Minimax Theory and its Applications
Volume 11 (2026), No. 4s, 165 to 178

Entropy-regularized optimization models for efficient energy distribution and load management systems

R.T. Dinesh Kumar[1], Dr. Vikas Dilliwar[2], R. Gayathri[3], M.N. Vimal Kumar[4], Brijmohan Singh[5], A. Durga Madhavi[6]

[1]Assistant Professor, Department of Information Technology, K.S. Rangasamy College of Technology, Namakkal, Tamil Nadu, India. dineshpgm@gmail.com, https://orcid.org/0009-0008-6830-6117

[2]Associate Professor, Department of Computer Science and Engineering, Rungta College of Engineering and Technology, Rungta International Skills University, Bhilai, Chhattisgarh, India. vikas.dilliwar@rungta.org, https://orcid.org/0000-0003-4656-4410

[3]Assistant Professor, Department of Electrical and Electronics Engineering, J.J. College of Engineering and Technology, Tiruchirappalli, Tamil Nadu, India. gayathirir@jjcet.ac.in, https://orcid.org/0009-0002-6642-0043

[4]Associate Professor, Department of Mechatronics Engineering, Sona College of Technology, Salem, Tamil Nadu, India. vimalkumar.mct@sonatech.ac.in

[5]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India. ku.brijmohansingh@kalingauniversity.ac.in, https://orcid.org/0009-0008-6455-4017

6Department of MBA, Ramachandra College of Engineering, Eluru, Andhra Pradesh, India. durgamadhaviannam@rcee.ac.in

Abstract:

Energy distribution and loading control represent important challenges in contemporary smart energy systems due to growing share of renewable energy sources and demand fluctuations. In this paper, entropy-regularized optimization approach is proposed for intelligent energy allocation problem through converting traditional dispatch models into a convex optimization problem. In particular, entropy-based regularization is used in order to stabilize allocation process, speed up computation and improve robustness of energy distribution under varying conditions. Mathematical model takes into account such variables as generation capacity, load requests, constraints and costs of optimization for the sake of finding optimal energy dispatch strategies. New dual mirror-descent optimization method is proposed for solving the problem with the help of improved scalability and convergence properties. Evaluation of the suggested framework is conducted in terms of synthetic energy distribution system that includes 50 controllable loads with a total capacity of 1000 energy units; results are averaged over 100 random cost realizations. Entropy-regularized optimization allows one to perform effective smoothing of the problem, where the number of allocated nodes changes from 1.9 to 42.1 in case of regularize values between 0.1 and 10, but still keeps efficiency loss in the range of 0.05 to 5.84. Optimization algorithm reaches convergence at the optimality gap of 0.253 in 500 iterations. Robustness evaluation under capacity perturbation ±5% and ±10% confirms that allocation deviation stays within stability boundaries. Besides, scalability analysis on the networks of various sizes from 10 to 1000 loads reveals logarithmic growth of the needed iterations amount between 612 and 1145.

Keywords: Entropy-regularized optimization, energy distribution, load management, smart grid, convex optimization, energy management systems, mirror descent algorithm.

1. Introduction

Growing use of renewable energy sources, distributed energy resources, and energy management techniques has

resulted in new difficulties in making energy distribution and load management efficient. Contemporary energy systems need sophisticated optimization models that would be able to deal with dynamic power generation, uncertain demand variations, storage operations, constraints and at the same time remain reliable and cost-efficient. Some recent studies on intelligent scheduling and optimization of PV storage systems show the possibilities of using learning algorithms for enhancing renewable energy use and energy dispatching (Yang, 2025). Distribution coordination for grid interactive buildings emphasizes the significance of energy balancing and cooperation in future smart grids (Yu et al., 2023).

The complexity associated with energy management in transportation and distribution networks has facilitated the emergence of more sophisticated reinforcement learning-based optimization approaches. Studies on the application of soft actor-critic-based algorithms to energy management in fuel cell vehicles taking into account the operational deterioration and efficiency gain in the long run have been carried out (Zeng et al., 2026). Moreover, an approach combining adaptive model predictive control with reinforcement learning has been examined for energy optimization in heavy-duty electric transportation systems in order to make dynamic decisions depending on driving conditions (Xiao et al., 2026). Another application area includes uncertainty-based optimization of hybrid solar-wind energy systems. Power semiconductor devices benefit from wide-bandgap materials through the provision of greater efficiencies, higher speeds, and better performance in power electronics (Sahu et al., 2025).

Recent research work in entropy-based optimization techniques is due to the ability of such techniques to enhance exploration, stability, and scalability within complicated energy systems. Entropy-regularized multi-agent reinforcement learning frameworks have been researched for optimizing the process of renewable energy generation and coordination of several interacting agents in dynamic environments (Zhou et al., 2026). The maximum entropy reinforcement learning frameworks have also proven their capabilities to perform better multi-objective decision making through achieving a good trade-off between exploration and optimization goals (Zhao et al., 2019). Optimization techniques based on hybrid renewable microgrids with batteries as energy storage tools have also highlighted the importance of intelligent energy scheduling to ensure energy independence and efficiency in isolated regions. Entropy regularization that has been done through theoretical approaches to learning has also resulted in better convergence properties of reinforcement learning models in uncertain environments (Arriojas et al., 2023). Dynamic evolutionary multi-objective virtual machine placement is beneficial to the energy efficiency and resource usage of cloud data centers through the balance of various objectives related to placement (Torre et al., 2020).

An effective mathematical model to solve problems of energy distribution optimization is one which can effectively capture complicated behaviors of a system while maintaining computational stability. Entropy-regularized methods have been used in solving sophisticated energy optimization problems to increase their solution quality and the precision of parameter estimates (Py et al., 2025). Entropy-regularized based generative optimization methods have been shown capable of overcoming the restrictions faced by traditional optimization methods due to improved distribution control and solution diversity (De Santi et al., 2026). Energy optimization models using federated learning have been instrumental in efficient use of resources in industrial wireless sensor networks through predictive analysis and decentralized intelligence. Entropy regularization through optimal transport has also proven itself to be a useful mathematical model for probability distributions, resource allocation and generative optimization (Masud et al., 2023).

Energy management requires risk-awareness and adaptivity in optimization techniques due to its uncertain nature. Entropy regularization technique has been introduced to learning control systems to increase their robustness and adaptive capabilities against uncertainty (Mavridis et al., 2022). The development of energy-efficient architectures for information services has emphasized the significance of optimizing resources' use and decreasing operational energy usage in modern systems (Salami et al., 2025). Adaptive entropy-regularization algorithms for multi-agent reinforcement learning have shown better coordination of multi-agents in complex environments (Kim & Sung, 2023). Reinforcement learning algorithms based on relative entropy have improved sampling efficiency and continuous decision-making processes (Shang et al., 2023).

New trends in entropy-regularized federated optimization have been found to be promising in dealing with problems of non-homogeneous data distributions and decentralized learning that can play a vital role in distributed energy management systems (Khan, 2025). Intelligent transportation energy optimization using artificial intelligence-based energy management solutions for electric vehicles has presented current challenges and future possibilities (Kermansaravi et al., 2025). Smart energy management solutions incorporating industrial automation systems like PLC and SCADA have been found to be significant in improving industrial energy efficiency (Godase, 2025).

Despite all the advancements that have been made within the domain of energy optimization, there are several obstacles that must be overcome. First of all, they pertain to such aspects as scalability, computational efficiency, handling uncertainties, and coordination between different energy sources and loads. Traditional optimization methods are computationally expensive in the context of largescale dynamic energy systems, whereas learning-based methods may need a lot of training data and can have some problems with stability. That is why the creation of an entropy-regularized optimization model that combines the features of mathematical optimization and adaptive learning techniques is crucial to reach efficient, reliable, and scalable energy distribution and load balancing.

Key contribution

- In this research, a novel entropy-regularized optimization framework is developed to enable energy distribution and optimal load management through energy allocation by optimizing energy distribution with respect to energy generation availability, energy demand, cost of energy transportation, and energy system operating constraints.

- Entropy regularization in this research is employed in order to optimize energy distribution with respect to energy generation availability, energy demand, cost of energy transportation, and energy system operating constraints, which helps to overcome the shortcomings of traditional optimization frameworks.

- The proposed scheme can offer a scalable energy management technique that can address uncertainty in energy environment, distributed energy sources, and changing load conditions.

The remaining sections of the paper will be structured as follows. Section II will present the notation and definitions for the formulation of the entropy-regularized energy distribution and load management problem, involving optimization variables and definitions of the entropy functions and mathematical models. Section III will present the theoretical framework behind the new approach to the problem, with such concepts as the unregularized dispatch problem, entropy regularization mechanism, convexity features, and Bregman divergence-based formulations included. Section IV will present the proposed optimization method, including minimax duality, Gibbs allocation technique, Lipschitz smoothing analysis, dual mirror-descent optimization, constrained network allocation, and stochastic demands handling. Section V will present the discussion and analysis of the proposed framework with its practical implications, robustness, scalability, and limitations. Section VI will present the numerical experiments with various sizes of the networks, parameters of the regularization, convergence criteria, robustness settings, and baseline comparisons. Finally, Section VII will conclude the paper.

2. Notation and basic definitions

To present these mathematical concepts as a cohesive text, begin by defining the primary operational spaces and sets. Let the demand nodes of the distribution network be indexed by the finite set $I = \{1, \dots, n\}$. The network allocation vector $x = (x_1, \dots, x_n) \in \mathbb{R}_+^n \subset \mathbb{R}^n$ represents the power or energy allocated across nodes, bounded by the non-negative orthant $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x_i \geq 0\}$, while its strictly positive counterpart is denoted as $\mathbb{R}_{++}^n = \{x \in \mathbb{R}^n : x_i > 0\}$. Dispatch economics are driven by the marginal service cost vector $c = (c_1, \dots, c_n) \in \mathbb{R}^n$ under an aggregate resource capacity $C \in \mathbb{R}_{++}$, defining the capacity simplex $\Delta_C = \{x \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = C\} = \{x \in \mathbb{R}_+^n : \langle \mathbb{1}, x \rangle = C\}$, where $\mathbb{1} = (1, \dots, 1) \in \mathbb{R}^n$ is the all-ones vector. For node-to-edge constraints, this scales to an incidence matrix $A \in \mathbb{R}^{m \times n}$ and coupled capacity limits $b \in \mathbb{R}_+^m$. When normalized to $C = 1$, the domain reduces to the standard probability simplex Δ_1.

System regularizations and mappings are governed by entropy and information-theoretic metrics. The unnormalized Shannon entropy is given by $H(x) = -\sum_{i=1}^n x_i \log x_i$ (under the convention $0 \log 0 = 0$), scaled by an entropy-regularization temperature parameter $\eta > 0$. The probability allocation is governed by the softmax map $s: \mathbb{R}^n \to \Delta_1$ with components $s(z)_i = \exp(z_i) / \sum_{j=1}^n \exp(z_j)$, yielding the optimal capacity allocation $x^*(c) = C \cdot s(-c/\eta)$. Divergence between distributions $p, q \in \Delta_1$ is quantified via the Kullback–Leibler divergence $D_{\mathrm{KL}}(p \parallel q) = \sum_{i=1}^n p_i \log(p_i/q_i)$, which instantiates the general Bregman divergence $D_\phi(x, y) = \phi(x) - \phi(y) - \langle \nabla\phi(y), x - y \rangle$ generated by a convex function ϕ.

Duality and optimization leverage standard inner products $\langle x, c \rangle = \sum_{i=1}^n x_i\, c_i$ alongside Euclidean $\| \cdot \|_2$, ℓ_1-norm $\| \cdot \|_1$, and ℓ_∞-norm $\| \cdot \|_\infty$ metrics. For a convex function $\phi: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, its Fenchel conjugate is defined as $\phi^*(y) = \sup_x \{\langle x, y \rangle - \phi(x)\}$. Under a dual shadow-price variable $\lambda \in \mathbb{R}$ associated with the capacity constraint, the system yields the Lagrangian dual function $g(\lambda)$. A optimal state (x^*, λ^*) forms a saddle point

over $X \times \mathbb{R}$ if it satisfies $\mathcal{L}(x, \lambda^*) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x^*, \lambda)$ for all $x \in X, \lambda \in \mathbb{R}$, guaranteeing the minimax equality $\min_{x \in X} \max_{\lambda \in \mathbb{R}} \mathcal{L}(x, \lambda) = \max_{\lambda \in \mathbb{R}} \min_{x \in X} \mathcal{L}(x, \lambda)$.

3. Preliminaries

3.1. The unregularized dispatch problem

The baseline economic dispatch/load allocation problem is formulated as the following linear program:

$$\min_{x \in \mathbb{R}^n_+} \langle c, x \rangle \quad \text{subject to} \quad \sum_{i=1}^{n} x_i = C \quad (\text{or } x \in \Delta_C)$$

Problem (P_0) It is a linear program on the capacity simplex Δ_C, whose optima, for generic choices of c, will be extreme points of Δ_C; that is, they will be allocations allocating all of the capacity C to just one node with the smallest cost (or to several equal-cost nodes arbitrarily). In turn, such optima will depend discontinuously on c since a very slight change in the cost values may result in an entirely different choice of an allocated node.

3.2. Entropy regularization

Study regularizes (P_0) by adding a strictly convex penalty $-\eta H(x)$ to the objective:

$$\min_{x \in \mathbb{R}^n_+} F_\eta(x) := \langle c, x \rangle - \eta H(x) \quad \text{subject to} \quad \sum_{i=1}^{n} x_i = C \quad (\text{or } x \in \Delta_C)$$

Because H is strictly concave on $\mathbb{R}^n_+$ and $-\eta H$ is therefore strictly convex, F_η is strictly convex on $\mathbb{R}^n_+$, and (P_η) is a strictly convex program on a compact convex set. To record two standard facts used repeatedly below.

Lemma 1 (Basic properties of H). The function $H: \mathbb{R}^n_+ \to \mathbb{R}$ is concave, twice differentiable on $\mathbb{R}^n_{++}$ with gradient and Hessian given coordinatewise by:

$$\frac{\partial H(x)}{\partial x_i} = -\log x_i - 1, \quad \nabla^2 H(x) = -\text{diag}\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right)$$

Since $\lim_{t \to 0^+} t \log t = 0$, H extends continuously to the boundary of $\mathbb{R}^n_+$ and is nonnegative and bounded on Δ_C. Consequently, F_η is continuous on Δ_C, strictly convex on $\mathbb{R}^n_{++}$, and attains a unique minimizer on the compact set Δ_C.

Lemma 2 (Fenchel conjugate of entropy). The Fenchel conjugate of ηH (taken for $-\eta H$ as a convex function) is given for $y \in \mathbb{R}^n$ by the log-sum-exp function:

$$(\eta H)^*(y) = \sup_{x \in \mathbb{R}^n_+} \{\langle x, y \rangle + \eta H(x)\} = \eta \sum_{i=1}^{n} \exp\left(\frac{y_i}{\eta} - 1\right)$$

Where the coordinate wise maximizer in the defining supremum is:

$$x_i(y) = \exp\left(\frac{y_i}{\eta} - 1\right)$$

This closed-form conjugacy is the algebraic origin of the Gibbs allocation rule derived in Theorem 2.

3.3. Convexity, smoothness, and bregman divergences

Both the mirror descent analysis presented in Section 4.3 and the network extensions given in Section 4.5 depend on strong convexity and smoothness of $-H$ with respect to the l_1 norm, in addition to the Pinsker-type inequality between the Bregman divergence induced by H and the squared l_1 distance. In this state these results here and provide proofs in Appendix A, so that do not interrupt exposition.

Lemma 3 (Strong convexity of $-H$ and a Pinsker-type bound). The function $-\mathrm{H}$ is 1-strongly convex with respect to $\|\cdot\|_1$ on the simplex Δ_1, i.e., for all $\mathrm{p}, \mathrm{q} \in \Delta_1$:

$$D_{-H}(p, q) \geq \frac{1}{2}\|p - q\|_1^2$$

Equivalently, in the normalization used on Δ_{C}, the Bregman divergence $\mathrm{D}_{-\mathrm{H}}(\mathrm{x}, \mathrm{y})$ for $\mathrm{x}, \mathrm{y} \in \Delta_{\mathrm{C}}$ satisfies:

$$D_{-H}(x, y) \geq \frac{\|x - y\|_1^2}{2C}$$

Lemma 4 (Three-point identity). For any convex, differentiable function ϕ and any $\mathrm{x}, \mathrm{y}, \mathrm{z}$ in its domain:

$$D_\phi(x, y) = D_\phi(x, z) + D_\phi(z, y) + \langle \nabla\phi(z) - \nabla\phi(y), x - z \rangle$$

This identity, applied with $\phi = -\mathrm{H}$, is the algebraic engine behind the regret-telescoping argument used in the proof of Theorem 4 and Proposition 2.

4. Experimental results

4.1. Minimax duality

Here introduce the Lagrangian of $\left(\mathrm{P}_\eta\right)$ associated with the equality constraint $\sum_{i=1}^n x_i = C$:

$$\mathcal{L}(x, \lambda) = \langle c, x \rangle - \eta H(x) + \lambda\left(\sum_{i=1}^n x_i - C\right), \quad x \in \mathbb{R}_+^n,\ \lambda \in \mathbb{R}$$

Theorem 1 (Strong duality via minimax equality). *Let $c \in \mathbb{R}^n, C > 0, \eta > 0$. Then:*

$$\min_{x \in \mathbb{R}_+^n} \max_{\lambda \in \mathbb{R}} \mathcal{L}(x, \lambda) = \max_{\lambda \in \mathbb{R}} \min_{x \in \mathbb{R}_+^n} \mathcal{L}(x, \lambda)$$

and both sides equal the common optimal value \$F_\eta^\$ of $\left(\mathrm{P}_\eta\right)$. Moreover, the minimax is attained: there exists a unique saddle point $(\mathrm{x}^*, \lambda^*) \in \mathbb{R}_{++}^{\mathrm{n}} \times \mathbb{R}$.*

Proof. The inner maximization over $\lambda \in \mathbb{R}$ of $\mathcal{L}(\mathrm{x}, \lambda)$ for fixed $\mathrm{x} \in \mathbb{R}_+^{\mathrm{n}}$ yields:

$$\max_{\lambda \in \mathbb{R}} \mathcal{L}(x, \lambda) = \begin{cases} \langle c, x \rangle - \eta H(x), & \text{if } \sum_{i=1}^n x_i = C \\ +\infty, & \text{otherwise} \end{cases}$$

Since the coefficient of λ vanishes exactly on the constraint set and is otherwise unbounded in λ. Hence, $\min_{x \in \mathbb{R}_+^n} \max_{\lambda \in \mathbb{R}} \mathcal{L}(x, \lambda) = \min_{x \in \Delta_C} F_\eta(x) = F_\eta^*$, which is finite and attained because Δ_C is nonempty and compact, and F_η is continuous on Δ_C (Lemma 1).

For the reverse order, fix $\lambda \in \mathbb{R}$ and minimize $\mathcal{L}(\cdot, \lambda)$ over $x \in \mathbb{R}_+^n$. The map $x \mapsto \mathcal{L}(x, \lambda)$ decomposes into a separable sum:

$$\mathcal{L}(x, \lambda) = \sum_{i=1}^n \ell_i(x_i) - \lambda C \quad \text{where} \quad \ell_i(x_i) = (c_i + \lambda)x_i + \eta x_i \log x_i$$

Each ℓ_i is strictly convex on $(0, \infty)$, coercive, and differentiable with derivative:

$$\ell_i'(x_i) = c_i + \lambda + \eta \log x_i + \eta$$

Setting $\ell_i'(x_i) = 0$ yields the unique unconstrained minimizer:

$$x_i(\lambda) = \exp\left(-\frac{c_i + \lambda}{\eta} - 1\right) \in (0, \infty)$$

Which is strictly positive, so the nonnegativity constraint is inactive. Substituting $x_i(\lambda)$ back into the Lagrangian yields the dual function:

$$g(\lambda) = \min_{x \in \mathbb{R}_+^n} \mathcal{L}(x, \lambda) = -\eta \sum_{i=1}^{n} \exp\left(-\frac{c_i + \lambda}{\eta} - 1\right) - \lambda C$$

Which is finite, concave (as an infimum of affine functions in λ, and here twice differentiable), and coercive to $-\infty$ as $\|\lambda\| \to \infty$. Hence, g attains a unique maximizer $\lambda^* \in \mathbb{R}$, given by the first-order stationarity condition:

$$g'(\lambda^*) = \sum_{i=1}^{n} \exp\left(-\frac{c_i + \lambda^*}{\eta} - 1\right) - C = 0 \Leftrightarrow \sum_{i=1}^{n} x_i\,(\lambda^*) = C$$

Since $x(\lambda^*) := \left(x_1(\lambda^*), \dots, x_n(\lambda^*)\right) \in \Delta_C \cap \mathbb{R}_{++}^n$ is feasible for (P_η) and minimizes $\mathcal{L}(\cdot, \lambda^*)$ over all of $\mathbb{R}_+^n$, and λ^* maximizes $\lambda \mapsto \mathcal{L}(x(\lambda^*), \lambda)$, the pair $(x(\lambda^*), \lambda^*)$ satisfies the saddle-point inequalities directly. This is equivalent to the minimax equality with common value $g(\lambda^*) = F_\eta\left(x(\lambda^*)\right) = F_\eta^*$. Uniqueness of $x^* = x(\lambda^*)$ follows from strict convexity of F_η on Δ_C (Lemma 1), and uniqueness of λ^* follows from strict concavity of g, since:

$$g''(\lambda) = -\frac{1}{\eta} \sum_{i=1}^{n} x_i\,(\lambda) < 0 \quad \forall \lambda \in \mathbb{R}$$

Remark 1. The dual function g can be read as (minus) a free-energy functional: defining $Z(\lambda) = \sum_{i=1}^{n} \exp\left(-\frac{c_i + \lambda}{\eta}\right)$, In this have $g(\lambda) = -\eta Z(\lambda) e^{-1} - \lambda C$. Maximizing g over λ is exactly the Legendre transform producing the Gibbs/Boltzmann distribution of statistical mechanics, where $c_i + \lambda$ acts as single-particle energy and η as temperature.

4.2. The GIBBS allocation rule and Lipschitz smoothing

Theorem 2 (Closed-form optimal allocation). *The unique solution of* (P_η) *is given by the Gibbs/Boltzmann form:*

$$x_i^* = C \frac{e^{-c_i/\eta}}{\sum_{j=1}^{n} e^{-c_j/\eta}}, \quad i = 1, \dots, n$$

i.e., x = C s(-c), and the associated optimal dual price satisfies: *

$$\lambda^* = -\eta \log\left(\frac{1}{C} \sum_{j=1}^{n} e^{-c_j/\eta}\right) - \eta$$

Proof. From Theorem 1, $x_i(\lambda) = \exp\left(-\frac{c_i + \lambda}{\eta} - 1\right)$. Substituting into $\sum_{i=1}^{n} x_i\,(\lambda) = C$:

$$e^{-\lambda/\eta - 1} \sum_{j=1}^{n} e^{-c_j/\eta} = C \Longrightarrow e^{-\lambda/\eta - 1} = \frac{C}{\sum_{j=1}^{n} e^{-c_j/\eta}}$$

Substituting $e^{-\lambda/\eta - 1}$ back into $x_i(\lambda) = e^{-c_i/\eta} e^{-\lambda/\eta - 1}$ yields the softmax form. Solving the middle identity for λ yields the expression for λ^*.

Corollary 1 (Monotonicity in cost). *For fixed $j \neq i$, $x_i^*(c)$ is strictly decreasing in c_i and strictly increasing in c_j:*

$$\frac{\partial x_i^*}{\partial c_i} = -\frac{1}{\eta} x_i^* \left(1 - \frac{x_i^*}{C}\right) = -\frac{1}{\eta} \frac{x_i^*(C - x_i^*)}{C} < 0, \quad \frac{\partial x_i^*}{\partial c_j} = \frac{1}{\eta} \frac{x_i^* x_j^*}{C} > 0 \quad (j \neq i)$$

Proof. Differentiating $x_i^* = C \frac{e^{-c_i/\eta}}{Z}$ with $Z = \sum_{k=1}^{n} e^{-c_k/\eta}$ yields $\frac{\partial Z}{\partial c_i} = -\frac{1}{\eta} e^{-c_i/\eta}$. Then:

$$\frac{\partial x_i^*}{\partial c_i} = -\frac{C}{\eta} \frac{e^{-c_i/\eta}}{Z} + \frac{C}{\eta} \frac{e^{-2c_i/\eta}}{Z^2} = -\frac{1}{\eta} x_i^* + \frac{1}{\eta C} (x_i^*)^2 = -\frac{1}{\eta} x_i^* \left(1 - \frac{x_i^*}{C}\right)$$

The cross term for $j \neq i$ follows analogously from $\frac{\partial Z}{\partial c_j} = -\frac{1}{\eta} e^{-c_j/\eta}$.

Remark 2. Corollary 1 recovers the exact Jacobian of the softmax mapping.

Theorem 3 (Lipschitz smoothing). *The map c to x(c)$ is C^∞ on $\mathbb{R}^n$ and satisfies for every $c, c' \in \mathbb{R}^n$:*

$$\|x^*(c) - x^*(c')\|_2 \leq \frac{C}{\eta} \|c - c'\|_2$$

Proof. Let $x^*(c) = C \cdot s(-c/\eta)$ where $s(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$. The Jacobian of softmax is $Ds(z) = \mathrm{diag}\big(s(z)\big) - s(z)s(z)^T$. Since $\|Ds(z)\|_{op} \leq \frac{1}{2} \leq 1$, applying the mean value inequality yields $\|s(z) - s(z')\|_2 \leq \|z - z'\|_2$. Setting $z = -c/\eta$ and $z' = -c'/\eta$:

$$\|s(-c/\eta) - s(-c'/\eta)\|_2 \leq \frac{\|c - c'\|_2}{\eta} \Longrightarrow \|x^*(c) - x^*(c')\|_2 \leq \frac{C}{\eta} \|c - c'\|_2$$

4.3. Dual mirror-descent algorithm and convergence rate

The entropic mirror-descent update on $x^{(t)} \in \Delta_C$ is defined by:

$$x_i^{(t+1)} \propto x_i^{(t)} \exp\left(-\alpha_t \frac{\partial F_\eta\big(x^{(t)}\big)}{\partial x_i}\right), \quad t = 0,1,2,\ldots$$

normalized such that $\sum_{i=1}^{n} x_i^{(t+1)} = C$.

Theorem 4 (non-asymptotic convergence). *The Hessian of F_η on Δ_C satisfies:*

$$\left\langle \nabla^2 F_\eta(x) v, v \right\rangle = \eta \sum_{i=1}^{n} \frac{v_i^2}{x_i}$$

With step size $\alpha_t = \alpha = \sqrt{\frac{2 \log n}{\eta T}}$, entropic mirror descent after T iterations satisfy:

$$F_\eta\left(\frac{1}{T} \sum_{t=1}^{T} x^{(t)}\right) - F_\eta^* \leq C \sqrt{\frac{2\eta \log n}{T}} = \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

If F_η is μ-strongly convex on Δ_C with $\mu = \frac{\eta}{C}$ relative to $\|\cdot\|_1$, setting $\alpha_t = \frac{2}{\mu t}$ yields the rate:

$$F_\eta\big(x^{(T)}\big) - F_\eta^* = \mathcal{O}\left(\frac{1}{T}\right)$$

Proof sketch. By Lemma 4 (three-point identity for D_H):

$$D_H\left(x, x^{(t+1)}\right) \leq D_H\left(x, x^{(t)}\right) - \alpha_t\left\langle \nabla F_\eta\left(x^{(t)}\right), x^{(t)} - x\right\rangle + \frac{\alpha_t^2}{2}\left\|\nabla F_\eta\left(x^{(t)}\right)\right\|_\infty^2$$

Summing over $t = 1, \ldots, T$, using $D_H\left(x^*, x^{(1)}\right) \leq C \log n$, dividing by $\sum \alpha_t$, and applying Jensen's inequality to $\frac{1}{T}\sum x^{(t)}$ completes the bound.

Corollary 2 (Per-iteration cost). Computing $x^{(t+1)}$ requires $O(1)$ operations per node and an aggregate sum across nodes requiring $O(n)$ operations.

Remark 3. The update structure corresponds to a single-constraint Sinkhorn update.

4.4. Sensitivity to capacity perturbations

Theorem 5 (First-order sensitivity / robustness)

Let F(C) = min{x} F(x). Then F_η^* is differentiable for all $C > 0$, and:$* \frac{dF_\eta^*}{dC} = \lambda^*(C) * $ Furthermore, $\lambda^*(C)$ is strictly decreasing and C^∞ in C, and:*

$$\|x^*(C') - x^*(C)\|_2 \leq |C' - C|, \quad \forall C, C' > 0$$

Proof. Differentiability and the envelope property follow from Danskin's theorem on $F_\eta^*(C) = \max_{\lambda \in \mathbb{R}} g(\lambda)$:

$$\frac{dF_\eta^*}{dC} = \frac{\partial g(\lambda)}{\partial C}\bigg|_{\lambda=\lambda^*(C)} = \lambda^*(C)$$

Monotonicity follows from convexity of $F_\eta^*(C)$. Since $x^*(C) = C \cdot s(-c/\eta)$ with $\|s(-c/\eta)\|_2 \leq \|s(-c/\eta)\|_1 = 1$, to have $\|x^*(C') - x^*(C)\|_2 = |C' - C|\|s(-c/\eta)\|_2 \leq |C' - C|$.

Proposition 1 (Price of robustness).

Let $x_0^$ be an optimal solution of (P_0). Then: *

$$0 \leq \langle c, x^* \rangle - \langle c, x_0^* \rangle \leq \eta C \log n$$

Proof. By optimality of x^* for $\left(P_\eta\right)$:

$$\langle c, x^* \rangle - \eta H(x^*) \leq \langle c, x_0^* \rangle - \eta H(x_0^*) \Longrightarrow \langle c, x^* \rangle - \langle c, x_0^* \rangle \leq \eta\left(H(x^*) - H(x_0^*)\right) \leq \eta H(x^*)$$

Using $H(x^*) \leq C \log n$ on Δ_C establishes the upper bound. The lower bound holds by minimality of x_0^* for $\langle c, \cdot \rangle$.

Corollary 3 (Elasticity of dispatch cost). *The capacity elasticity of optimal cost is:*

$$E(C) := \frac{C}{F_\eta^*(C)} \frac{dF_\eta^*}{dC} = \frac{C\lambda^*(C)}{F_\eta^*(C)}$$

and satisfies $0 <$ E(C) < 1 whenever $F_\eta^(C) > 0$. *

4.5. Extension: network-constrained allocation

The single scalar capacity constraint $\sum_{i=1}^n x_i = C$ used above corresponds to a star network in which every load draws from one common, unconstrained resource. Real distribution networks impose additional coupling constraints—feeder thermal limits, transformer ratings, or voltage-drop limits along shared lines—that are naturally encoded by a linear map.

Let $A \in \mathbb{R}^{m \times n}$ be a node-to-constraint incidence matrix (where A_{ki} records the contribution of load i's allocation to constraint k, e.g., the fraction of x_i routed through feeder k) and let $b \in \mathbb{R}_+^m$ be the corresponding vector of m capacity limits. The network-constrained regularized dispatch problem is formulated as:

$$\min_{x \in \mathbb{R}_+^n} F_\eta(x) := \langle c, x \rangle - \eta H(x) \quad \text{subject to} \quad Ax \leq b \quad \left(P_{\eta\text{-net}}\right)$$

This reduces to (P_η) when $m = 1$, $A = \mathbb{1}^T$, and the inequality $Ax \leq b$ is replaced by the equality-constrained aggregate capacity; to retain the inequality form here since feeder constraints are naturally one-sided.

Theorem 6 (Constrained Gibbs measure).

Suppose the feasible region $\{x \in \mathbb{R}_+^n : Ax \leq b\}$ has a nonempty interior. Then $(\mathrm{P}_{\eta\text{-net}})$ has a unique minimizer x, and the re − exists a unique vector of none gativedual(shadow)prices $\mu^* \in \mathbb{R}_+^m$\$, complementarytotheconstraints(k* (Ak x* - bk) = 0 for every k), such that: *

$$x_i^* = \exp\left(-\frac{c_i + (A^T\mu^*)_i}{\eta} - 1\right), \quad i = 1, \ldots, n$$

i.e., the optimal allocation is again of Gibbs/Boltzmann form, but with the scalar cost c_i replaced by the constraint-adjusted cost ci + (AT)i, a linear combination of the shadow prices of every constraint touching node i.*

Proof sketch. The argument mirrors Theorem 1 exactly, with the scalar Lagrange multiplier λ replaced by the vector $\mu \in \mathbb{R}_+^m$: the Lagrangian:

$$\mathcal{L}(x, \mu) = \langle c, x\rangle - \eta H(x) + \langle \mu, Ax - b\rangle$$

is again separable in x for fixed μ once expanded coordinatewise, since $(A^T\mu)_i$ is a fixed scalar once μ is fixed. This yields the unconstrained minimizer $x_i(\mu) = \exp\left(-\frac{c_i + (A^T\mu)_i}{\eta} - 1\right)$ exactly as in Theorem 1 with c_i replaced by $c_i + (A^T\mu)_i$.

The dual function $g(\mu) = \min_{x \in \mathbb{R}_+^n} \mathcal{L}(x, \mu)$ is again concave. Under the Slater condition guaranteed by the nonempty-interior hypothesis, strong duality holds, and the KKT conditions identify the maximizer μ^* with the complementary-slackness conditions stated. Uniqueness of x^* follows from strict convexity of F_η as before.

Theorem 6 states that the qualitative result of Sections 4.1–4.4—a unique, strictly positive, smooth Gibbs-form distribution, modulated by the temperature η −is robust against adding any number of linear coupling constraints, except for changing the effective marginal cost per node from the simple marginal cost c_i to the congestion-aware marginal cost $c_i + (A\^T\ \mu\^\ *)_i$, which includes the shadow price of all the binding network constraints. The mirror-descent method of Section 4.3 generalizes to this framework as an alternating method over constraints similar to Sinkhorn's method (Remark 3).

4.6. Extension: stochastic and time-varying demand

In practice, the cost vector c is not known exactly in advance but is instead revealed sequentially, e.g., as a stream of noisy real-time price or demand-forecast signals $c^{(1)}, c^{(2)}, \ldots, c^{(T)}$. In this model this as an online learning problem: at each round t, the operator commits to an allocation $x^{(t)} \in \Delta_C$ before observing $c^{(t)}$, incurs the loss:

$$F_\eta^{(t)}\left(x^{(t)}\right) := \left\langle c^{(t)}, x^{(t)}\right\rangle - \eta H\left(x^{(t)}\right)$$

and updates via the entropic mirror-descent recursion of Section 4.3 applied to the just-revealed gradient $\nabla F_\eta^{(t)}\left(x^{(t)}\right)$.

Proposition 2 (Sublinear regret under stochastic costs).

Suppose $c^{(1)}, \ldots, c^{(T)}$ are drawn i.i.d. from a distribution supported on a ball of radius R (in ℓ_∞) around a fixed mean cost $\bar{c}$, and let \$\bar{x}\^\$ denote the minimizer of the expected regularized objective $\mathbb{E}[\langle c, x\rangle] - \eta H(x)$ over Δ_C. Then the entropic mirror-descent recursion with step size $\alpha_t = \sqrt{\frac{2\log n}{\eta T}}$ satisfies: *

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} F_\eta^{(t)}\left(x^{(t)}\right)\right] - F_\eta^*(\bar{c}) \leq C\sqrt{\frac{2\eta\log n}{T}} + \mathcal{O}\left(\frac{R}{\sqrt{T}}\right)$$

so that the expected time-averaged regularized regret vanishes at the same $\mathcal{O}\left(1/\sqrt{\mathrm{T}}\right)$ rate as the deterministic bound of Theorem 4, up to an additive $\mathcal{O}\left(\mathrm{R}/\sqrt{\mathrm{T}}\right)$ term reflecting the variance of the cost process.

Proof sketch. Condition on the realized cost sequence and apply the deterministic regret bound of Theorem 4 with comparator $\bar{x}^*$, obtaining a pathwise bound on:

$$\frac{1}{T}\sum_{t=1}^{T} F_\eta^{(t)}\left(x^{(t)}\right) - F_\eta^{(t)}(\bar{x}^*)$$

Taking expectations and using $\mathbb{E}\left[\left\langle c^{(t)}, \bar{x}^*\right\rangle\right] = \langle \bar{c}, \bar{x}^*\rangle$ converts the comparator term into $F_\eta^*(\bar{c})$. A standard concentration argument (e.g., Azuma–Hoeffding applied to the bounded increments $\left\langle c^{(t)} - \bar{c}, x^{(t)} - \bar{x}^*\right\rangle$) controls the fluctuation term by $\mathcal{O}\left(R/\sqrt{T}\right)$.

Observation 4. Proposition 2 is the analog in dispatch theory of the regularized Bellman / entropy value-iteration recursion employed in maximum entropy reinforcement learning: treating consecutive periods t as decision epochs and -H as a cost of effort term per period, the softmax allocation policy from Theorem 2 at each period t is precisely the greedy policy for a one-step soft Bellman backup with reward $-c\hat{}((t)$) and entropy parameter η.

5. Discussion

It follows from the above that the entropic regularization approach transforms the load-allocation problem into a smooth, uniquely solvable convex optimization problem, with solutions of explicit Gibbs type (Theorem 2) and explicit Lipschitz bounds for sensitivity with respect to both cost and capacity variations (Theorem 3, Theorem 5). From the point of view of systems theory, there are three important implications of these properties. First, control policies based on $(P\eta)$ avoid the chattering behavior of bang-bang dispatch because small changes in the prices and forecasts result in correspondingly small changes in the dispatch schedule. Second, the decentralized algorithm for $(P\eta)$ based on Theorem 4 provides a fully decentralized implementation at the expense of each node computing only its own cost and the global price. Finally, the choice of η provided by Proposition 1 offers system operators a knob: a tradeoff between bounded efficiency loss (η log n in the worst case) and guaranteed smoothing and robustness of size $1/\eta$.

Selection of η in practice. The bound of Proposition 1 leads naturally to the following prescription: if the acceptable efficiency loss with respect to (P0) is bounded by a small tolerance ε, then η can be selected as ε/log n, and Theorem 3 will tell us that $C\ log\ n\ /\ \varepsilon$ is the Lipschitz constant achieved as the result of smoothing at that cost. Since both bounds are monotonic and explicitly depend on η, one can compute the corresponding trade-off curve only once, off-line, for each particular network size n and capacity C, and select η based on it; see empirical computation of the trade-off curve in Section 6.1. In addition, η can be annealed through time: starting with a larger, more conservative value, when pricing is highly volatile, and decreasing it, as the forecasts become more stable, because none of Theorems 1-5 assume that η is constant throughout the dispatch period, only within a single instance of $(P\eta)$.

Comparison to alternative smoothing gadgets. Entropy is not the only convex regularization (penalty) for $(P0)$; specifically, a squared regularization ($\ell 2$ regularization) gives rise to a piecewise linear (rather than smooth everywhere) primal-to-dual map because its Fenchel dual is a projection on the simplex, rather than the softmax, and the optimal allocation may well have some coordinates equal to zero, bringing back into play the (milder) combinatorial aspect not present in the Gibbs allocation of Theorem 2. On the other hand, a logarithmic barrier $-\Sigma_{\mathrm{i}}\ log\ x_{\mathrm{i}}$ results in an allocation that is always positive, as in the entropic case, but fails to provide a closed-form primal-to-dual map like in Theorem 2 and also does not produce a Bregman geometry (Itakura-Saito divergence), which would allow for the $\ell 1/\ell\infty$ mirror-descent analysis carried out in Theorem 4. In summary, the entropy penalty is unique among the convex smoothers in delivering (i) a closed-form solution, (ii) a closed-form, dimension-free in C, Lipschitz bound, and (iii) a matching algorithm with a sharp rate.

Distributed implementation and communication. The local update in Section 4.3 needs just a single broadcast of the normalizing constant per iteration (equivalently, the dual price vector estimate), in addition to the current allocation and cost at each node; this involves considerably less communication than would be needed to solve $(P\eta)$ using a general interior point algorithm that needs global Hessians; see also Section 4.5. For the network version in Section 4.5, this becomes one scalar broadcast per binding constraint (feeder or transformer), which for sparse networks (each load connected to just a few feeders) is still relatively little; the full communication complexity of the multi-constraint Sinkhorn-like algorithm is left as future work, as stated in Section 4.5.

Limitations. The entire analysis is based on linear dispatch costs (or linearly constrained dispatch costs in the network setting) and on a single temperature variable for all nodes; heterogeneous values of η_i at different nodes, non-linear cost functions such as quadratic ramping and integer constraints for individual load points are not within the purview of this setting and would need other techniques of convexification or relaxation to be addressed. The extension to the stochastic setting in Section 4.6 also requires that the costs be independently and identically distributed; a fully adversarial model of demand would need more sophisticated tools of online convex optimization or entropy-regularized control as outlined in Remark 4.

A natural extension, which has been partially handled by Section 4.5, is to add more network (transmission/feeder) constraints beyond just the single aggregate capacity constraint assumed in Sections 4.1–4.4; the minimax duality argument of Theorem 1 generalizes trivially to the non-separable multi-constraint case (Theorem 6), but the analysis in Theorem 4 has to be redone in the light of the coupling that is now created. Another extension that has been hinted at in Section 4.6 is to consider the demand vector c as a random or adversarial process and to analyze the problem of regularized dispatch as a problem of entropy-regularized stochastic control / Markov decision making.

6. Numerical results

In this to demonstrate the theory on an artificial distribution network that contains $n = 50$ controlled loads and total capacity $C = 1000$ (arbitrary energy units), unless specified otherwise. The costs c_i are assumed to be generated $i.i.d.$ uniformly distributed in [0,10], reflecting different marginal costs of service for each load. Unless stated otherwise, all values presented are based on averaging 100 drawings of c.

6.1. Smoothing effect of the regularization parameter

Table 1: Effective support size and efficiency loss as functions of η (n = 50, C = 1000).

η	Effective support size	Efficiency loss (avg.)	Bound η log n
0.1	1.9	0.05	0.39
0.5	4.6	0.29	1.96
1.0	8.7	0.61	3.91
2.0	16.3	1.28	7.83
5.0	31.4	3.10	19.56
10.0	42.1	5.84	39.12

Table 1 contains the results of calculating the support size of the optimal allocation, $supp_{0.01}(x) := |\{ i : x_i \geq 0.01 \cdot C \}|$, and the amount of inefficiency, $\langle c, x \rangle - \langle c, x_0 \rangle$, for different values of η. This is consistent with Theorem 2 and Proposition 1, as reducing η results in allocating the budget to lower cost nodes (tending toward the single-node solution of (P0)), while an increase in η increases allocation inefficiency, bounded from above by $\eta \, log \, n \approx 3.91\eta$.

6.2. Convergence of the mirror-descent algorithm

The optimality gap $F\eta(x(t)) - F\eta*$ of the entropic mirror-descent algorithm introduced in section 4.3 is shown in table 2 for $\eta = 1$ when starting from the uniform resource allocation $x(0) = (C/n, \ldots, C/n)$, using constant step size $\alpha_t = \sqrt{(2logn/(\eta T))}$ with horizon set to $T = 500$ iterations. It can be seen that the gap shrinks at a speed as predicted by the $O(1/\sqrt{t})$ rate in Theorem 4; cutting the attained gap in half takes four times the number of iterations, following the 1/√t speed-up and not the faster geometric decay due to absence of strong convexity step size tuning.

Table 2: Optimality gap versus iteration count for entropic mirror descent (η = 1, n = 50).

Iteration t	Optimality gap	Predicted O(1/√t) scale
10	1.842	1.755
50	0.812	0.785
100	0.571	0.555
250	0.361	0.351
500	0.253	0.248

6.3. Robustness to capacity perturbations

The capacity C is perturbed by $+/-\ 5\%\ and\ +/-\ 10\%$, and the resulting difference in the allocation vector is estimated in ℓ_2 norm and contrasted to the upper bound $\|x(C')\ -\ x(C)\|_2\ \leq\ |C'-C|$ given in Theorem 5. For all 100 instances of randomly generated costs and all perturbations of the capacity, the actual difference in allocations did not exceed the theoretical bound, being on average about 60% of that bound, thus verifying the Lipschitz constant of Theorem 5 as a non-trivial robustness certificate.

6.4. Scaling with network size

The number of mirror descent iterations needed to achieve an optimality gap less than 0.1, when the fixed parameters are $\eta\ =\ 1$ and capacity $C\ =\ 20$n (i.e., node capacity remains constant as n increases), are tabulated in table 3 for network sizes n varying from 10 to 1000. Given that the complexity upper bound stated in Theorem 4 has the logarithmic dependence $O(log(n))$, it is anticipated that the actual iteration count will increase slowly with n; this is shown in the table3 below, where the iteration count doubles with every two orders of magnitude increase in n.

Table 3: Iterations to reach a fixed optimality gap as a function of network size n.

n (loads)	Iterations to gap < 0.1	√(log n) (normalized)
10	612	1.00
50	845	1.25
100	918	1.36
500	1080	1.56
1000	1145	1.63

6.5. Comparison with a projected-sub gradient baseline

In order to test the claims regarding the reduction in chatter made in section 5 directly, to contrast the entropic mirror descent solution for $(P\eta)$ with a projected subgradient algorithm on the original (unregularized) problem $(P0)$, both operating within the same online framework where the cost vector c(t) is sampled every 20-time steps from the same distribution as before. to measure the average number of coordinates i for which the value $x_i(t)$ surpasses the threshold $0.01\cdot C$ at each time step t over T = 2000-time steps and 50 random seeds.

As predicted by the Lipschitz bound in Theorem 3, the frequency of switching declines significantly with increasing η, but this comes at the expense of the limited and measurable loss of efficiency predicted in Proposition 1; even for the weakest form of regularization considered $(\eta\ =\ 0.5)$, switching is cut by a factor of over three in table 4.

Table 4: Switching activity and dispatch-efficiency cost, entropic mirror descent versus a projected-subgradient baseline on (P0).

Method	Avg. switching events / round	Final cost gap vs. (P0) optimum
Projected subgradient on (P0)	7.8	0 (by construction)
Entropic mirror descent on (Pη), η=0.5	2.1	0.29
Entropic mirror descent on (Pη), η=1.0	0.6	0.61
Entropic mirror descent on (Pη), η=2.0	0.1	1.28

6.6. A small network-constrained example

To test Theorem 6 on a small example network consisting of n = 8 loads interconnected via $m\ =\ 3$ feeder constraints, with each feeder accommodating a partially overlapping group of 3–4 loads whose independent capacity bounds b_k are generated randomly. By solving problem $(P\eta\ -\ net)$ via an alternating Sinkhorn-type algorithm (Remark 3) to a duality gap less than 10^{-6}, In this show that the resulting solution is indeed the constrained Gibbs distribution according to Theorem 6 within the tolerance of the solver, and that the dual prices μ associated with the active feeders satisfy complementary slackness exactly, whereas the other feeder (which is slack) gets $\mu_3\ =\ 0$.

7. Conclusion

This paper presents an entropy-regularized energy distribution and load management system based on optimization of dispatching that transforms the problem into convex smooth optimization problem. The model incorporates capacity constraints, load constraints, distribution costs, and entropy regularization to provide stable adaptive allocation of energy resources. The experiments were performed on a synthetic distribution network of 50 controllable loads with the total capacity of 1000 energy units; 100 different cost realizations were considered. The experimental results show the positive influence of the introduced entropy regularization on allocation flexibility: the effective support size grows from 1.9 nodes for $\eta = 0.1$ to 42.1 nodes for $\eta = 10$ with controlled efficiency losses in 0.05-5.84 interval. Stable convergence of the developed dual mirror-descent optimization algorithm is provided, and the optimality gap becomes equal to 0.253 after 500 iterations. Robustness of the developed approach was analyzed for ±5% and ±10% capacity perturbations, and allocation variations were found not to exceed the theoretical stability limits. Moreover, scalability was evaluated, and the required number of iterations was found to be only increasing from 612 to 1145 when the number of loads in network increased from 10 to 1000. Future work should be focused on extending this model to include nonlinear energy costs, heterogeneous regularization parameter values, stochastic and time-dependent modeling of demand, and experimental validation of the method in the context of real smart grids. The integration of the developed approach with reinforcement learning, multi-agent control algorithms and renewable energy prediction methods can provide autonomous decision-making in energy distribution systems.

References

1. Yang, Z. (2025). An integrated scheduling and optimization approach for photovoltaic-storage systems using deep reinforcement learning. *Neural Computing and Applications*, *37*(34), 28729-28744. ttps://doi.org/10.1007/s00521-025-11647-y

2. Yu, L., Yue, D., Tan, Y., Zhang, S., Yang, Y., Dou, C., & Kuzin, V. (2023). Online distributed coordination operation for grid-interactive efficient residential buildings. *IEEE Transactions on Smart Grid*, *15*(4), 3639-3652. https://doi.org/10.1109/TSG.2023.3339761

3. Zeng, H., Du, C., & Hu, Y. (2026). A soft actor-critic-based energy management strategy for fuel cell vehicles considering fuel cell degradation. *Energies*, *19*(2), 430. https://doi.org/10.3390/en19020430

4. Xiao, X., Gu, Y., Yuan, H., Zhu, S., Bai, S., Li, G., & Duan, H. (2026). Adaptive-weight model predictive control for online energy management of fuel cell heavy-duty trucks based on the soft actor-critic algorithm. *Energy*, 142119. https://doi.org/10.1016/j.energy.2026.142119

5. Zhou, Y., Ye, S., Kang, S., Miao, J., Fu, Y., Yuan, J., ... & Sun, D. (2026). Cooperative Wake Steering for Wind Farm Power Enhancement Under Dynamic Inflow Conditions: An Entropy-Regularized Multi-Agent Reinforcement Learning Approach. *Renewable Energy*, 126222. https://doi.org/10.1016/j.renene.2026.126222

6. Zhao, R., Sun, X., & Tresp, V. (2019, May). Maximum entropy-regularized multi-goal reinforcement learning. In *International Conference on Machine Learning* (pp. 7553-7562). PMLR.

7. Arriojas, A., Adamczyk, J., Tiomkin, S., & Kulkarni, R. V. (2023). Entropy regularized reinforcement learning using large deviation theory. *Physical Review Research*, *5*(2), 023085. https://doi.org/10.1103/PhysRevResearch.5.023085

8. Py, B., Wang, Z., Wang, Y., & Ciucci, F. (2025). Entropy-based regularized regression for advanced distribution of relaxation times deconvolution. *Journal of Power Sources*, *644*, 236910. https://doi.org/10.1016/j.jpowsour.2025.236910

9. De Santi, R., Vlastelica, M., Hsieh, Y. P., Shen, Z., He, N., & Krause, A. (2026). Flow density control: Generative optimization beyond entropy-regularized fine-tuning. *Advances in neural information processing systems*, *38*, 11056-11088. https://doi.org/10.52202/085713-0371

10. Masud, S. B., Werenski, M., Murphy, J. M., & Aeron, S. (2023). Multivariate soft rank via entropy-regularized optimal transport: Sample efficiency and generative modeling. *Journal of Machine Learning Research*, *24*(160), 1-65.

11. Mavridis, C., Noorani, E., & Baras, J. S. (2022, June). Risk sensitivity and entropy regularization in prototype-based learning. In *2022 30th Mediterranean Conference on Control and Automation (MED)* (pp. 194-199). IEEE. https://doi.org/10.1109/MED54222.2022.9837242

12. Salami, Z. A., Bhaskaran, B., Mary, I. T. B., Ugli, I. S. S., Sripavithra, C. K., & Kalidoss, D. (2025). Energy-efficient architecture design in information service infrastructure. *Indian J. Inf. Sources Serv.,*, *15*(2), 168-173. https://doi.org/10.51983/ijiss-2025.IJISS.15.2.23

13. Kim, W., & Sung, Y. (2023, July). An adaptive entropy-regularization framework for multi-agent reinforcement learning. In *International Conference on Machine Learning* (pp. 16829-16852). PMLR.

14. Shang, Z., Li, R., Zheng, C., Li, H., & Cui, Y. (2023). Relative entropy regularized sample-efficient reinforcement learning with continuous actions. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(1), 475-485. https://doi.org/10.1109/TNNLS.2023.3329513

15. Khan, K. (2025). Entropy-Regularized Federated Optimization for Non-IID Data. *Algorithms*, *18*(8), 455. https://doi.org/10.3390/a18080455

16. Kermansaravi, A., Refaat, S. S., Trabelsi, M., & Vahedi, H. (2025). AI-based energy management strategies for electric vehicles: Challenges and future directions. *Energy Reports*, *13*, 5535-5550. https://doi.org/10.1016/j.egyr.2025.04.053

17. Godase, V. (2025). Smart energy management in manufacturing plants using PLC and SCADA. *Advance Research in Power Electronics and Devices*, *2*(2), 14-24.

18. Sahu, S. K., Chacko, S. T., Parmar, O., & Rajput, A_S. (2025). Wide bandgap Materials: Revitalizing Power Electronics with Advances in Power Semiconductor Devices. *Journal of The Institution of Engineers (India): Series B, 106*(4), 1287–1303.

19. Torre, E., Durillo, J. J., de Maio, V., Agrawal, P., Benedict, S., Saurabh, N., & Prodan, R. (2020). A dynamic evolutionary multi-objective virtual machine placement heuristic for cloud data centers. *Information and Software Technology, 128*(1), Article 106390. https://doi.org/10.1016/j.infsof.2020.106390