noise: Stochastic Aspects of Dynamics,* 2nd ed., New York, NY, USA: Springer, 1994.

[2] A. Boyarsky and P. Góra, *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension,* 1st ed., Boston, MA, USA: Birkhäuser, 1997.

[3] T-Y. Li, "Finite approximation for the Frobenius-Perron operator. A solution to Ulam's conjecture," *J. Approximation Theor.*, vol. 17, no. 2, pp. 177–186, 1976, 10.1016/0021-9045(76)90037-X.

[4] S. H. Isabelle and G. W. Wornell, "Statistical analysis and spectral estimation techniques for one-dimensional chaotic signals," *IEEE Trans. Signal Process.*, vol. 45, no. 6, pp. 1495–1506, 1997, 10.1109/78.599984.

[5] A. M. McDonald, M. A. van Wyk and G. Chen, "The inverse Frobenius-Perron problem: A survey of solutions to the original problem formulation," *AIMS Math.*, vol. 6, no. 10, pp. 11200-11232, 2021, 10.3934/math.2021650.

[6] S. Großman and S. Thomae, "Invariant distributions and stationary correlation functions of one-dimensional discrete processes," *Z. Naturforsch*, vol. 32, no. 12, pp. 1353–1363, 1977, 10.1515/zna-1977-1204.

[7] S. V. Ershov and G. G. Malinetskii, "The solution of the inverse problem for the Perron-Frobenius equation," *USSR Comput. Math. Math. Phys.*, vol. 28, no. 5, pp. 136–141, 1988, 10.1016/0041-5553(88)90022-5.

[8] P. Góra and A. Boyarsky, "A matrix solution to the inverse Perron-Frobenius problem," *Proc. Amer. Math. Soc.*, vol. 118, no. 2, pp. 409–414, 1993, 10.1090/S0002-9939-1993-1129877-8.

[9] F. K. Diakonos and P. Schmelcher, "On the construction of one-dimensional iterative maps from the invariant density: The dynamical route to the beta distribution," *Phys. Lett. A*, vol. 211, no. 4, pp. 199–203, 1996, 10.1016/0375-9601(95)00971-X.

[10] D. Pingel, P. Schmelcher and F. K. Diakonos, "Theory and examples of the inverse Frobenius-Perron problem for complete chaotic maps," *Chaos: Interdisciplinary J. Nonlinear Sci.*, vol. 9, no. 2, pp. 357–366, 1999, 10.1063/1.166413.

[11] N. Santitissadeekorn, "Transport analysis and motion estimation of dynamical systems of time-series data ," Ph.D. dissertation, Clarkson Univ., New York, NY, USA, 2008.

[12] A. G. Lozowski, M. Lysetskiy and J. M. Zurada, "Signal processing with temporal sequences in olfactory systems," *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1268–1275, 2004, 10.1109/TNN.2004.832730.

[13] A. Lasota and P. Rusek, "An application of ergodic theory to the determination of the efficiency of cogged drilling bits," *Archiwum Górnictwa*, vol. 3, pp. 281–295, 1974.

[14] A. M. McDonald and M. A. van Wyk, "Efficient generation of random signals with prescribed probability distribution and spectral bandwidth via ergodic transformations," in *26th IEEE Eur. Signal Process. Conf.,* Rome, Italy, Sept. 3-7, 2018, pp. 331–335, 10.23919/EUSIPCO.2018.8553152.

[15] X. Nie and D. Coca, "A new approach to solving the inverse Frobenius-Perron problem," in *2013 IEEE Eur. Control Conf.,* Zurich, Switzerland, July. 17-19, 2013, pp. 2916–2920, 10.23919/ECC.2013.6669502.

[16] X. Nie and D. Coca, "Reconstruction of one-dimensional chaotic maps from sequences of probability density functions," *Nonlinear Dyn.*, vol. 80, no. 3, pp. 1373-1390, 2015, 10.1007/s11071-015-1949-9.

[17] E. M. Coen, R. G. Gilbert, B. R. Morrison, H. Leube and S. Peach, "Modelling particle size distributions and secondary particle formation in emulsion polymerisation," *Polym.*, vol. 39, no. 26, pp. 7099–7112, 1998, 10.1016/S0032-3861(98)00255-9.

[18] H. Wang, H. Baki and P. Kabore, "Control of bounded dynamic stochastic distributions using square root models: An applicability study in papermaking systems," *Trans. Inst. Meas. Control*, vol. 23, no. 1, pp. 51–68, 2001, 10.1177/014233120102300104.

[19] T. Wigren, "Soft uplink load estimation in WCDMA," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 760–772, 2009, 10.1109/tvt.2008.926210.

[20] A. Baranovsky and D. Daems, "Design of one-dimensional chaotic maps with prescribed statistical properties," *Int. J. Bifurcation Chaos*, vol. 5, no. 06, pp. 1585–1598, 1995, 10.1142/S0218127495001198.

[21] G. D. Birkhoff, "Proof of the ergodic theorem," *Proc. Nat. Acad. Sci. USA*, vol. 17, no. 12, pp. 656–660, 1931, 10.1073/pnas.17.2.656.

[22] X. Nie and D. Coca, "A matrix-based approach to solving the inverse Frobenius-Perron problem using sequences of density functions of stochastically perturbed dynamical systems," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 54, pp. 248–266, 2018, 10.1016/j.cnsns.2017.05.011.

[23] X. Nie, J. Luo, D. Coca, M. Birkin and J. Chen, "Identification of stochastically perturbed autonomous systems from temporal sequences of probability density functions," *J. Nonlinear Sci.*, vol. 28, no. 4, pp. 1467–1487, 2018, 10.1007/s00332-018-9455-0.

[24] X. Nie, D. Coca, J. Luo and M. Birkin, "Solving the inverse Frobenius-Perron problem using stationary densities of dynamical systems with input perturbations," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 90, pp. 1–15, 2020, 10.1016/j.cnsns.2020.105302.

[25] A. M. McDonald and M. A. van Wyk, "A novel approach to solving the generalized inverse Frobenius-Perron problem," in *2020 IEEE Int. Symp. Circuits Syst.,* Online, Oct. 12-14, 2020, pp. 1-5, 10.1109/ISCAS45731.2020.9181115.

[26] R. Penrose, "On best approximate solution of linear matrix equations," *Proc. Cambridge Philosophical Society*, vol. 52, no. 1, pp. 17–19, 1956, 10.1017/S0305004100030929.

André M. McDonald received the B.Eng degree (Computer Engineering) in 2004, and the B.Eng (Hons) and M.Eng degrees (Electronic Engineering) in 2010 from the University of Pretoria, South Africa. He is currently pursuing the Ph.D. degree with the University of the Witwatersrand, South Africa.

Michaël A. van Wyk received the B.Eng. and M.Eng. degrees (Electrical Engineering) and the Ph.D. degree (Applied Mathematics) from Rand Afrikaans University, Johannesburg, South Africa, in 1986, 1989, and 1996, respectively. Dr. van Wyk is Professor of Electrical Engineering, holder of the Carl and Emily Fuchs Research Chair in Systems and Control Engineering, and the Head of the Systems and Control Group at the University of the Witwatersrand, Johannesburg, South Africa since December 2008.