Fully differentiable optimization protocols for non-equilibrium steady states

Rodrigo A. Vargas-Hernández,[1,2] Ricky T. Q. Chen,[2,3] Kenneth A. Jung,[1] and Paul Brumer[1]

[1] *Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada.*
[2] *Vector Institute for Artificial Intelligence, Toronto, Canada.*
[3] *Department of Computer Science, University of Toronto, Canada.*

(Dated: March 24, 2021)

The time-evolution or equations of motions for many systems are usually described by a set of first-order ordinary differential equations, and for a variety of physical observables, the long-time limit or steady state solution is desired. In the case of open quantum systems, the time-evolution of the reduced density matrix is described by the Liouvillian. For inverse design or optimal control of such systems, the common approaches are based on brute-force search strategies. Here, we present a novel methodology, based on automatic differentiation, capable of differentiating the steady state solution with respect to any parameter of the Liouvillian. Our approach has a low memory cost, and is agnostic to the exact algorithm for computing the steady state. We illustrate the advantage of this method by inverse designing the parameters of a quantum heat transfer device that maximizes the heat current and the rectification coefficient. We also optimize the parameters of various Lindblad operators used in the simulation of energy transfer under natural incoherent light.

arXiv:2103.12604v1 [quant-ph] 23 Mar 2021

Introduction.– A common approach to engineer and study physical systems is through inverse design protocols, where a search mechanism is used to navigate the system's parameters space. The goal is to identify the optimal set of parameters that reproduce a target property. To compute these properties we must simulate the time-evolution of the system, which is commonly described by differential equations of the form,

$$\frac{d\rho(t)}{dt} = F(\rho(t);\mathbf{\Theta}), \tag{1}$$

where $F(\cdot)$ describes the change of the physical state ρ over time, and $\mathbf{\Theta} \in \mathbb{R}^d$ is the set of parameters of $F(\cdot)$.

In open quantum systems (OQSs), the time-evolution of a subsystem interacting with either a single or multiple environments, is modeled by a quantum master equation (QME) where ρ is the reduced density matrix, and F is the Liouvillian [1]. The simulation of open quantum many-body systems has lead to the development of various algorithms, e.g., renormalization group [2–4], meanfield methods [5–7], tensor networks [8–12], hierarchical equations of motion[13–15], Heisenberg equation of motion approaches[16, 17], secular and non-secular Redfield theory [18, 19], tensor transfer methods [20–22], and mixed quantum-classical methods [23–25].

Often, the state of interest is the long-time or steady state solution (ρ^{ss}) that satisfies $\frac{d\rho^{\mathbf{ss}}(t)}{dt} = 0$, given that it is required to compute physical observables of the form $\langle\hat{O}\rangle_{\rho^{ss}} = \mathrm{Tr}[\hat{O}\rho^{ss}]$. Recent work to alleviate the large cost in obtaining such results has focused on the parametrization of the steady state solution [26–29], and determining the Liouvillian gap [30] using various deep learning architectures. For time dependent problems, the memory kernel has also been approximated with deep learning methodologies [31, 32].

For the computation of $\langle\hat{O}\rangle_{\rho^{ss}}$, these approaches could alleviate some of the computational cost associated with obtaining ρ^{ss}, however, for the inverse design of an OQS with a target $\langle\hat{O}\rangle_{\rho^{ss}}$ one requires more efficient methodologies to navigate/search in the space of physical models. If one needs to compute the steady state multiple times to optimize an observable the optimization procedure becomes an insurmountable task.

While speeding up the calculation of the steady state solution for an OQS is important, our goal is different; we desire to automate the inverse design of an OQS through properties or observables that depend on the steady state, $\mathbf{\Theta}^* = \arg\max_{\mathbf{\Theta}} f(\rho^{ss}, \mathbf{\Theta})$, where f is a differentiable function and $\mathbf{\Theta}$ are parameters that may affect the steady state. If one could efficiently compute the change of ρ^{ss} with respect to $\mathbf{\Theta}$, without the need to carry out unnecessary steady state computations, it would alleviate this high cost and allow optimizing and studying more complex open quantum systems.

Here, we present a new methodology capable of studying the dependence of quantum observables that depend on ρ^{ss} using automatic differentiation (AD). This procedure is completely general and compatible with any algorithm used to obtain the steady state solution. By knowing the change of ρ^{ss} with respect to any parameter of the system (H_S), bath (H_B) or system-bath (H_{SB}) Hamiltonians for a target property, we could inverse design a OQS through an optimization procedure. We apply this approach to two commonly used open quantum models, the first being the device performance of a quantum heat transfer (QHT) model and the other an incoherently excited exciton transport system. These systems are commonly optimized or studied by brute force search methods, Refs. [33–36]. Here we present an efficient gradient-based approach for inverse design.

Method.– In general, any physical observable that depends on the steady state is a scalar function, *e.g.* $f(\rho^{ss}, \mathbf{\Theta}) = \langle\hat{O}\rangle_{\rho^{ss}}$, whose gradient with respect to

the parameters can be decomposed with the chain rule $\frac{\partial f}{\partial \rho^{ss}}\frac{d\rho^{ss}}{d\mathbf{\Theta}} + \frac{\partial f}{\partial \mathbf{\Theta}}$. Generally, the derivatives $\frac{\partial f}{\partial \rho^{ss}}$ and $\frac{\partial f}{\partial \mathbf{\Theta}}$ can be automatically computed efficiently. On the other hand, the gradient of the steady state with respect to some parameters, $\frac{d\rho^{ss}}{d\mathbf{\Theta}}$, is not readily available, and the standard ways of obtaining these gradients is either analytically (when possible) or via finite differences. However, the latter approach requires computing $\frac{d\rho^{ss}}{d\Theta_i}$ for each single parameter separately, for a total of $\mathcal{O}(d)$ evaluations of the steady state; making it intractable for systems that depend on a larger number of parameters. To make this process more versatile and efficient, we propose using implicit differentiation [37].

Computing gradients through the chain rule is the role of an automatic differentiation (AD) framework, though we note that naïvely using AD is insufficient for computing $\frac{d\rho^{ss}}{d\mathbf{\Theta}}$. While we can solve for ρ^{ss} by running an appropriate ODE solver for a sufficiently long period of time, differentiating through the internals of the ODE solver is prohibitive as it requires storing all intermediate quantities of the solver. There exists low-memory methods for computing gradients of ODE solutions but they require either the trajectory $\rho(t)$ to be stored in memory or solving $\rho(t)$ in reverse time for constant memory [38]. However, the reversing approach is not applicable as the steady state, once reached, cannot be reversed.

To differentiate the steady state solution with respect to any parameter with constant memory usage, we view ρ^{ss} as the solution of a fixed point problem,

$$F(\rho^{ss}; \mathbf{\Theta}) = 0, \tag{2}$$

The Jacobian $\frac{d\rho^{ss}}{d\mathbf{\Theta}}$ then follows from the implicit function theorem [37],

$$\frac{d\rho^{ss}}{d\mathbf{\Theta}} = -\left(\frac{dF(\rho^{ss};\mathbf{\Theta})}{d\rho}\right)^{-1}\left[\frac{dF(\rho^{ss};\mathbf{\Theta})}{d\mathbf{\Theta}}\right]. \tag{3}$$

The Jacobian $\frac{d\rho^{ss}}{d\mathbf{\Theta}}$ only depends on the derivatives of $F(\cdot)$ evaluated at ρ^{ss} and $\mathbf{\Theta}$. Libraries that implement reverse-mode automatic differentiation are comprised of algorithms that can compute vector-Jacobian products, $\mathbf{v}^\top \frac{d\rho^{ss}}{d\mathbf{\Theta}}$, where $\mathbf{v}$ is any vector. We thus propose an efficient algorithm for computing $\mathbf{v}^\top \frac{d\rho^{ss}}{d\mathbf{\Theta}}$ in order to allow any black-box steady state solver to be placed within an AD framework, such as JAX [39].

Computing this vector-Jacobian product using Eq. (3) requires solving $\mathbf{v}^\top J^{-1}$ where $J = \frac{dF(\rho^{ss};\mathbf{\Theta})}{d\rho}$. For large systems where J^{-1} cannot be tractably computed, we propose directly computing the product $\mathbf{v}^\top J^{-1}$. This can be done by recognizing that it is the steady state solution to an ODE of the form, $\frac{d\mathbf{y}}{dt} = \mathbf{y}^\top J - \mathbf{v}$, *i.e.* the steady state satisfies $\mathbf{y}^{ss} = \mathbf{v}^\top J^{-1}$. Notably, simulating this ODE only requires vector-Jacobian products that can be efficiently computed with AD. Thus, to differentiate through the steady state solution of a QME, we essentially solve two steady state problems, both of which can be computed using any black-box steady state solver.

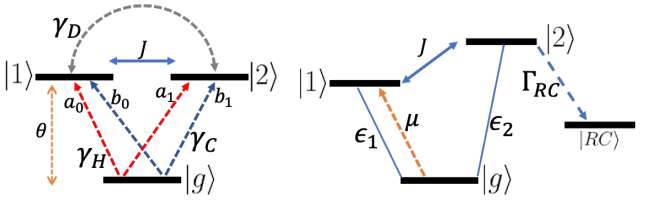

FIG. 1. (upper panel) Diagram of the methodology proposed here. (lower left panel) Diagram for a three level quantum heat transfer model in the local site basis. The dashed arrows represent the interactions with the hot (red) and cold (blue) bath. (lower right panel) Diagram for a three level system with incoherent excitation coupled to a reaction center.

To summarize, we propose a method of efficiently computing exact derivatives for all parameters simultaneously with just a single evaluation of ρ^{ss}, compared to finite difference approximations computed using $\mathcal{O}(d)$ evaluations. In the following sections, we carry out inverse design and sensitivity analysis by differentiating through the steady state of open quantum systems.

Results.- Quantum heat transfer (QHT) models have been proposed as rich systems to study quantum effects in thermodynamics [33, 40–43]. QHT models are commonly studied within the framework of OQS where various approaches of the QME have been applied. In the limit of weak interactions between the system and the baths, the time evolution of a QHT model can be described in a perturbative manner using Redfield theory (RT). RT assumes that the baths are prepared in a canonical thermal state, that the system-bath interaction can be factorized, and the environments are Markovian. The Redfield master equation for multiple non-interacting baths is [1, 19],

$$\frac{\partial \rho_{\mu,\nu}(t)}{\partial t} = -i\left[H_S, \rho(t)\right] + \sum_\alpha \mathcal{D}^\alpha \rho(t) \tag{4}$$

$$= -i\omega_{\mu,\nu}\rho_{\mu,\nu}(t) + \sum_\alpha \sum_{\kappa,\lambda} R^\alpha_{\mu,\nu,\kappa,\lambda}\, \rho_{\kappa,\lambda}(t), \tag{5}$$

where $\mathcal{D}^\alpha$ is the dissipator for each environment, and the Redfield tensor for the α environment is R^α. Each Redfield tensor depends on the system-environment interactions, the spectral density function, $G^\alpha(\omega)$, and the average phonon occupation number, $n_\alpha(\omega)$. Additionally, each environment is characterized by the temperature T_α. For more details about Redfield theory we refer the reader to the Supplemental Material and to Refs. [1, 18, 19].

Quantum heat transfer models are characterized by the change in the system's energy, $\mathrm{d}\langle H_S\rangle/\mathrm{d}t = \mathrm{Tr}\left[H_S\dot{\rho}(t)\right]$, corresponding to the heat flow. This equality only holds if the system's Hamiltonian H_S is time independent. In the long-time limit, $\dot{\rho}(t) = 0$, the energy exchange is comprised of the flow of heat, where the rate of heat

exchange with the α-bath is given by,

$$J_\alpha = \mathrm{Tr}\left[H_S \mathcal{D}^\alpha \rho^{ss}\right]. \tag{6}$$

If one tries to inverse design the system, baths, or the system-baths interactions to maximize J_α, one needs to understand the effect each parameter has in the QME, e.g., $\frac{\partial D^\alpha}{\partial \theta}$ or $\frac{\partial \rho^{ss}}{\partial \gamma_\alpha}$. Here, we combine automatic differentiation and Eq. 3, to inverse design the heat current using gradient based methods. We illustrate this new methodology on a QHT model with a three-level system interacting with three baths, where the system Hamiltonian is,

$$H_S = \epsilon_g \left|g\right\rangle \left\langle g\right| + \theta \sum_{i=1}^{2} \left|i\right\rangle \left\langle i\right| + J\big(\left|1\right\rangle \left\langle 2\right| + \text{h.c.}\big). \tag{7}$$

θ and J describe the energy of sites $|1\rangle$ and $|2\rangle$ and the hopping between them. We set $\epsilon_g = 0$ for reference. The shorthand "h.c." denotes the Hermitian conjugate of former terms in the expressions. For the construction of the Redfield tensors we invoked the Markovian approximation, and each environment's states were described with an ohmic spectral density function and a friction coefficient γ_α. We also neglect the Lamb shift. For more details refer to the Supplemental Material.

The most general type of interactions with the hot (H), cold (C) and decoherence (D) baths are,

$$S^H = a_0 \left|g\right\rangle \left\langle 1\right| + a_1 \left|g\right\rangle \left\langle 2\right| + \text{h.c.} \tag{8}$$

$$S^C = b_0 \left|g\right\rangle \left\langle 1\right| + b_1 \left|g\right\rangle \left\langle 2\right| + \text{h.c.} \tag{9}$$

$$S^D = \left|1\right\rangle \left\langle 2\right| + \text{ h.c.}, \tag{10}$$

where $[a_0, a_1]$ and $[b_0, b_1]$ are the coupling strength parameters to the H and C bath. The $D-$bath is a control mechanism to study the role of coherences between the sites. For this system, analytic results for J_α can only be derived in the secular limit [33].

The parameters of this QHT model can be efficiently optimized by maximizing the heat exchange with the hot bath, J_H. For these simulations, the temperatures are held fixed for all three baths. The final space of parameters is $\mathbf{\Theta} = [\theta, J, a_0, a_1, b_0, b_1, \gamma_H, \gamma_C, \gamma_D]$.

For each optimization procedure, all initial parameters were randomly sampled. Values of a_i and b_i were constrained to $[0, 1]$ to avoid physically incorrect models, and the values of θ and J were constrained to the positive domain by casting them as the exponential function of unconstrained variables. To maximize J_H, we used the Adam [44] optimization algorithm with a learning rate of 0.02. In Ref. [33], this system was studied using fixed parameters $a_0 = b_1 = 1$ and $a_1 = b_0 = 0$. Through optimization, we found that the majority of the optimized systems also recovered these parameter values; however, some of the solutions have both baths interact with a single site, e.g., $a_0 = b_0 \approx 1$ while $a_1 = b_1 \approx 0$ or vice-versa. Fig. 2 contains a *collection* of optimizations, with the optimized values of θ and J indicated and those of the remaining parameters not explicitly shown. In addition,

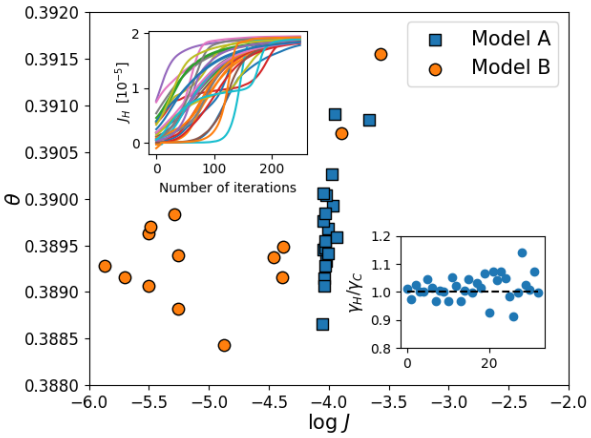

FIG. 2. The markers in the main and lower inset panels represent θ and J the optimal parameter of a quantum heat transfer system, (Fig. 1), obtained by maximizing the heat exchange, Eq. 6, with the Adam algorithm; $\mathbf{\Theta} = [\theta, J, a_0, a_1, b_0, b_1, \gamma_H, \gamma_C, \gamma_D]$. Model A is defined for $a_0 = b_1 \approx 1$ or $a_1 = b_0 \approx 1$, the H and C baths interact with different sites. Model B is defined for $a_i = b_i \approx 1$, the H and C baths interact with the same site. The upper inset figure is the value of J_H as function of the optimization procedure for different random initializations of $\mathbf{\Theta}$.

the the upper inset shows that all of these optimizations converge to essentially the same value of J_H.

γ_H and γ_C are the friction parameters that describe the strength of the coupling to each bath. For each individual optimization we found that at the end of the search procedure, $\gamma_H/\gamma_C \approx 1$. This is an interesting property that can depend on the fixed values of the temperatures. We also found that the value of $\partial J_H/\partial \gamma_D$ is zero in the regions where J_H is maximum. Given the weak-interacting assumption in the Redfield theory, the maximum value allowed for all the γ_i was 0.0025 and the the initial values for these parameters were only sampled from $[10^{-5}, 10^{-4}]$. The gradient of $\partial J_H/\partial \gamma_H$ and $\partial J_H/\partial \gamma_C$ show that their values must increase in order to maximize the heat exchange.

For this case, each iteration requires only two steady state evaluations with our approach, while a finite difference approach would have needed $2\times|\mathbf{\Theta}| = 18$ evaluations of the steady state. Additionally, each gradient-based optimization took less than 150 iterations to find optimal parameters with high precision, Fig. 2. A standard grid-search approach of 10 points for each parameter would have required over 10^9 steady state evaluations.

We also considered a quantum heat transfer model with non-degenerate states,

$$H_S = \epsilon_g \left|g\right\rangle \left\langle g\right| + \sum_{i=1}^{2} \theta_i \left|i\right\rangle \left\langle i\right| + J\big(\left|1\right\rangle \left\langle 2\right| + \text{H.c.}\big), \tag{11}$$

using the same procedure and the same system-bath interactions, Eqs. 9–10. We found that a degenerate system, $\theta_1/\theta_2 \approx 1$ is a system with the highest J_H. The

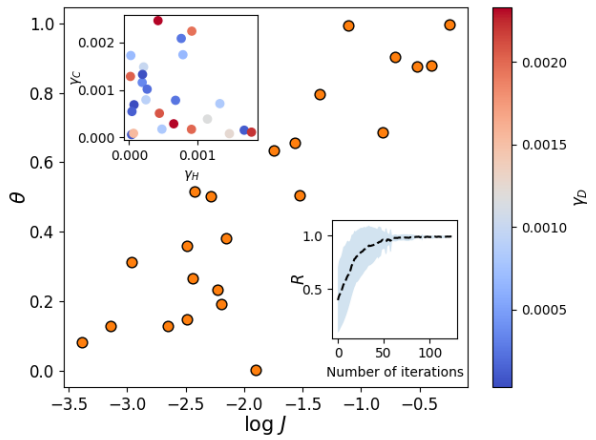

FIG. 3. The markers in the main and upper inset panels represent the optimal parameter of a quantum heat transfer system, (Fig. 1), obtained by maximizing the rectification of heat transfer, R, Eq. 12, with the Adam algorithm; $\mathbf{\Theta} = [\theta, J, \gamma_H, \gamma_C, \gamma_D]$. The hot and cold bath only interact to different sites. The lower inset figure is the value of R as function of the optimization procedure averaged over different random initializations of $\mathbf{\Theta}$. Note that different random initializations lead, after > 50 iterations, to the same value of R.

optimal values of θ_i were similar to the ones for the non-degenerate case, $\theta_1 = \theta_2 \approx 0.389$ (Fig. 2).

Quantum heat transfer models can also be quantified by the rectification coefficient, R, the net heat current when the temperature difference, $T_H - T_C$, in the H and C reservoirs is reversed [45],

$$R = \frac{|J_H| - |J'_H|}{|J_H| + |J'_H|}, \tag{12}$$

where J_H is the heat current when the hot bath interacts with site one and cold bath with site two, and J'_H is computed by swapping the temperatures of the H and C bath. The same methodology used to optimize a quantum heat transfer system where the surrogate observable was J_H can be used to tune the free parameters for R. We use the same gradient-based algorithm, Adam, to maximize R; we only tune θ, J and the friction coefficients for all three baths, $[\gamma_H, \gamma_C, \gamma_D]$. Results of the optimal parameters, as in Fig. 2 for a set of obtained cases, are displayed in Fig. 3. As stressed through this letter, inverse designing quantum heat transfer model with modern gradient methods like Adam, requires very few number of iterations—roughly 100 or less for this case—to find optimal parameter values.

As a second system, we consider a simplified model of energy transfer shown in the right panel in Fig. 1. When the radiation incident on the donor state $|1\rangle$ is taken to originate from an incoherent source, such as the sun, the system is a useful minimal model of biological energy transfer. The steady state efficiency, η_{loc}, is quantified by;

$$\eta_{loc} = \frac{\Gamma_{RC}}{r}\rho_2^{ss}. \tag{13}$$

Here ρ_2^{ss} is the probability of being in the site neighboring the reaction center, ($|2\rangle$), Fig. 1, and Γ_{RC} is the rate of energy transfer from the acceptor state $|2\rangle$ to the reaction center $|RC\rangle$. The time evolution of this system is modeled by,

$$\frac{\partial \rho_S}{\partial t} = \mathcal{L}_0[\rho] + \mathcal{L}_{rad}[\rho] + \mathcal{L}_{deph}[\rho] + \mathcal{L}_{rec}[\rho] + \mathcal{L}_{RC}[\rho], \tag{14}$$

where the first term is the unitary evolution of the system, $\mathcal{L}_0[\rho] = -i\,[H_S, \rho]$, and the rest of the terms, $\mathcal{L}_i$, describe the radiation (rad), the trapping of the excitons at the reaction center (RC), environmental dephasing (deph), and the recombination of the excitons (rec). See the Supplemental Material and Ref. [34] for more information.

We optimize $\boldsymbol{\theta} = [\Gamma, \gamma_d, |\epsilon_1 - \epsilon_2|, J]$ by maximizing η_{loc} using the Adam algorithm, Fig. 4. $|\epsilon_1 - \epsilon_2|$ is the energy difference between site $|1\rangle$ and $|2\rangle$, J is the hopping amplitude, both are system parameters. Γ corresponds to the recombination rate, and γ_d is the phonon bath dephasing rate. For each optimization, we randomly sampled different values for the parameters. We fixed the values of Γ_{RC} and r to $\Gamma_{RC} = 0.5$ ps^{-1} and the incoherent-light pumping rate $r = 6.34 \times 10^{-10}$ ps^{-1}; parameters taken from Ref. [36]. By combining AD and Eq. 3, we used only ~ 20 iterations to find the parameters where $\eta_{loc} \approx 1$, Fig. 4. Given the set of optimal parameters we found a linear correlation between $|\epsilon_1 - \epsilon_2|$ and J. For γ_d the optimal possible values expand a wider range, from 10^9 to 10^{13} Hz; however, for larger values of γ_d the $|\epsilon_1 - \epsilon_2|$ optimal parameters must be greater too. The optimal range for Γ was less wide than the rest and it is aligned with the values used in physical simulations of Ref. [35]. For these simulations, the initial parameters were sampled from a region where η_{loc} is not optimal, however, we managed to optimize the parameters, $\eta_{loc} \approx 1$, regardless of their initial values.

One of the clear applications of gradients is for the maximization/minimization of functions. However, they can also be used to understand the effect that a parameter has on a quantity of interest. For example, studying the effects of the attenuation of the incident radiation is important as photon absorbing centers are found in a variety of environments [46–48]. By computing $\partial\rho^{ss}/\partial r$ we found that $\rho^{ss}_{e_-,e_-}$ is the most sensitive to r, i.e. $\partial\rho^{ss}_{e_-,e_-}/\partial r > \partial\rho^{ss}_{e_+,e_+}/\partial r$, Fig. 5. Here, $e_\pm$ are the eigenstates of the system Hamiltonian, $H_S\,|e_\pm\rangle = \epsilon_\pm\,|e_\pm\rangle$, and $\rho^{ss}_{e_\pm,e_\mp}$ are the matrix elements of the steady state density matrix in the eigenbasis. Furthermore, the imaginary part of the coherence, $\Im(\rho_{e_+,e_-})$, which relates to the exciton flux, Ref. [34], is less sensitive until $r \approx 0.510^{-9}$ ps^{-1}. Indicating that the pumping rate will not increase the flux bellow

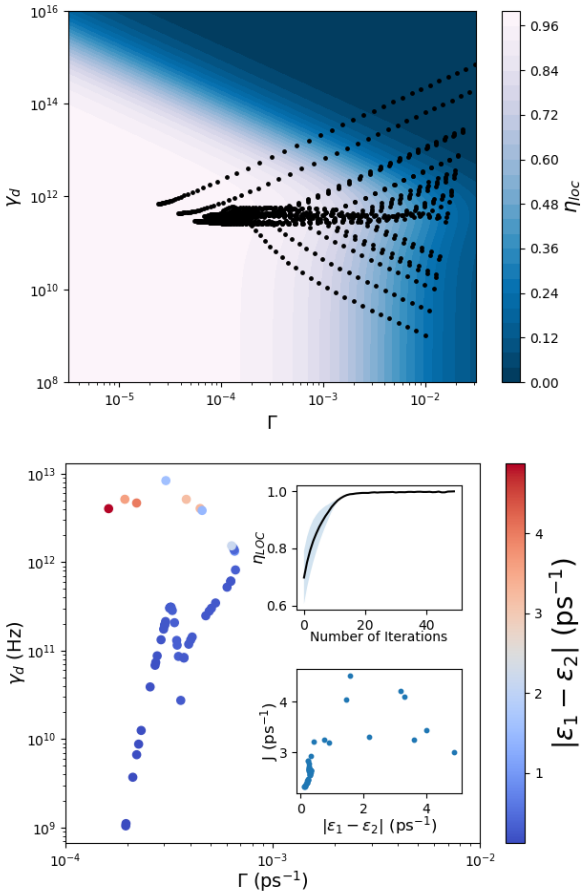

FIG. 4. (upper panel) Optimization steps for η_{loc} for different random initializations, starting in the right with $\eta_{loc} \sim 0$. The markers represent the optimization trajectories that the Adam algorithm follows. (lower panel) Optimal parameters found by maximizing η_{loc} for fixed values of $\Gamma_{RC} = 0.5$ ps^{-1} and $r = 6.34 \times 10^{-10}$ ps^{-1} [35]. The optimal values of Γ, γ_d, J and $|\epsilon_1 - \epsilon_2|$ are presented in the main and inset lower panels. The upper inset figure presents η_{loc} averaged throughout the optimization procedure for different random initialization.

$r \sim 0.510^{-9}$ ps^{-1}.

Summary.– Inverse design and optimal control protocols to optimize open quantum systems that are quantified through observables that depend on the steady state are commonly done with brute-force search methods. On the other hand, gradient-based algorithms have proven to be efficient tools to minimize(maximize) functions. The technical limitation was the inability to efficiently compute the gradient of the steady state with respect to any parameter of the master equation. We circumvent this by combining AD and implicit differentiation. Furthermore, we believe that the present work is the first example of the application of gradient-based methods to efficiently inverse design non-equilibrium steady state systems.

The optimal design of non-equilibrium systems is still driven by physical intuition. However, with the possibility to compute $\partial\rho^{ss}/\partial\mathbf{\Theta}$, we could engineer more robust systems, baths, and system-bath interactions. Furthermore, this methodology could also be used to study the sensitivity of ρ^{ss} with respect to any parameter in the Liouvillian, and, for example, could lead to more insight in how light affects biological processes.

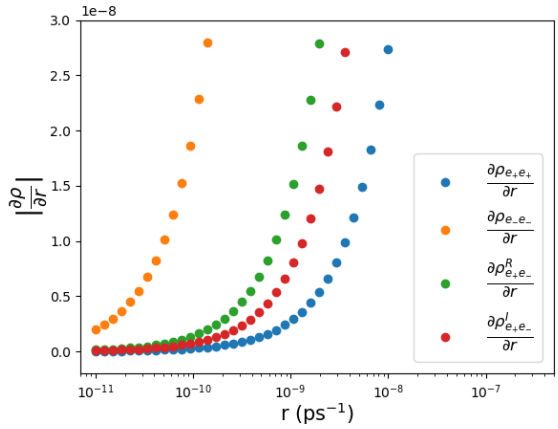

FIG. 5. Sensitivity analysis of ρ^{ss} with respect to the pumping rate, r. We compute $\partial\rho^{ss}_{i,j}/\partial r$ using Eq. 2. For this calculations we used $\Gamma_{RC} = 0.5$ ps^{-1}, $\gamma_d = 1$ Hz, $J = 0.12$ ps^{-1}, and $|\epsilon_1 - \epsilon_2| = 1.3$ ps^{-1}.

In the cases introduced, optimization of a single target quantity (e.g. the heat transfer, or energy efficiency) was carried out in models parametrized by several quantities. As a result, numerous optimized models were obtained, each with similar values of the optimized target. That is, interestingly, the parameter surface has multiple maxima of similar depth. In cases where one is attempting to achieve optimization of multiple quantities (e.g. heat transfer *and* verification) the proposed methodology could be integrated into gradient-based algorithms for multi-objective optimization to construct the Pareto front [49].

In all the simulations presented here, an ODE solver was used to obtain ρ^{ss}; however, Eq. 3 is agnostic to the exact method used to obtain ρ^{ss}. This makes the present methodology particularly valuable. Additionally, this method could be used to study any other system whose observables depend on a long-time solution for equations of the form of Eq. 1. Finally, this methodology opens the possibility of studying more complex quantum heat transfer devices or natural-light induced processes.

We acknowledge fruitful discussions with Dvira Segal. This work was funded by components of two grants from the US Air Force Office of Scientific Research, FA9550-19-1-0267 and FA9550-20-1-0354.

[1] H. P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, Great Clarendon Street, 2002).

[2] S. Finazzi, A. Le Boité, F. Storme, A. Baksic, and C. Ciuti, Corner-space renormalization method for driven-dissipative two-dimensional correlated systems, Phys. Rev. Lett. **115**, 080604 (2015).

[3] R. Rota, F. Minganti, C. Ciuti, and V. Savona, Quantum critical regime in a quadratically driven nonlinear photonic lattice, Phys. Rev. Lett. **122**, 110405 (2019).

[4] R. Rota, F. Storme, N. Bartolo, R. Fazio, and C. Ciuti, Critical behavior of dissipative two-dimensional spin lattices, Phys. Rev. B **95**, 134431 (2017).

[5] A. Biella, J. Jin, O. Viyuela, C. Ciuti, R. Fazio, and D. Rossini, Linked cluster expansions for open quantum systems on a lattice, Phys. Rev. B **97**, 035103 (2018).

[6] J. Jin, A. Biella, O. Viyuela, L. Mazza, J. Keeling, R. Fazio, and D. Rossini, Cluster mean-field approach to the steady-state phase diagram of dissipative spin systems, Phys. Rev. X **6**, 031011 (2016).

[7] O. Scarlatella, A. A. Clerk, R. Fazio, and M. Schiró, Dynamical mean-field theory for open markovian quantum many body systems (2020), arXiv:2008.02563 [cond-mat.stat-mech].

[8] E. Mascarenhas, H. Flayac, and V. Savona, Matrix-product-operator approach to the nonequilibrium steady state of driven-dissipative quantum arrays, Phys. Rev. A **92**, 022116 (2015).

[9] J. Cui, J. I. Cirac, and M. C. Bañuls, Variational matrix product operators for the steady state of dissipative quantum systems, Phys. Rev. Lett. **114**, 220601 (2015).

[10] A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, Positive tensor network approach for simulating open quantum many-body systems, Phys. Rev. Lett. **116**, 237201 (2016).

[11] D. Jaschke, S. Montangero, and L. D. Carr, One-dimensional many-body entangled open quantum systems with tensor network methods, Quantum Science and Technology **4**, 013001 (2018).

[12] A. Kshetrimayum, H. Weimer, and R. Orús, A simple tensor network algorithm for two-dimensional steady states, Nature Communications **8**, 1291 (2017).

[13] Y. Tanimura and R. Kubo, Time evolution of a quantum system in contact with a nearly gaussian-markoffian noise bath, Journal of the Physical Society of Japan **58**, 101 (1989).

[14] Y. Tanimura, Nonperturbative expansion method for a quantum system coupled to a harmonic-oscillator bath, Physical Review A **41**, 6676 (1990).

[15] C. Duan, Z. Tang, J. Cao, and J. Wu, Zero-temperature localization in a sub-ohmic spin-boson model investigated by an extended hierarchy equation of motion, Physical Review B **95**, 10.1103/physrevb.95.214308 (2017).

[16] J. Liu and D. Segal, Generalized input-output method to quantum transport junctions. i. general formulation, Physical Review B **101**, 10.1103/physrevb.101.155406 (2020).

[17] J. Liu and D. Segal, Generalized input-output method to quantum transport junctions. II. applications, Physical Review B **101**, 10.1103/physrevb.101.155407 (2020).

[18] A. G. Redfield, The theory of relaxation processes, in *Advances in Magnetic Resonance*, Advances in Magnetic and Optical Resonance, Vol. 1, edited by J. S. Waugh (Academic Press, 1965) pp. 1 – 32.

[19] D. Egorova, M. Thoss, W. Domcke, and H. Wang, Modeling of ultrafast electron-transfer processes: Validity of multilevel redfield theory, J. Chem. Phys. **119**, 2761 (2003).

[20] J. Cerrillo and J. Cao, Non-markovian dynamical maps: Numerical processing of open quantum trajectories, Physical Review Letters **112**, 10.1103/physrevlett.112.110401 (2014).

[21] A. A. Kananenka, C.-Y. Hsieh, J. Cao, and E. Geva, Accurate long-time mixed quantum-classical liouville dynamics via the transfer tensor method, The Journal of Physical Chemistry Letters **7**, 4809 (2016).

[22] A. Gelzinis, E. Rybakovas, and L. Valkunas, Applicability of transfer tensor method for open quantum system dynamics, The Journal of Chemical Physics **147**, 234108 (2017).

[23] J. C. Tully, Mixed quantum–classical dynamics, Faraday Discussions **110**, 407 (1998).

[24] R. Kapral and G. Ciccotti, Mixed quantum-classical dynamics, The Journal of Chemical Physics **110**, 8919 (1999).

[25] J. E. Subotnik, A. Jain, B. Landry, A. Petit, W. Ouyang, and N. Bellonzi, Understanding the surface hopping view of electronic transitions and decoherence, Annual Review of Physical Chemistry **67**, 387 (2016).

[26] N. Yoshioka and R. Hamazaki, Constructing neural stationary states for open quantum many-body systems, Phys. Rev. B **99**, 214306 (2019).

[27] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Variational neural-network ansatz for steady states in open quantum systems, Phys. Rev. Lett. **122**, 250503 (2019).

[28] A. Nagy and V. Savona, Variational quantum monte carlo method with a neural-network ansatz for open quantum systems, Phys. Rev. Lett. **122**, 250501 (2019).

[29] C. Guo and D. Poletti, Scheme for automatic differentiation of complex loss functions with applications in quantum physics, Phys. Rev. E **103**, 013309 (2021).

[30] D. Yuan, H. Wang, Z. Wang, and D.-L. Deng, Solving the liouvillian gap with artificial neural networks (2020), arXiv:2009.00019 [quant-ph].

[31] M. J. Hartmann and G. Carleo, Neural-network approach to dissipative quantum many-body dynamics, Phys. Rev. Lett. **122**, 250502 (2019).

[32] D. Luo, Z. Chen, J. Carrasquilla, and B. K. Clark, Autoregressive neural network for simulating open quantum systems via a probabilistic formulation (2020), arXiv:2009.05580 [cond-mat.str-el].

[33] M. Kilgour and D. Segal, Coherence and decoherence in quantum absorption refrigerators, Phys. Rev. E **98**, 012117 (2018).

[34] K. A. Jung and P. Brumer, Energy transfer under natural incoherent light: Effects of asymmetry on efficiency, J. Chem. Phys. **153**, 114102 (2020).

[35] T. V. Tscherbul and P. Brumer, Partial secular bloch-redfield master equation for incoherent excitation of multilevel quantum systems, J. Chem. Phys. **142**, 104107 (2015).

[36] T. V. Tscherbul and P. Brumer, Non-equilibrium sta-

tionary coherences in photosynthetic energy transfer under weak-field incoherent illumination, The Journal of Chemical Physics **148**, 124114 (2018).

[37] S. G. Krantz and H. R. Parks, *The implicit function theorem: history, theory, and applications* (Springer Science & Business Media, 2012).

[38] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, Neural ordinary differential equations, in *Advances in neural information processing systems* (2018) pp. 6571–6583.

[39] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, and S. Wanderman-Milne, JAX: composable transformations of Python+NumPy programs (2018).

[40] J. Klatzow, J. N. Becker, P. M. Ledingham, C. Weinzetl, K. T. Kaczmarek, D. J. Saunders, J. Nunn, I. A. Walmsley, R. Uzdin, and E. Poem, Experimental demonstration of quantum effects in the operation of microscopic heat engines, Phys. Rev. Lett. **122**, 110601 (2019).

[41] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—a topical review, Journal of Physics A: Mathematical and Theoretical **49**, 143001 (2016).

[42] R. Kosloff, Quantum thermodynamics: A dynamical viewpoint. entropy, Entropy **15**, 2100 (2013).

[43] N. Linden, S. Popescu, and P. Skrzypczyk, How small can thermal machines be? the smallest possible refrigerator, Phys. Rev. Lett. **105**, 130401 (2010).

[44] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv **1412.6980** (2017).

[45] T. Motz, M. Wiedmann, J. T. Stockburger, and J. Ankerhold, Rectification of heat currents across nonlinear quantum chains: a versatile approach beyond weak thermal contact, New Journal of Physics **20**, 113020 (2018).

[46] S. Axelrod and P. Brumer, An efficient approach to the quantum dynamics and rates of processes induced by natural incoherent light, The Journal of Chemical Physics **149**, 114104 (2018).

[47] S. Axelrod and P. Brumer, Multiple time scale open systems: Reaction rates and quantum coherence in model retinal photoisomerization under incoherent excitation, The Journal of Chemical Physics **151**, 014104 (2019).

[48] C. Chuang and P. Brumer, LH1–RC light-harvesting photocycle under realistic light–matter conditions, The Journal of Chemical Physics **152**, 154101 (2020).

[49] Multiple-gradient descent algorithm (mgda) for multiobjective optimization, Comptes Rendus Mathematique **350**, 313 (2012).

Supplemental material for "Fully differentiable optimization protocols for non-equilibrium steady states"

Rodrigo A. Vargas-Hernández

Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada. and Vector Institute for Artificial Intelligence, Toronto, Canada.

Ricky T. Q. Chen

Vector Institute for Artificial Intelligence, Toronto, Canada. and Department of Computer Science, University of Toronto, Canada.

Kenneth A. Jung and Paul Brumer

Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada.

(Dated: March 24, 2021)

arXiv:2103.12604v1 [quant-ph] 23 Mar 2021

Abstract

The purpose of this supplemental material is to provide details of the numerical calculations we present in this work. Section I presents the core of the algorithm, and sections II B and II A describe the details of the quantum systems used in the main work. In section III we present our numerical solutions of the optimal parameters presented in the main work.

I. AUTOMATIC DIFFERENTIATION

Automatic differentiation (AD) is a framework designed to compute exact vector-Jacobian ($\mathbf{v}^\top J$) and Jacobian-vector ($J\mathbf{v}$) products [1]. The former is often referred to as reverse-mode automatic differentiation, and the latter as forward-mode. For a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ the full Jacobian evaluated at an input value $\mathbf{x}$ is a matrix of m by n:

$$J(\mathbf{x}) = \left[\frac{\partial \mathbf{f}}{\partial x_1}, \ \cdots, \ \frac{\partial \mathbf{f}}{\partial x_n}\right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}, & \cdots, & \frac{\partial f_1}{\partial x_n} \\ \vdots, & \frac{\partial f_i}{\partial x_j}, & \vdots \\ \frac{\partial f_m}{\partial x_1}, & \cdots, & \frac{\partial f_m}{\partial x_n} \end{bmatrix}. \tag{1}$$

For scalar functions, $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^1$ it is clear that with a single vector-Jacobian product it is possible to construct the entire Jacobian. For this case, $\mathbf{v}$ is the scalar 1. In our setting, the optimization objective is a quantum observable, $\langle \hat{O} \rangle = \mathrm{Tr}\,[\hat{O}\rho]$, which can is a scalar function. Thus, with reverse-mode AD, it is possible to construct the entire Jacobian with the same asymptotic cost as a single computation of $\langle \hat{O} \rangle$.

$$\frac{\partial \langle \hat{O} \rangle}{\partial \mathbf{\Theta}} = \left[\frac{\partial \langle \hat{O} \rangle}{\partial \Theta_1}, \ \cdots, \ \frac{\partial \langle \hat{O} \rangle}{\partial \Theta_n}\right] \tag{2}$$

In this work we take advantage of efficient gradient computation to illustrate that complex quantum observables could be numerically optimized by combining AD and gradient-based optimization methods such as Adam [2]. Since quantum observables are often defined as a function of the steady state, we require efficient evaluation of $\frac{\partial \rho^{ss}}{\partial \mathbf{\Theta}}$, where ρ^{ss} is the steady state solution, $\frac{\partial \rho(t)}{\partial t} = 0$, of a first-order differential equation,

$$\frac{\partial \rho(t)}{\partial t} = F(\rho(t); \mathbf{\Theta}). \tag{3}$$

Θ is any set of parameters of $F(\cdot)$, and ρ is the state of interest. Here we studied the evolution of quantum systems that are affected by multiple external environments, and Θ can contain controllable parameters of the environment.

A. Fixed-point problem: computing the Jacobian for the steady state solution

In the steady state limit, $\frac{\partial \rho(t)}{\partial t} = 0$, we can recognize that,

$$F(\rho^{ss}; \mathbf{\Theta}) = 0. \tag{4}$$

Conditioned on a given set of parameters $\mathbf{\Theta}_0$, we denote ρ^{ss} as the solution that satisfies the above equation. By differentiating with respect to $\mathbf{\Theta}_0$ and evaluating at the point $(\rho^{ss}, \mathbf{\Theta}_0)$ we get,

$$\frac{\partial F(\rho^{ss}, \mathbf{\Theta}_0)}{\partial \mathbf{\Theta}_0} + \frac{\partial F(\rho^{ss}, \mathbf{\Theta}_0)}{\partial \rho^{ss}} \frac{\partial \rho^{ss}}{\partial \mathbf{\Theta}_0} = 0. \tag{5}$$

We can re-arrange this equation and obtain an expression for the term of interest:

$$\frac{\partial \rho^{ss}}{\partial \mathbf{\Theta}_0} = -\left[\frac{\partial F(\rho^{ss}, \mathbf{\Theta}_0)}{\partial \rho^{ss}}\right]^{-1} \frac{\partial F(\rho^{ss}, \mathbf{\Theta}_0)}{\partial \mathbf{\Theta}_0}. \tag{6}$$

Any of the terms in Eq. 5 can be computed exactly with AD. In other words, given an $\mathbf{\Theta}_0$ no matter how we solve for ρ^{ss}, we can still compute the Jacobian using just derivative information at the solution point $\mathbf{\Theta}_0$. Also, the above procedure is agnostic to how the steady state is solved, and avoids differentiating through the ODE solver. The memory cost is thus constant and does not scale with the complexity of the problem.

All the simulations in this work were coded in the suite JAX [3]. Figure 1 is the computational diagram of the methodology, presented in the main work, to compute $\frac{d\rho^{ss}}{d\mathbf{\Theta}}$, where $\mathbf{\Theta}$ is the collection of parameters of $F(\ )$.

II. MODEL DESCRIPTIONS

The time-evolution of systems of interest in this work are described within an open quantum system formalism. We describe the steady state of the quantum heat transfer using Redfield theory [4, 5], and for the energy transfer model for the donor-acceptor excitonic model we applied Lindblad [6–8]. Both systems and theories are briefly presented in the following sections. We optimize the parameters of both models using a gradient descent method (Adam).

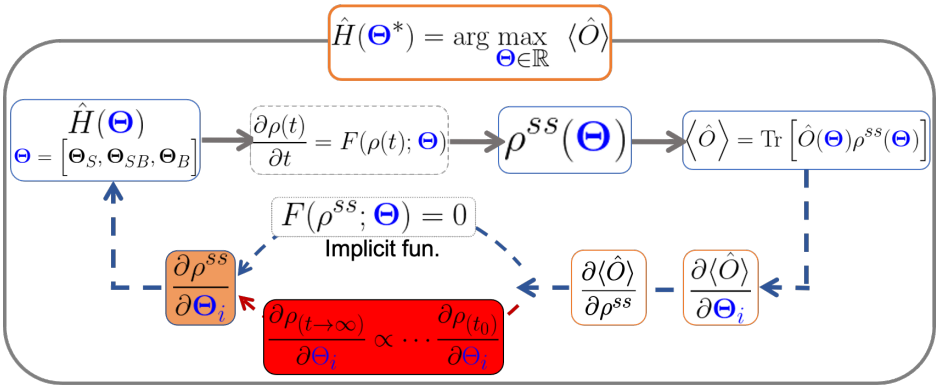

FIG. 1. Diagram of the methodology proposed here. $\boldsymbol{\Theta}$ represents the entire set of parameters of $F(\,)$.

A. Heat transfer model

The system Hamiltonian for a quantum heat transfer model is,

$$H_S = \epsilon_g \left|g\right\rangle \left\langle g\right| + \theta \sum_{i=1}^{2} \left|i\right\rangle \left\langle i\right| + J\big(\left|1\right\rangle \left\langle 2\right| + \text{h.c.}\big), \tag{7}$$

where θ and J describe the energy of sites $|1\rangle$ and $|2\rangle$ and the hopping between them. We set $\epsilon_g = 0$ for reference. This system will be interacting with two baths and is treated using Redfield theory (RT), which is a second-order time-local master equation [4, 6]. For multiple non-interacting baths with factorized initial conditions the Redfield quantum master equation is

$$\frac{\partial\rho_{\mu,\nu}(t)}{\partial t} = -i\left[H_S, \rho(t)\right] + \sum_{\alpha} \mathcal{D}^{\alpha}\rho(t) \tag{8}$$

$$= -i\omega_{\mu,\nu}\rho_{\mu,\nu}(t) + \sum_{\alpha}\sum_{\kappa,\lambda} R^{\alpha}_{\mu,\nu,\kappa,\lambda}\ \rho_{\kappa,\lambda}(t), \tag{9}$$

where $\mathcal{D}^{\alpha}$ is the dissipator for each environment, and the Redfield tensors for the α environment R^{α}, are,

$$R^{\alpha}_{\mu,\nu,\kappa,\lambda} = {}^{\alpha}\Gamma^{+}_{\lambda,\nu,\mu,\kappa} + {}^{\alpha}\,\Gamma^{-}_{\lambda,\nu,\mu,\kappa} - \delta_{\nu,\lambda}\sum_{\ell}\ {}^{\alpha}\Gamma^{+}_{\mu,\ell,\ell,\kappa}$$

$$-\delta_{\mu,\kappa}\sum_{\ell}\ {}^{\alpha}\Gamma^{-}_{\lambda,\ell,\ell,\nu}, \tag{10}$$

where ${}^{\alpha}\Gamma^{\pm}$ are the transition rates,

$$ {}^{\alpha}\Gamma^{+}_{\lambda,\nu,\mu,\kappa} = \langle\lambda|S^{\alpha}|\nu\rangle\langle\mu|S^{\alpha}|\kappa\rangle \int_0^{\infty} \mathrm{d}\tau\ \Omega_{\alpha}(\tau) e^{-i\omega_{\mu,\kappa}\tau}, \tag{11} $$

$$ {}^{\alpha}\Gamma^{-}_{\lambda,\nu,\mu,\kappa} = \langle\lambda|S^{\alpha}|\nu\rangle\langle\mu|S^{\alpha}|\kappa\rangle \int_0^{\infty} \mathrm{d}\tau\ \Omega^{*}_{\alpha}(\tau) e^{-i\omega_{\lambda,\nu}\tau}. \tag{12} $$

$\Omega_{\alpha}(\tau)$ is the correlation function for the α bath,

$$ \Omega_{\alpha}(\tau) = \int_0^{\infty} \mathrm{d}\omega\ G^{\alpha}(\omega) \left[(2n_{\alpha}(\omega)+1)\cos(\omega\tau) - i\sin(\omega\tau)\right], \tag{13} $$

where $G^{\alpha}(\omega)$ is the spectral density function, and $n_{\alpha}(\omega)$ is the average phonon occupation number for the $\alpha-$bath, characterized by the temperature T_{α}. $[\mu,\nu,\kappa,\lambda]$ are the index of the eigenstates of H_S, i.e. they satisfy $H_S|\mu\ \rangle = \epsilon_{\mu}|\mu\rangle$, and $\omega_{\mu,\nu}$ is the difference between eigenvalues ϵ_{μ} and ϵ_{ν}.

The spectral density function of each bath is taken to be of the Ohmic form with an exponential cutoff,

$$ G^{\alpha}(\omega) = \gamma_{\alpha}\omega e^{-\omega/\omega_c}. \tag{14} $$

The golden rule rates in Eq. 12 can be solved in closed form in the Markovain limit to give

$$ {}^{\alpha}\Gamma^{\pm}_{\lambda,\nu,\mu,\kappa} = \langle\lambda|S^{\alpha}|\nu\rangle\langle\mu|S^{\alpha}|\kappa\rangle \left[\Upsilon^{\alpha}(\omega) + i\Delta^{\alpha}(\omega)\right], \tag{15} $$

where Δ^{α} is the Lamb shift, and

$$ \Upsilon^{\alpha}(\omega) = \begin{cases} \frac{1}{2}G^{\alpha}(|\omega|)n_{\alpha}(|\omega|;T_{\alpha}) & \omega < 0 \\ \frac{1}{2}G^{\alpha}(\omega)\left[n_{\alpha}(\omega;T_{\alpha})+1\right] & \omega > 0. \end{cases} \tag{16} $$

The heat transfer will be characterized by the rate of heat exchange with the *hot*-bath,

$$ J_H = \mathrm{Tr}\left[H_S\mathcal{D}^{H}\rho^{ss}\right]. \tag{17} $$

And the rectification coefficient,

$$ R = \frac{|J_H| - |J'_H|}{|J_H| + |J'_H|}. \tag{18} $$

For R, J_H is the heat current when the hot bath interacts with site one and the cold bath with site two, while J'_H is computed by swapping the temperatures of the H and C baths. As we illustrate in the main draft, both functions, J_H and R, can be maximized using gradient ascent based methods.

B. Energy transfer model

The quantum master equation for a donor-acceptor exciton composite system is,

$$\frac{\partial \hat{\rho}_S}{\partial t} = \mathcal{L}_0[\hat{\rho}] + \mathcal{L}_{rad}[\hat{\rho}] + \mathcal{L}_{deph}[\hat{\rho}] + \mathcal{L}_{rec}[\hat{\rho}] + \mathcal{L}_{RC}[\hat{\rho}], \tag{19}$$

where $\mathcal{L}_0[\hat{\rho}] = -i\left[\hat{H}_S, \hat{\rho}_S\right]$ and $\hat{H}_S$ is the Hamiltonian of the system of interest. For this excitonic system,

$$H_S = \sum_{i=1}^{2} \epsilon_i \left|i\right\rangle \left\langle i\right| + J\big(\left|1\right\rangle \left\langle 2\right| + \left|2\right\rangle \left\langle 1\right|\big), \tag{20}$$

where ϵ_i are the onsite-energies for each state $|i\rangle$, and J is the hopping constant [9–11]. The eigenstates of H_S are $|e_\pm\rangle$; $H_S\left|e_\pm\right\rangle = \epsilon_\pm \left|e_\pm\right\rangle$.

The remaining operators in Eq. 19, $\mathcal{L}_i[\cdot]$, describe the radiation (rad), the trapping of the excitons at the reaction center (RC), environmental dephasing (deph), and the recombination of the excitons (rec). Each of the bath operators, except the radiative contribution, are taken to be of the Lindblad form

$$\mathcal{L}_i[\hat{\rho}] = 2K_i \sum_{j=1}^{2} \left(L_i \hat{\rho} L_i^\dagger - \frac{1}{2}[L_i^\dagger L_i, \hat{\rho}]_+ \right), \tag{21}$$

where $[A, B]_+$ is the anticommutator of A and B, and K_i are the dephasing rates. For $\mathcal{L}_{deph}$ we take $K_{deph} = \gamma_d$ as the phonon bath dephasing rate and $L_{deph} = |k\rangle\langle k|$. For $\mathcal{L}_{rec}$ the dephasing corresponding to the recombination rate $K_{rec} = \Gamma$ and $L_{rec} = |g\rangle\langle k|$. Lastly, $K_{RC} = \Gamma_{RC}$ and $L_{RC} = |RC\rangle\langle 2|$. The light-mater interaction term, $\mathcal{L}_{rad}$, derived from the partially-secular Redfield approximation [10–12] and is written in the eigenbasis of H_S as

$$\begin{aligned}
(\mathcal{L}_{\text{rad}}\rho)_{e_\pm e_\pm} &= -(r_{e_\pm} + \gamma_{e_\pm})\rho_{e_\pm e_\pm} + r_{e_\pm}\rho_{gg} \\
&\quad -(\sqrt{r_{e_+} r_{e_-}} + \sqrt{\gamma_{e_+}\gamma_{e_-}})\rho^R_{e_+ e_-} \qquad (22) \\
(\mathcal{L}_{\text{rad}}\rho)_{e_+ e_-} &= -\frac{1}{2}(r_{e_+} + \gamma_{e_+} + r_{e_-} + \gamma_{e_-} + 2i\Delta)\rho_{e_+ e_-} \\
&\quad -\frac{1}{2}(\sqrt{r_{e_+} r_{e_-}} + \sqrt{\gamma_{e_+}\gamma_{e_-}})(\rho_{e_+ e_+} + \rho_{e_- e_-}) \\
&\quad +\sqrt{r_{e_+} r_{e_-}}\rho_{gg}, \qquad (23)
\end{aligned}$$

where $\Delta = \sqrt{(\epsilon_1 - \epsilon_2)^2 - 4J^2}$ is the excitonic splitting, $\gamma_\pm$ are the incoherent emission rates, and $r_\pm$ are the pumping rates for each eigenstate of H_S. The total quantum master

equation is,

$$
\begin{aligned}
\dot{\rho}_{e_+e_+} &= -[2\Gamma + \gamma_d \sin^2(2\theta) + 2\Gamma_{RC}\cos^2(\theta)]\rho_{e_+e_+} \\
&\quad +\gamma_d \sin^2(2\phi)\rho_{e_-e_-} + r_{e_+}\rho_{gg} \\
&\quad +[\Gamma_{RC}\sin(2\phi) - 2\gamma_d \sin(2\theta)\cos(2\phi)]\rho^R_{e_+e_-}, \\
\dot{\rho}_{e_-e_-} &= -[2\Gamma + \gamma_d \sin^2(2\phi) + 2\Gamma_{RC}\sin^2(\phi)]\rho_{e_-e_-} \\
&\quad +\gamma_d \sin^2(2\phi)\rho_{e_+e_+} + r_{e_-}\rho_{gg} \\
&\quad +[\Gamma_{RC}\sin(2\phi) + 2\gamma_d \sin(2\phi)\cos(2\phi)]\rho^R_{e_+e_-}, \\
\dot{\rho}^R_{e_+e_-} &= -[2\Gamma + \Gamma_{RC} + 2\gamma_d(1 - \sin^2(2\phi))]\rho^R_{e_+e_-} \\
&\quad +\Delta\rho^I_{e_+e_-} + \sqrt{r_{e_+}r_{e_-}}\rho_{gg} \\
&\quad +\frac{1}{2}\Gamma_{RC}\sin(2\phi)(\rho_{e_+e_+} + \rho_{e_-e_-}) \\
&\quad +\gamma_d \sin(2\phi)\cos(2\phi)(\rho_{e_-e_-} - \rho_{e_+e_+}), \\
\dot{\rho}^I_{e_+e_-} &= -[2\Gamma + 2\gamma_d + \Gamma_{RC}]\rho^I_{e_+e_-} - \Delta\rho^R_{e_+e_-},
\end{aligned}
\tag{24}
$$

with $r_+ = 2r\cos^2(\phi)$ and $r_- = 2r\sin^2(\phi)$. More details pertaining to this model can be found in Ref. [9].

In energy transfer scenarios one needs a metric to characterize the amount of energy transferred through the system. This is quantified by the exciton efficiency

$$
\eta_{loc} = \frac{\Gamma_{RC}}{r}\ \rho_2^{ss}. \tag{25}
$$

From the above definition, we can observe that η_{loc} is a function of the parameters of the quantum master equation, Eq. 24.

To optimize η_{loc} with gradient descent we require the knowledge of $\frac{\partial\rho_2^{ss}}{\partial\Theta_i}$, where Θ_i is any parameter of the quantum master equation, and ρ_2 is the density matrix of site $|2\rangle$ in the steady state limit.

III. RESULTS

In this section we present the values of the optimal models for a quantum heat transfer model.

For the simulations we fixed $\gamma_H = \gamma_C = 0.002$, $\gamma_D = 0.01$. $T_H = 0.15$, $T_C = 0.1$, and $T_D = 0.12$. We use the Adam optimizer with a learning rate = 0.01 to optimize J_H. To ensure the values of a_0, a_1, b_0, and b_1 are between [0,1], we used the softmax function. θ and J were restricted to only positive values by an exponential function. We report the collection of optimal parameters in Table I.

θ	J	γ_H	γ_C	γ_D	a_0	a_1	b_0	b_1	J_H
0.3900	0.0164	0.0021	0.0020	0.0001	0.9995	0.0005	0.0006	0.9994	1.598e-05
0.3890	0.0172	0.0022	0.0024	0.0001	0.0001	0.9999	0.9998	0.0002	1.789e-05
0.3897	0.0044	0.0025	0.0023	0.0001	0.0007	0.9993	0.0001	0.9999	1.867e-05
0.4022	0.0664	0.0025	0.0025	0.0004	0.9999	0.0001	0.9999	0.0001	1.831e-05
0.3891	0.0180	0.0025	0.0025	0.0002	0.0021	0.9979	0.9999	0.0001	1.920e-05
0.3896	0.0041	0.0025	0.0024	0.0018	0.0002	0.9998	0.0001	0.9999	1.882e-05
0.3898	0.0177	0.0025	0.0024	0.0002	0.9999	0.0001	0.0002	0.9998	1.869e-05
0.3888	0.0052	0.0024	0.0025	0.0001	0.9992	0.0008	0.9998	0.0002	1.908e-05
0.3895	0.0126	0.0025	0.0025	0.0001	0.9992	0.0008	0.9999	0.0001	1.934e-05
0.3893	0.0178	0.0025	0.0025	0.0001	0.0002	0.9998	0.9998	0.0002	1.927e-05
0.8258	0.4442	0.0024	0.0024	0.0014	0.0001	0.9999	0.9999	0.0001	1.013e-05
0.9959	0.6213	0.0025	0.0025	0.0000	0.9997	0.0003	0.9995	0.0005	1.014e-05
0.3892	0.0124	0.0024	0.0025	0.0014	0.9998	0.0002	0.9994	0.0006	1.910e-05
0.3897	0.0041	0.0025	0.0024	0.0019	0.9998	0.0002	0.9999	0.0001	1.888e-05
0.3895	0.0178	0.0025	0.0024	0.0001	0.9994	0.0006	0.0002	0.9998	1.904e-05
0.3885	0.0174	0.0022	0.0025	0.0001	0.0003	0.9997	0.9989	0.0011	1.807e-05
0.3894	0.0060	0.0024	0.0023	0.0004	0.9998	0.0002	0.9998	0.0002	1.821e-05
0.3899	0.0190	0.0025	0.0024	0.0008	0.9998	0.0002	0.0004	0.9996	1.884e-05
0.3893	0.0085	0.0025	0.0025	0.0021	0.9990	0.0010	0.9999	0.0001	1.938e-05
0.3904	0.0178	0.0025	0.0022	0.0005	0.9998	0.0002	0.0002	0.9998	1.816e-05
0.4032	0.0651	0.0025	0.0024	0.0021	0.0015	0.9985	0.9998	0.0002	1.811e-05

TABLE I. Optimal parameters for a quantum heat transfer model found by maximizing J_H with Adam. For each set of parameters, the initial values were sampled randomly. We held fixed $\gamma_H = \gamma_C = 0.002$, $\gamma_D = 0.01$. $T_H = 0.15$, $T_C = 0.1$, and $T_D = 0.12$.

For the quantum heat transfer system were we maximize the rectification coefficient, we optimized $\theta, J, \gamma_H, \gamma_C$, and γ_D. We held fixed $T_H = 0.15$, $T_C = 0.1$, and $T_D = 0.12$. We use the Adam optimizer with a learning rate = 0.01. For these simulations the hot and cold bath only interact with opposite sites. θ and J were restricted to only positive values by an exponential function. We report the collection of optimal parameters in Table II.

θ	J	γ_H	γ_C	γ_D
0.190	0.111	0.0002	0.0012	0.0004
0.313	0.052	0.0001	0.0007	0.0000
0.502	0.102	0.0017	0.0002	0.0001
0.875	0.594	0.0002	0.0008	0.0009
0.127	0.043	0.0008	0.0017	0.0007
0.001	0.150	0.0004	0.0025	0.0023
0.516	0.089	0.0009	0.0002	0.0020
0.146	0.083	0.0002	0.0015	0.0010
0.265	0.087	0.0013	0.0007	0.0008
0.233	0.108	0.0003	0.0010	0.0002
0.685	0.441	0.0007	0.0008	0.0002
0.281	0.356	0.0008	0.0021	0.0003
0.993	0.330	0.0001	0.0001	0.0015
0.997	0.781	0.0002	0.0013	0.0001
0.880	0.669	0.0000	0.0013	0.0020
0.359	0.083	0.0004	0.0005	0.0021
0.128	0.071	0.0000	0.0017	0.0007
0.506	0.218	0.0012	0.0004	0.0012
0.382	0.116	0.0000	0.0005	0.0001
0.635	0.175	0.0005	0.0002	0.0008
0.656	0.210	0.0018	0.0001	0.0022
0.082	0.034	0.0009	0.0022	0.0020
0.797	0.259	0.0015	0.0001	0.0013
0.902	0.494	0.0007	0.0003	0.0023

TABLE II. Optimal parameters for a quantum heat transfer model found by maximizing the rectification coefficient, Eq. 18, with Adam. For each set of parameters, the initial values were sampled randomly. We held fixed $\gamma_H = \gamma_C = 0.002$, $\gamma_D = 0.01$. $T_H = 0.15$, $T_C = 0.1$, and $T_D = 0.12$.

A. Energy transfer model

The optimal collection of parameters depicted in Fig. 4 in the main manuscript are reported in Table III.

J(ps^{-1})	$\lvert\epsilon_1 - \epsilon_2\rvert$ (ps^{-1})	$log(\gamma_d)$ (Hz)	$log(\Gamma)$ (ps^{-1})
2.690	0.257	26.353	-7.613
2.399	0.194	25.021	-8.206
2.652	0.326	25.593	-7.974
2.659	0.266	25.682	-7.801
3.184	0.889	27.915	-7.330
2.642	0.327	26.054	-7.998
2.317	0.118	20.826	-8.536
2.658	0.266	25.606	-7.818
2.373	0.175	24.383	-8.268
2.649	0.312	25.176	-7.954
2.645	0.267	25.134	-7.893
2.825	0.238	27.129	-7.374
2.340	0.145	22.894	-8.396
4.100	3.304	29.026	-7.713
4.511	1.576	29.752	-8.094
3.245	0.760	27.998	-7.345
2.463	0.239	26.003	-8.106
2.406	0.200	25.192	-8.190
2.682	0.260	26.234	-7.650
2.655	0.267	25.494	-7.840
2.698	0.253	26.430	-7.588
2.713	0.255	26.567	-7.542
2.451	0.232	25.886	-8.119
2.828	0.235	27.139	-7.372
3.217	0.413	27.955	-7.335
2.582	0.303	26.451	-8.035
2.460	0.238	25.983	-8.108
3.246	3.596	29.266	-8.549
2.554	0.289	26.449	-8.047
2.631	0.271	24.032	-7.928
2.785	0.255	26.979	-7.414
2.653	0.323	25.470	-7.968
4.203	3.193	29.265	-7.867
2.541	0.284	26.433	-8.053
2.471	0.245	26.084	-8.097
3.295	2.166	28.055	-7.364
2.645	0.267	25.150	-7.892
2.328	0.132	22.032	-8.461
2.315	0.117	20.767	-8.539
2.923	0.324	27.427	-7.321
2.572	0.299	26.459	-8.039
2.564	0.295	26.457	-8.043
3.441	4.014	29.169	-8.419
2.996	4.876	29.022	-8.729
4.039	1.448	28.974	-7.691
2.335	0.140	22.619	-8.418
2.608	0.314	26.371	-8.022
2.397	0.192	24.946	-8.213
2.348	0.151	23.250	-8.368
2.430	0.217	25.610	-8.147

TABLE III. Optimal parameters for a quantum energy transfer model found by maximizing the exciton steady state efficiency, Eq. 25, with Adam. For each set of parameters, the initial values were sampled randomly. We held fixed $\Gamma_{RC} = 0.1$ ps-1, and $r = 6.34 \times 10^{-10}$ ps-1 [12].

IV. REFERENCES

[1] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Automatic differentiation in machine learning: a survey, Journal of Machine Learning Research **18**, 1 (2018).

[2] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv **1412.6980** (2017).

[3] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, and S. Wanderman-Milne, JAX: composable transformations of Python+NumPy programs (2018).

[4] D. Egorova, M. Thoss, W. Domcke, and H. Wang, Modeling of ultrafast electron-transfer processes: Validity of multilevel redfield theory, J. Chem. Phys. **119**, 2761 (2003).

[5] A. G. Redfield, The theory of relaxation processes, in *Advances in Magnetic Resonance*, Advances in Magnetic and Optical Resonance, Vol. 1, edited by J. S. Waugh (Academic Press, 1965) pp. 1 – 32.

[6] H. P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, Great Clarendon Street, 2002).

[7] D. Manzano, A short introduction to the lindblad master equation, AIP Advances **10**, 025106 (2020).

[8] G. Lindblad, On the generators of quantum dynamical semigroups, Communications in Mathematical Physics **48**, 119 (1976).

[9] K. A. Jung and P. Brumer, Energy transfer under natural incoherent light: Effects of asymmetry on efficiency, J. Chem. Phys. **153**, 114102 (2020).

[10] T. V. Tscherbul and P. Brumer, Non-equilibrium stationary coherences in photosynthetic energy transfer under weak-field incoherent illumination, The Journal of Chemical Physics **148**, 124114 (2018).

[11] T. V. Tscherbul and P. Brumer, Long-lived quasistationary coherences in a v-type system driven by incoherent light, Phys. Rev. Lett. **113**, 113601 (2014).

[12] T. V. Tscherbul and P. Brumer, Partial secular bloch-redfield master equation for incoherent excitation of multilevel quantum systems, J. Chem. Phys. **142**, 104107 (2015).