Check for
updates

Placing Green Bridges Optimally, with a Multivariate Analysis

Till Fluschnik[1] · Leon Kellerhals[1]

Abstract

We study the problem of placing wildlife crossings, such as green bridges, over human-made obstacles to challenge habitat fragmentation. The main task herein is, given a graph describing habitats or routes of wildlife animals and possibilities of building green bridges, to find a low-cost placement of green bridges that connects the habitats. We develop three problem models for this task and study them from a computational complexity and parameterized algorithmics perspective.

Keywords Wildlife crossings · Computational sustainability · Parameterized algorithmics · NP-hardness · Connected subgraphs

1 Introduction

Sustainability is a major concern impacting today's politics, economy, and industry. Accordingly, sustainability sciences are well-established by now. Yet, the *interdisciplinary* scientific field "computational sustainability" [20, 21], which combines practical and theoretical computer science with sustainability sciences, is quite young.

✉ Leon Kellerhals
  leon.kellerhals@tu-berlin.de

  Till Fluschnik
  till.fluschnik@tu-clausthal.de

[1] Technische Universität Berlin, Faculty IV, Algorithmics and Computational Complexity, Berlin, Germany

🖄 Springer

For instance, the Institute for Computational Sustainability at Cornell University was founded in 2008, the 1st International Conference on Computational Sustainability (CompSust'09) took place in 2009, and special tracks on computational sustainability and AI were established at AAAI [17] and IJCAI [35]. This work contributes to computational sustainability: We model problems of elaborately placing wildlife crossings and give complexity-theoretical and algorithmic analysis for each. Wildlife crossings are constructions (mostly bridges or tunnels [33]) that allow wildlife animals to safely cross human-made transportation lines (e.g., roads). We will refer to wildlife crossings as *green bridges*.

There are numerous reports on wildlife-vehicle collisions [22, 26, 37]. Huijser et al. [26] identify several endangered animal species suffering from high road mortality and estimate the annual cost associated with wildlife-vehicle collisions with around 8 billion US dollars. Wildlife fencing with wildlife crossings can reduce collisions by over 80% [26], enables populations to sustain [36], and are thereby among the most cost-effective [25]. The implementation, though, is a delicate problem, as depicted by Huijser et al. [26, p.16]:

> The location, type, and dimensions of wildlife crossing structures must be carefully planned with regard to the species and surrounding landscape. For example, grizzly bears, deer, and elk tend to use wildlife overpasses to a greater extent than wildlife underpasses, while black bears and mountain lions use underpasses more frequently than overpasses. In addition, different species use different habitats, influencing their movements and where they want to cross the road.

Apart from these delicacies, another challenge is to obtain good data about the specific areas inhabited by a species [39]: While it is arguably easier to answer whether some animal species habitates a certain patch of land in the positive, it seems more challenging to rule it out. Clearly, high data quality is a crucial for deciding on where to place green bridges.

In this work, we consider the task of (re-)connecting habitats under varying connectivity requirements by placing as few green bridges as possible, thus minimizing the cost. We assume to be given a set of land patches which are disconnected by roads, the set of inhabited patches for each animal, and possible locations for green bridges, each of which connects two patches. This is canonical to model as a graph: vertices represent the land patches, edges represent the possible locations for green bridges, and for each animal species we are given a vertex subset of the inhabited patches. The goal in the model now is to find an edge set that sufficiently connects the habitats of each species. In particular, we comparatively study in terms of computational complexity and parameterized algorithmics the following three different (families of) decision problems.[1]

---

[1] The d-th power G^d of a graph G contains edge $\{v, w\} \in \binom{V(G)}{2}$ if and only if $\text{dist}_G(v, w) \le d$.
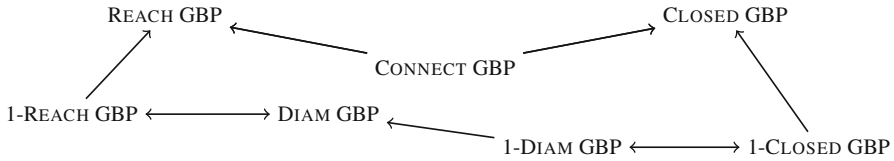
Fig. 1 A diagram of interconnections between the problems (for the definition of CONNECT GBP see Problem 1). An edge from problem A to problem B means that any solution to A is also a solution to B. Problems with d omitted from the problem name require that there is a solution for some value of d

Π GREEN BRIDGES PLACEMENT (Π GBP)

Input: An undirected graph $G = (V, E)$, a set $\mathcal{H} = \{V_1, \ldots, V_r\}$ of habitats where $V_i \subseteq V$ for all $i \in \{1, \ldots, r\}$, and $k \in \mathbb{N}_0$.

Question: Is there an edge set $F \subseteq E$ with $|F| \leq k$ such that for every $i \in \{1, \ldots, r\}$, it holds that $V_i \subseteq V(G[F])$ and

Π ≡ d- REACH:	$G[F]^d[V_i]$ is connected?	(Problem 2)	(Section 3)
Π ≡ d- CLOSED:	$G[F]^d[V_i]$ is a clique?	(Problem 6)	(Section 4)
Π ≡ d- DIAM(ETER):	$\mathrm{diam}(G[F][V_i]) \leq d$?	(Problem 7)	(Section 5)

Our problems address both the challenge in obtaining high quality data as well as the question to what connectivity is sufficient. Connectivity is addressed by the different requirements on the solution: While d- REACH GBP simply ensures connectivity of each habitat along length-d paths, d- CLOSED GBP additionally requires every two patches of each habitat to be connected by such a path. The latter is also true for d- DIAM GBP, which additionally requires that such a path only uses the habitat's patches. In this sense, d- DIAM GBP generalizes 1- REACH GBP. Moreover, d- DIAM GBP and d- CLOSED GBP are equivalent for $d = 1$. See Fig. 1 for relationships between the problems in terms of Karp reductions.

As for the data quality, recall that it is arguably easier to tell with sufficient certainty that some animal species inhabits a certain area, but harder to rule it out with the same certainty, especially for areas that are adjacent to habitated areas. This property is captured very well by d- REACH GBP and d- CLOSED GBP. Herein, one should choose d antiproportionally to the data quality. For instance, with perfect data quality, that is, perfect knowledge about each species' habitat, one may choose $d = 1$ (and hence, d- DIAM GBP is also amenable). Imperfect data quality is reflected by a choice of $d > 1$. Here, we relax the connectivity constraints and allow for "hops" within the connected habitat. If for example $d = 2$ and a possibly uninhabited area v is adjacent to two inhabited areas u and w, then u and w may be connected by $\{u, v\}$ and $\{v, w\}$, thus "hopping" over v.

Our Contributions Our results are summarized in Table 1. We settle the classic complexity and parameterized complexity (regarding the number k of green bridges and the number r of habitats) of the three problems. While d- REACH GBP is (surprisingly) already NP-hard for $d = 1$ on planar or maximum degree $\Delta = 4$ graphs, d- CLOSED GBP and d- DIAM GBP become NP-hard for $d \geq 2$, but admit an $(r + \Delta)^{O(1)}$-sized problem kernel and thus are linear time solvable if $r + \Delta$ is constant. All variants are para-NP-hard when parameterized by r. d- REACH GBP and d- CLOSED GBP are

Table 1 Overview of our results

Problem (Π GBP)		Comput. Complex.	Parameterized Algorithmics			Section
			k	r	$k + r$	
d-REACH	$d = 1$	NP-c.[a]	$2k$-vertex K[†]	p-NP-h.[‡]	$O(rk + k^2)$ PK	3.2
	$d = 2$	NP-c.[b]	$O(k^k)$-vertex K[†,*]	p-NP-h.[c]	FPT[†]	3.3
	$d \geq 3$	NP-c.	XP, W[1]-h.	p-NP-h.[c]	XP, W[1]-h.	3.4
d-CLOSED	$d = 1$	Lin. time	—	—	—	4
	$d = 2$	NP-c.[d]	$O(k^k)$-vertex K[†,*]	p-NP-h.[e,g]	FPT[†]	4.3
	$d \geq 3$	NP-c.	XP, W[1]-h.	p-NP-h.[e,g]	XP, W[1]-h.	4.4
d-DIAM	$d = 1$	Lin. time	—	—	—	5
	$d = 2$	NP-c.[f]	$2k$-vertex K†	p-NP-h.[g]	$O(rk + k^2)$ PK	5

NP-c., P, K, W[1]-h., and p-NP-h. stand for NP-complete, "polynomial-size", "problem kernel", W[1]-hard, and para-NP-hard, respectively

[a](even on planar graphs) [b](even on bipartite graphs with $\Delta = 4$ or graphs of diameter four) [c](even if $r = 1$ or if $r = 2$ and $\Delta = 4$) [d](even on bipartite graphs of diameter three and $r = 1$, *but* linear-time solvable when $r + \Delta$ is constant) [e](admits a linear-size problem kernel if Δ is constant) [f](linear-time solvable when $r + \Delta$ is constant) [g](even if $r = 1$) [†](no polynomial problem kernel unless NP $\subseteq$ coNP / poly) [*](but an $O(k^3)$-vertex problem kernel on planar graphs) [‡](if $r \geq 7$, linear-time solvable if $r \leq 2$)

fixed-parameter tractable regarding k when $d \leq 2$, but become W[1]-hard (yet XP) regarding k and $k + r$ when $d > 2$. Additionally, we prove that d- REACH GBP admits an rd-approximation in $O(mn + rnd)$ time.

Further related work Our problems deal with finding (small) spanning connected subgraphs obeying some (connectivity) constraints. These problems are applicable in a wide range of areas and typically take the form of a special case or variant of 1- REACH GBP. Areas include computer networks [6], social networks [2], graph drawing [4], combinatorial auctions [7], reconfigurable computing [16], vacuum technology [14], and structural biology [1].

1- REACH GBP on cliques is also known as the SUBSET INTERCONNECTION DESIGN problem: Given sets $V_1, \ldots, V_r$, find a graph G with $V(G) = V_1 \cup \cdots \cup V_r$ with the minimum number of edges such that $G[V_i]$ is connected for each i. This problem was first introduced by [13] and proven to be NP-hard by [15]. It was also studied in terms of its approximability [2] and its parameterized complexity [5].

Closely related to our problems are also Steiner multigraph problems [18, 34], which were also studied in the context of wildlife corridor construction [29, 30]. Requiring small diameter appears also in the context of spanning trees [32] and Steiner forests [11]. An edge-weighted version of 4- DIAM GBP is proven to be NP-hard even if there are only two different weights [31]. Kim et al. [27] study the problem of deleting few edges to augment a graph's diameter to a constant. Gionis et al. [19] studied a variant of 2- DIAM GBP in which for any solution F and habitat V_i, $G[F][V_i]$ must induce a star, and gave an efficient approximation algorithm for it. Herrendorf [24] studied the same variant as well as the 1- REACH GBP problem (under a different name) in terms of their parameterized complexity.

Connecting habitats arbitrarily The following obvious model *just* requires that each habitat is connected.

Problem 1 CONNECTED GREEN BRIDGES PLACEMENT (CONNECT GBP)

Input: An undirected graph $G = (V, E)$, a set $\mathscr{H} = \{V_1, \ldots, V_r\}$ of habitats where $V_i \subseteq V$ for all $i \in \{1, \ldots, r\}$, and an integer $k \in \mathbb{N}_0$.
Question: Is there a subset $F \subseteq E$ with $|F| \leq k$ such that for every $i \in \{1, \ldots, r\}$ it holds that in $G[F]$ there exists a connected component containing V_i?

CONNECT GBP with edge costs is also known as STEINER FOREST [18] and generalizes the well-known NP-hard STEINER TREE problem. Gassner [18] proved STEINER FOREST to be NP-hard even if every so-called terminal net contains two vertices, if the graph is planar and has treewidth three, and if there are two different edge costs, each being upper-bounded linearly in the instance size. It follows that CONNECT GBP is also NP-hard in this case. Bateni et al. [3] proved that STEINER FOREST is polynomial-time solvable on treewidth-two graphs and admits approximation schemes on planar and bounded-treewidth graphs.

From a modeling perspective, solutions for CONNECT GBP may be highly scattered: Patches of the same species' habitat may be arbitrarily far away from another; thus, to reach another patch of their habitat, animals may need to take long walks through areas of their habitats when only using green bridges to cross streets. It is likely that species with scattered habitats will not make use of the green bridges. With our models we avoid such solutions.

2 Preliminaries

Let $\mathbb{N}$ and $\mathbb{N}_0$ be the natural numbers without and with zero, respectively. We use basic definitions from graph theory [10] and parameterized algorithmics [8].

Graph Theory Let $G = (V, E)$ be an undirected graph with vertex set V and edge set $E \subseteq \binom{V}{2}$. We also denote by $V(G)$ and $E(G)$ the vertices and edges of G, respectively. For $V' \subseteq V$, let $G[V'] = (V', E \cap \binom{V'}{2})$ denote the graph G induced by a vertex set V'. For $F \subseteq E$ let $V(F) := \{v \in V \mid \exists e \in F : v \in e\}$ and $G[F] := (V(F), F)$ denote the graph G induced by the edge set F. A path P is a graph with $V(P) := \{v_1, \ldots, v_n\}$ and $E(P) := \{\{v_i, v_{i+1}\} \mid 1 \leq i < n\}$. The length of the path P is $|E(P)|$. The distance $\mathrm{dist}_G(v, w)$ between vertices $v, w \in V(G)$ is the length of the shortest path between v and w in G. The diameter $\mathrm{diam}(G)$ is the length of longest shortest path over all vertex pairs. For $p \in \mathbb{N}$, the graph G^p is the p-th power of G containing the vertex set V and edge set $\{\{v, w\} \in \binom{V}{2} \mid \mathrm{dist}_G(v, w) \leq p\}$. For $F \subseteq E, V' \subseteq V$, and $d \in \mathbb{N}$, the graph $G[F]^d[V']$ is understood as $((G[F])^d)[V']$. Let $N_G(v) := \{w \in V \mid \{v, w\} \in E\}$ be the (open) neighborhood of v, and $N_G[v] := N_G(v) \cup \{v\}$ be the closed neighborhood of v. For $p \in \mathbb{N}$, let $N_G^p(v) := \{w \in V \mid \{v, w\} \in E(G^p)\}$ be the (open) p-neighborhood of v, and $N_G^p[v] := N_G^p(v) \cup \{v\}$ be the closed p-neighborhood of v. Two vertices $v, w \in V$

are called twins if $N_G(v) = N_G(w)$. The (vertex) degree $\deg_G(v) := |N_G(v)|$ of v is the number of its neighbors. The maximum degree $\Delta(G) := \max_{v \in V} \deg_G(v)$ is the maximum over all (vertex) degrees.

3 Connecting Habitats with a Patch at Short Reach

The following problem ensures that any habitat patch can reach the other patches via patches of the same habitat and short strolls over "foreign" ground.

Problem 2 d- REACH GREEN BRIDGES PLACEMENT (d- REACH GBP)

Input: An undirected graph $G = (V, E)$, a set $\mathscr{H} = \{V_1, \ldots, V_r\}$ of habitats where $V_i \subseteq V$ for all $i \in \{1, \ldots, r\}$, and an integer $k \in \mathbb{N}_0$.
Question: Is there a subset $F \subseteq E$ with $|F| \leq k$ such that for every $i \in \{1, \ldots, r\}$ it holds that $V_i \subseteq V(G[F])$ and $G[F]^d[V_i]$ is connected?

Theorem 1 d- REACH GREEN BRIDGES PLACEMENT *is*

(i) *if* $d = 1$, NP-*hard even on planar graphs, or if* $r \geq 7$ *but solvable in linear time if* $r \leq 2$;
(ii) *if* $d = 2$, NP-*hard even on graphs with maximum degree four and* $r = 2$ *or graphs with diameter four and* $r = 1$, *and in* FPT *regarding* k;
(iii) *if* $d \geq 3$, NP-*hard and* W[1]-*hard regarding* $k + r$.

Moreover, d- REACH GBP admits an rd-approximation of the minimum number of green bridges in $O(mn + rnd)$ time.

We will first present the approximation algorithm. Afterwards, we will present the results in (i)-(iii) in the order above.

3.1 An ($r \cdot d$)-Approximation for d-REACH GBP

In this section we will present the approximation algorithm of Theorem 1. The approximation algorithm computes for every habitat V_i a spanning tree in $G^d[V_i]$, and adds the edges of the corresponding paths to the solution F. Each of the spanning trees then is a d-approximation for just the one habitat, hence the union of the spanning trees is an rd-approximation for all habitats.

Lemma 1 *For $r = 1$, d- REACH GBP admits a d-approximation of the minimum number of green bridges in $O(mn)$ time.*

Proof We start off by computing in $O(mn)$ time the graph $H := G^d$ as well as for every edge $e = \{u, v\} \in E(H)$ the corresponding path P_e from u to v of length at most d in G. If $H[V_1]$ is not connected, then return no. If not, then compute a minimum spanning tree $T \subseteq H[V_1]$ in $O(n \log n)$ time. For each edge $e = \{u, v\} \in E(T)$ compute in $O(m)$ time the corresponding path $P_e \subseteq G$ from u to v of length at most d. Finally, return the set $F := \bigcup_{e \in E(T)} E(P_e)$, computable in $O(m)$ time. Clearly, $G[F]^d[V_1]$ is

connected. As a minimum solution F^* has at least $|V_1| - 1$ edges, and every path P_e consists of at most d edges,

$$|F| = |\bigcup_{e \in E(T)} E(P_e)| \leq \sum_{e \in E(T)} E(P_e) \leq (|V_1| - 1) \cdot d \leq d|F^*|.$$

□

Proposition 1 d- REACH GBP *admits an rd-approximation of the minimum number of green bridges in $O(mn + rnd)$ time.*

Proof We initially compute the shortest paths between all vertex pairs in G in $O(mn)$ time. We obtain the graph $H := G^d$ as a byproduct. If for some $i \in \{1, \ldots, r\}$, $H[V_i]$ is not connected, then return no. If not, then compute for each $i \in \{1, \ldots, r\}$ a spanning tree T_i of $H[V_i]$, or return no if $H[V_i]$ is not connected. Let $F_i \subseteq E(G)$ be the edge set corresponding to T_i as in the proof of Lemma 1. As $G[F_i]^d[V_i]$ is connected, $F := \bigcup_{i=1}^r F_i$ is a solution.

Note that each of the r spanning trees T_i contains at most n edges, and for each of these edges $e \in F_i$ we can determine the corresponding paths $P_e \subseteq G$ of length at most d in $O(d)$ time. We obtain an overall running time of $O(mn + rnd)$.

As for the approximation ratio, let F^* be a minimum solution, and for every $i \in \{1, \ldots, r\}$ let $F_i^* \subseteq E(G)$ be a minimum-size edge set such that $G[F_i^*]^d[V_i]$ is connected. As $|F^*| \geq \max_{i \in \{1,\ldots,r\}} |F_i^*|$, we have

$$|F| \leq \sum_{i=1}^r |F_i| \leq \sum_{i=1}^r d|F_i^*| \leq r \cdot d|F^*|.$$

□

3.2 When a next habitat is directly reachable ($d = 1$)

Recall that setting $d = 1$ may reflect perfect knowledge about the habitats. In this case, we want that in $G[F]$, each habitat V_i forms a connected component.

Du and Miller [15] showed that 1- REACH GBP is NP-hard even when the input graph is complete. We give two reductions that show NP-hardness in some restricted cases. From the second reduction we can also derive that presumably there is no polynomial kernel with respect to the budget k. Lastly, we show that if there are only two habitats, then the problem can be solved in linear time.

We start with proving that 1- REACH GBP is NP-hard on series-parallel graphs. As every series-parallel graph is planar, we also obtain the same hardness result for planar graphs. Further, the provided reduction also shows that the problem is unlikely to admit a kernel whose size is bounded polynomially in the parameter.

Proposition 2 1- REACH GBP *is NP-hard and, unless* NP $\subseteq$ coNP / poly, *admits no problem kernel of size $k^{O(1)}$, even on series-parallel graphs.*

We will give a linear parametric transformation from the following problem:

Problem 3   HITTING SET (HS)

Input: A universe U, a set $\mathcal{F} \subseteq 2^U$ of subsets of U, and an integer k.
Question: Is there a *hitting set* $U' \subseteq U$ with $|U'| \leq k$ such that for all $F \in \mathcal{F}$ we have $F \cap U' \neq \emptyset$?

Note that HITTING SET admits no problem kernel of size polynomial in $|U|$ unless NP $\subseteq$ coNP / poly [12].

Construction 1 For an instance $\mathcal{I} = (U, \mathcal{F}, k)$ of HITTING SET with $U = \{1, \dots, n\}$ and $\mathcal{F} = \{F_1, \dots, F_m\}$, construct an instance $\mathcal{I}' := (G', \mathcal{H}, k')$ with habitats $\mathcal{H} = \{S, V_1, \dots, V_m\}$ and $k' := n + k$ as follows (see Fig. 2 for an illustration).
  Add to G' the vertex set $V_U := \{x_i \mid i \in U\}$ as well as the two vertices s and t, and the edge sets $E^* := \bigcup_{i=1}^n \{\{s, x_i\}\}$ and $E_U := \bigcup_{i=1}^n \{\{x_i, t\}\}$. Finally, let $S := \{s\} \cup \bigcup_{i=1}^n \{x_i\}$, and for each $F_j \in \mathcal{F}$ let $V_j := \{s, t\} \cup \bigcup_{i \in F_j} \{x_i\}$.

Observation 2 *The graph G' constructed in Construction 1 is planar and series-parallel.*

Observation 3 *Let $\mathcal{I}'$ be a* yes*-instance. Then every solution F contains all edges in E^*.*

Proof By construction, $G[S]$ is a star with center s. Hence, all edges in $G[S]$ are contained in every solution. Since $E^* = E(G[S])$, the claim follows. □

Lemma 4 *Let $\mathcal{I}'$ be the instance obtained from an instance $\mathcal{I}$ using Construction 1. Then, $\mathcal{I}$ is a* yes*-instance if and only if $\mathcal{I}'$ is a* yes*-instance.*

Proof ($\Rightarrow$)   Let $U' \subseteq U$ be a solution for instance $\mathcal{I}$. We claim that $F := E^* \cup \bigcup_{i \in U'} \{\{x_i, t\}\}$ is a solution for $\mathcal{I}'$. Note that $|F| \leq n + k$. Observe that $G'[F][S]$ is connected. Suppose now that there is V_j such that $G'[F][V_j]$ is not connected. Let F_j be the corresponding set. Since $E^* \subseteq F$, none of the edges $\{\{x_i, t\} \mid i \in F_j\}$ are contained in F. It follows that $F_j \cap U' = \emptyset$, contradicting the fact that U' is a solution for $\mathcal{I}$.
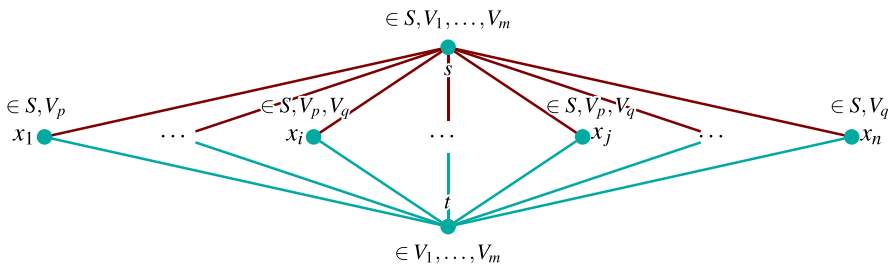

Fig. 2 Illustration to Construction 1 for 1- REACH GBP on series-parallel (and thus planar) graphs. In this example, there are e.g. $F_p \supseteq \{1, i, j\}$ and $F_q \supseteq \{i, j, n\}$. In case of a yes-instance, the red-colored edges are in every solution (Observation 3).

($\Leftarrow$) Let F be a solution to $\mathcal{I}'$. By Observation 11 we know that $E^* \subseteq F$. We claim that $U' := \{i \in U \mid \{x_i, t\} \in F\}$ is a solution for $\mathcal{I}$. Clearly $|U'| \leq k$. Suppose U' is not a solution. Then there is an $F_j \in \mathscr{F}$ with $F_j \cap U' = \emptyset$. But then $G'[F][V_j]$ is not connected, a contradiction. $\square$

Next, we prove that 1- REACH GBP is NP-hard even if we are given a constant number of habitats.

Proposition 3 1- REACH GBP *is* NP-*complete even if* $r = 7$.

We reduce from the following NP-hard problem.

Problem 4 DIRECTED HAMILTONIAN PATH (DHP)

Input: A directed graph $D = (W, A)$ and two distinct vertices $s, t \in W$ such that outdeg$(t) = $ indeg$(s) = 0$.
Question: Is there an s-t path that visits every vertex exactly once?

We first recall a well-known reduction to HAMILTONIAN PATH (HP), the undirected variant. Then, we reduce HP to 1- REACH GBP. For both constructions, we refer to Fig. 3 for an illustrative example.

Construction 2 Construct the undirected graph $G' = (V', E')$ as follows. For each vertex $v \in W \setminus \{s, t\}$, G' contains the path $P_v = (v_{in}, v, v_{out})$. Moreover, it contains the paths $P_s = (s, s_{out})$ and $P_t = (t_{in}, t)$. For every arc $(v, w) \in A$, add the edge (v_{out}, w_{in}).
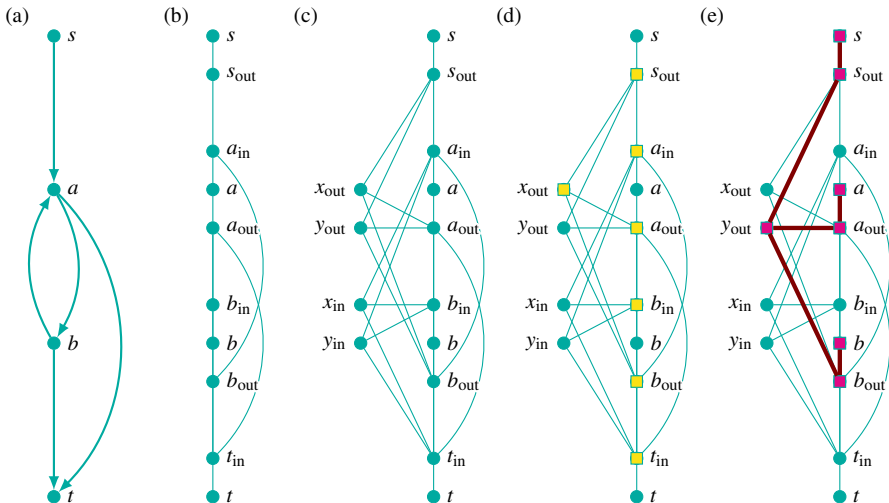

Fig. 3 Illustration to Constructions 2 and 3. Part (a) shows an exemplary directed graph which is a yes-instance for DHP. Applying Construction 2 on (a) yields (b). Applying Construction 3 on (b) yields the instance whose graph is depicted in (c) and two habitats of which are depicted in (d) and (e). Vertices marked yellow in (d) are contained in the habitat X_{out}. Vertices marked red in (e) are contained in the habitat Y_{out}. The graph induced by Y_{out} contains the red edges

Recall the following.

Observation 5 (D, s, t) *is a* yes-*instance of* DHP *if and only if* (G', s, t)*, obtained from* (D, s, t) *using Construction 2, is a* yes-*instance of* HP.

Next, we construct from G' the instance of 1- REACH GBP.

Construction 3 Let $G' = (W', E')$ obtained from $(D = (W, A), s, t)$ using Construction 2. We now construct the graph $G = (V, E)$ from G' with habitat set

$$\mathcal{H} = \{X_{\text{out}}, X_{\text{in}}, X_{\text{out}}^*, X_{\text{in}}^*, V_{\text{all}}, Y_{\text{out}}, Y_{\text{in}}\}$$

as follows.Add the new vertices $x_{\text{out}}, x_{\text{in}}, y_{\text{out}}, y_{\text{in}}$, that is, let

$$V := V' \cup \{x_{\text{out}}, x_{\text{in}}, y_{\text{out}}, y_{\text{in}}\}.$$

Moreover, make x_{out} adjacent to v_{out} for each $v \in W \setminus \{t\}$, make y_{out} adjacent to v_{out} for each $v \in W \setminus \{t\}$, make x_{in} adjacent to v_{in} for each $v \in W \setminus \{s\}$, make y_{in} adjacent to v_{in} for each $v \in W \setminus \{s\}$, Next, let $V_{\text{all}} := W'$ and

$$X_{\text{out}}^* := \{x_{\text{out}}\} \cup \bigcup_{v \in W \setminus \{t\}} \{v_{\text{out}}\}, \qquad X_{\text{out}} := X_{\text{out}}^* \cup \bigcup_{v \in W \setminus \{s\}} \{v_{\text{in}}\},$$

$$X_{\text{in}}^* := \{x_{\text{in}}\} \cup \bigcup_{v \in W \setminus \{s\}} \{v_{\text{in}}\}, \qquad X_{\text{in}} := X_{\text{in}}^* \cup \bigcup_{v \in W \setminus \{t\}} \{v_{\text{out}}\},$$

$$Y_{\text{out}} := \{y_{\text{out}}\} \cup \bigcup_{v \in W \setminus \{t\}} \{v_{\text{out}}, v\}, \text{ and} \qquad Y_{\text{in}} := \{y_{\text{in}}\} \cup \bigcup_{v \in W \setminus \{s\}} \{v_{\text{in}}, v\}.$$

Finally, let $k := 2(n-2) + 2 + 4(n-1) + (n-1) = 7(n-1)$, where $n = |W|$. ◇

As the habitats X_{out}^*, X_{in}^*, Y_{out}, and Y_{in} induce trees in G, we have the following.

Observation 6 *If* $(G, \mathcal{H}, k)$*, obtained from* (G', s, t) *using Construction 3, is a* yes-*instance of* 1- REACH GBP*, then every solution contains all edges contained in* P_v *for every* $v \in W$ *and all edges incident with* x_{out}*, with* x_{in}*, with* y_{out}*, and with* y_{in}.

Lemma 7 *The instance* $I' = (G', s, t)$*, obtained from* (D, s, t) *using Construction 2, is a* yes-*instance of* HP *if and only if* $I = (G, \mathcal{H}, k)$*, obtained from* (G', s, t) *using Construction 3, is a* yes-*instance of* 1- REACH GBP.

Proof Let $F' := \bigcup_{v \in W} E(P_v) \cup \{e \in E \mid e \cap \{x_{\text{out}}, x_{\text{in}}, y_{\text{out}}, y_{\text{in}}\} \neq \emptyset\}$ and let $n := |W|$. Note that $|F'| = 2(n-2) + 2 + 4(n-1)$.

($\Rightarrow$) Let $P = (s, v^2, \ldots, v^{n-2}, t)$ be an s-t path in D that visits every vertex exactly once. We claim that $F := F' \cup F''$ is a solution for instance I, where $F'' := \{u_{\text{out}}, v_{\text{in}} \mid (u, v) \in A(P)\}$. Clearly $|F| \leq |F'| + |F''| = k$. Note that the set F' already connects the habitats X_{out}^*, X_{in}^*, Y_{out}, and Y_{in}. Note that P is a subgraph of D that is weakly connected and in which every vertex has indegree one and every vertex has outdegree one, except for s (indegree zero) and t (outdegree zero). Hence,

for every $v \in W$ there exists an edge $(v, w) \in A(P)$, and thus $\{v_{\text{out}}, w_{\text{in}}\} \in F''$. Therefore, $G[F][X_{\text{in}}]$ is connected. The argumentation for $G[F][X_{\text{out}}]$ being connected is analogous. Finally, as P is a connected subgraph, F contains the edges of an s-t path that contains all vertices in V_{all}.

($\Leftarrow$) Let F be a solution to I. Due to Observation 6, we know that $F' \subseteq F$, and hence for $F'' := F \setminus F'$ we have $|F''| \leq k - (6(n-1)) = n - 1$. By definition of X_{out} and X_{in}, we know that in $G[F]$, every v_{out} is adjacent to at least one w_{in}, and every v_{in} is adjacent to at least one w_{out}. Thus, in the graph $P := (W, E^*)$ with $E^* = \{(v, w) \mid (v_{\text{out}}, w_{\text{in}}) \in F''\}$, every vertex has indegree and outdegree one, except for s (indegree zero) and t (outdegree zero). We claim that P is weakly connected. Consider any two vertices $v, w \in W$. By our assumption there exists a v-w path P' in $G[F][V_{\text{all}}]$. Note that $G[F'][V_{\text{all}}]$ has n connected components, each of which contains exactly one vertex in W. Hence, P' contains at least one edge in F'', and an additional edge in F'' for each additional vertex in W that is visited by P'. This edge set $E(P') \cap F''$ corresponds to the edges of an undirected v-w path in P. Hence, P is connected. Together with the above properties of P, it follows that P is a Hamiltonian s-t path. □

Lastly, we show that 1- REACH GBP becomes tractable for $r = 2$. Let $\alpha \colon \mathbb{N} \to \mathbb{N}$ be the inverse of the single-valued Ackermann function.

Proposition 4 1- REACH GBP *is solvable in* $O((n + m)\alpha(n))$ *time if* $r = 2$.

Proof Assume that both $G[V_1]$ and $G[V_2]$ are connected (otherwise, safely return no) and that $V_1 \cap V_2 \neq \emptyset$ (otherwise, a tree spanning over V_i for each $i \in \{1, 2\}$ is a valid, minimum-size solution). We first compute a spanning forest $T_\cap$ within $G[V_1 \cap V_2]$, using breadth-first search. Afterwards, for each $i \in \{1, 2\}$, we run Kruskal's [28] algorithm to extend the forest $T_\cap[V_i]$ to a spanning tree T_i that spans over the vertices in V_i. Let $F := E(T_1) \cup E(T_2)$. We return yes if and only if $|F| \leq k$. As each $v \in V_i$ is visited at most once, the algorithm runs in $O((n + m)\alpha(n))$ time by using the disjoint-set structure [38].

To prove the correctness of the algorithm, we show that F is a minimum-cardinality solution. Since both $G[V_1]$ and $G[V_2]$ are connected, $G[F][V_i]$ is connected for each $i \in \{1, 2\}$. It remains to show that F is of minimum cardinality. Consider some minimum-cardinality solution F'. Let $F'_i := E(G[F'][V_i])$ for each $i \in \{1, 2\}$, and let $F'_\cap := E(G[F'][V_1 \cap V_2])$. Observe that $|E(T_\cap)| \geq |F'_\cap|$ as otherwise there is cycle in $G[F'][V_1 \cap V_2]$ contradicting the fact that F' is of minimum-cardinality. It follows that

$$|F'| = |F'_1| + |F'_2| - |F_\cap| \geq |V_1| - 1 + |V_2| - 1 - |F_\cap|$$
$$\geq |V_1| - 1 + |V_2| - 1 - |E(T_\cap)| = |F|.$$

□

3.3 One hop between habitat patches ($d = 2$)

In this section we prove that 2- REACH GBP is already NP-complete even if there are two habitats and the graph has maximum degree four, or if there is only one habitat. Afterwards we show that 2- REACH GBP still admits a problem kernel with respect to k. If the graph is planar, we can show that the kernelization is polynomial in the number of vertices.

Proposition 5 d- REACH GBP *with* $d \geq 2$ *is* NP-*complete even if (i)* $r = 2$ *and* $\Delta \leq 4$ *or (ii)* $r = 1$ *and the input graph has diameter* $2d$.

For the sake of presentation, we prove Proposition 5(i) for $d = 2$. Afterwards, we briefly explain how to adapt the proof for $d > 2$ and for Proposition 5(ii).

Construction 4 Let $\mathscr{I} = (G, k)$ be an instance of 3- REGULAR VERTEX COVER with $G = (V, E)$ and $V = \{1, \ldots, n\}$ construct an instance of 2- REACH GBP with graph $G' = (V', E')$, habitat sets V_1 and V_2, and integer $k' := |E| + (n - 1) + k$ as follows (see Fig. 4(a) for an illustration).

Add the vertex set $V_E := \{v_e \mid e \in E\}$ and add v_e with $e = \{i, j\} \in E$ to habitat V_1. Next, add the vertex sets $V_G = \{v_i \mid i \in V\}$, and connect each v_i with all edge-vertices corresponding to an edge incident with i, i.e., add the edge set $E_G := \bigcup_{i \in V}\{\{v_i, v_e\} \mid i \in e\}$. Next, add the vertex set $V_X := \{x_i \mid i \in V\}$, connect each x_i with v_i, and add x_i to V_1 and to V_2. Finally, add the edge set $\{\{x_i, x_{i+1}\} \mid i \in \{1, \ldots, n - 1\}\}$. ◇

Observation 8 *Let* $\mathscr{I} = (G, k)$ *be an instance of* 3- REGULAR VERTEX COVER *and let* $\mathscr{I}' = (G', \{V_1, V_2\}, k')$ *be the instance obtained from* $\mathscr{I}$ *using Construction 4. If* $\mathscr{I}'$ *is a* yes-*instance, then every solution contains all edges in* $G[V_X]$.

Proof Suppose not, and let F be a solution without some edge $\{x_i, x_{i+1}\}$. Note that in $G - \{\{x_i, x_{i+1}\}\}$, the distance between x_i and x_{i+1} is at least four; thus $G[F]^2[V_X] = G[F]^2[V_2]$ is not be connected. A contradiction. □

Lemma 9 *Let* $\mathscr{I} = (G, k)$ *be an instance of* 3- REGULAR VERTEX COVER *and let* $\mathscr{I}' = (G', f, k')$ *be the instance obtained from* $\mathscr{I}$ *using Construction 4. If* $\mathscr{I}'$ *is a* yes-*instance, then there is a solution* $F \subseteq E(G')$ *such that* $\deg_{G'[F]}(v_e) = 1$ *for all* $e \in E(G)$.
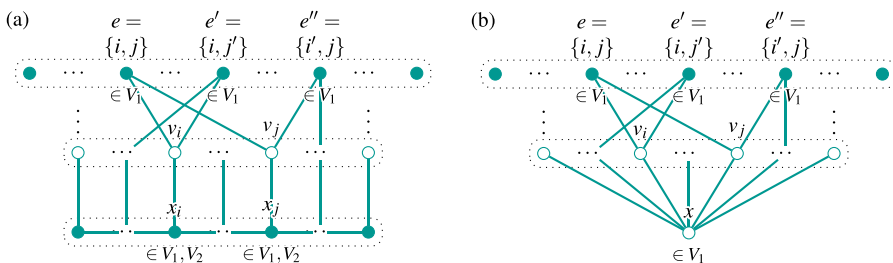

Fig. 4 Illustration for 2- REACH GBP with (a) $r = 2$ and $\Delta = 4$ ($k' = m + (n - 1) + k$) and (b) $r = 1$ ($k' = m + k$)

Proof Clearly, in every solution, we have $\deg_{G'[F]}(v_e) \geq 1$. Let F be a minimum solution with a minimum number of edges incident to vertices in $\{v_e \mid e \in E\}$. Suppose that there is at least one $e = \{i, j\} \in E$ such that $\deg_{G'[F]}(v_e) = 2$, that is, $\{v_e, v_i\}, \{v_e, v_j\} \in F$. Since F is a solution, there is a path P in $G'[F]$ from v_e to some x_i. Let $\{v_e, v_i\}$ be the first edge on this path. Let $F' := (F \setminus \{v_e, v_j\}) \cup \{v_j, x_j\}$. We claim that F' is a solution, yielding a contradiction to the fact that F is a solution with a minimum number of edges incident with vertices in V_E.

Only a vertex $v_{e'}$ can be disconnected from any V_X by removing $\{v_e, v_j\}$ from F. This vertex cannot be on the path P, and hence is connected to v_e via edge $\{v_e, v_j\}$. Since now edge $\{v_j, x_j\}$ is present, $v_{e'}$ is again connected to V_X. $\square$

Lemma 10 *Let $\mathscr{I} = (G, k)$ be an instance of* 3- REGULAR VERTEX COVER *and let $\mathscr{I}' = (G', \{V_1, V_2\}, k')$ be the instance obtained from $\mathscr{I}$ using Construction 4. Then $\mathscr{I}$ is a* yes*-instance if and only if $\mathscr{I}'$ is a* yes*-instance.*

Proof ($\Rightarrow$) Let $S \subseteq V$ be a vertex cover of size k in G. We construct a solution $F \subseteq E'$ as follows. Let $F_X = \bigcup_{i=1}^{n-1} \{\{x_i, x_{i+1}\}\}$ and $F_V := \{\{v_i, x_i\} \mid i \in S\}$. We define the auxiliary function $g \colon E \to V'$ with $g(\{i, j\}) = v_{\min(\{i,j\} \cap S)}$. Let $F_E := \bigcup_{e=\{i,j\} \in E} \{v_e, g(e)\}$. Let $F := F_X \cup F_V \cup F_E$. Note that $|F| = |F_X| + |F_V| + |F_E| \leq |E| + (n-1) + k = k'$. Moreover, every $v_e \in V_E$ is connected to x_i via a path (v_e, v_i, x_i), where $i \in (e \cap S)$. Finally, observe that $G'[F][V_X]$ is connected.

($\Leftarrow$) Let $\mathscr{I}'$ be a yes-instance. Due to Lemma 9 there is a solution $F \subseteq E'$ such that $\deg_{G'[F]}(v_e) = 1$ for all $e \in E$. Due to Observation 8, we know that the edges $\bigcup_{i=1}^{n-1} \{\{x_i, x_{i+1}\}\} \subseteq F$. Let $S := \{i \in V \mid \{v_i, x_i\} \in F\}$. We claim that S is a vertex cover. Suppose not, that is, there is an edge $e \in E$ such that $e \cap S = \emptyset$. That means that the unique neighbor of v_e, say v_i, is not adjacent with x_i in $G'[F]$. Since $\deg_{G'[F]}(v_e) = 1$ for all $e \in E$, $N_{G'[F]}[v_i]$ forms a connected component in $G'[F]^2$ not containing x_i. This contradicts the fact that F is a solution. $\square$

Remark 1

(i) To make the reduction work for $d \geq 3$, it is enough to subdivide each edge $\{v_e, v_i\}$ $(d-2)$ times and set $k' := (d-1)m + (n-1) + k$.

(ii) If we contract all x_i, set $V_2 = \emptyset$ (i.e., only one habitat remains), and set $k' := (d-1)m + k$, then the reduction is still valid (see Fig. 4(b) for an illustration). Thus, Proposition 5(ii) follows.

The reduction in the proof of Proposition 5 requires k to be linear in the input instance's size. We next prove that, indeed, 2- REACH GBP is fixed-parameter tractable with respect to k by showing that it admits a problem kernel of size exponential in k.

Proposition 6 2- REACH GBP *admits a problem kernel with at most $2k + \binom{2k}{k}$ vertices, at most $\binom{2k}{2} + k\binom{2k}{k}$ edges, and at most 2^{2k} habitats.*

Let $\bar{V} := V \setminus \bigcup_{V' \in \mathscr{H}} V'$ for a graph $G = (V, E)$ and habitat set $\mathscr{H} = \{V_1, \ldots, V_r\}$. The following reduction rules are immediate.

Reduction Rule 1

(i) *If $|V_i| = 1$ for some i, delete V_i.*
(ii) *If a vertex in $\bar{V}$ is of degree at most one, delete it.*
(iii) *If there is an $i \in \{1, \ldots, r\}$ with $|V_i| > 1$ and an $v \in V_i$ of degree zero, return a trivial* no-*instance.*
(iv) *If there is a vertex $v \in V \setminus \bar{V}$ of degree at most one, delete it (also from $V_1, \ldots, V_r$), and set $k := k - 1$.*

Clearly, k edges can connect at most $2k$ vertices; thus we obtain the following.

Reduction Rule 2 *If $|V \setminus \bar{V}| > 2k$, then return a trivial* no-*instance.*

So we have at most $2k$ vertices in habitats. Next, we upper-bound the number of non-habitat vertices. No minimal solution has edges between two such vertices.

Reduction Rule 3 *If there is an edge $e \in E$ with $e \subseteq \bar{V}$, then delete e.*

Moreover, no minimum solution connects through non-habitat twins.

Reduction Rule 4 *If $N(v) \subseteq N(w)$ for distinct $v, w \in \bar{V}$, then delete v.*

We still need to bound the number of vertices in $\bar{V}$. For an n-element set S let $\mathscr{F} \subseteq 2^S$ be a family of subsets such that for every $A, B \in \mathscr{F}$ we have $A \not\subseteq B$. Then $|\mathscr{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$ by Sperner's Theorem. Hence, after applying the reduction rules, we get an instance with at most $2k + \binom{2k}{k}$ vertices and $\binom{2k}{2} + 2k\binom{2k}{k}$ edges.

Finally, we can upper-bound the number of habitats by simply deleting duplicates.

Reduction Rule 5 *If $V_i = V_j$ for distinct $i, j \in \{1, \ldots, r\}$, then delete V_j.*

It follows that we can safely assume that $r \leq 2^{2k}$. Thus, Proposition 6 follows. Unfortunately, improving the problem kernel to polynomial-size appears unlikely.

Proposition 7 *Unless* NP $\subseteq$ coNP / poly, *d-* REACH GBP *for $d \geq 2$ admits no problem kernel of size $k^{O(1)}$, even if $r \geq 1$ is constant.*

We will give a linear parametric transformation from the following problem:

Problem 5 SET COVER (SC)

Input: A universe U, a set $\mathscr{F} \subseteq 2^U$ of subsets of U, and an integer k.
Question: Is there $\mathscr{F}' \subset \mathscr{F}$ with $|\mathscr{F}'| \leq k$ such that $\bigcup_{F \in \mathscr{F}'} F = U$?

The construction is basically the same as for Proposition 5(ii). Note that SET COVER admits no problem kernel of size polynomial in $|U| + k$, unless NP $\subseteq$ coNP / poly [12].

Proof Let $\mathscr{I} = (U, \mathscr{F}, k)$ be an instance of SET COVER, with $U = \{u_1, \ldots, u_n\}$. Construct an instance $\mathscr{I}' := (G, V_1, k')$ of 2- REACH GBP with $k' = |U| + k$ as follows (see Fig. 5). Let G be initially empty. Add the vertex set $V_U := U$, the vertex set $V_{\mathscr{F}} := \{v_F \mid F \in \mathscr{F}\}$, and the vertex x. Set $V_1 := V_U \cup \{x\}$. Make each vertex in $V_{\mathscr{F}}$ adjacent with x. Finally, for each $F \in \mathscr{F}$, add the edge set $\{\{v_i, v_F\} \mid u_i \in F\}$.
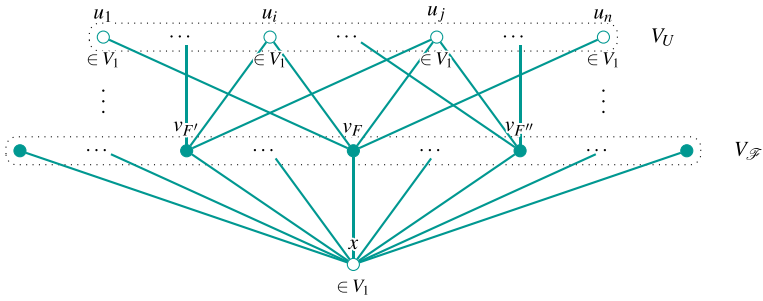
Fig. 5 Illustration for the construction in the proof of Proposition 7 for 2- REACH GBP with $r = 1$. In this example, $U = \{u_1, \ldots, u_n\}$ and we have $\{u_1, u_i, u_j, u_n\} = F \in \mathscr{F}$

The proof that $\mathscr{I}$ is a yes-instance if and only if $\mathscr{I}'$ is a yes-instance is analogous with the correctness proof for Proposition 5(ii).

Since SET COVER admits no problem kernel of size polynomial in $|U| + k$, unless NP $\subseteq$ coNP / poly [12], neither does 2- REACH GBP when parameterized by $k' = |U| + k$. □

Proposition 7 holds for general graphs. In fact, for planar graphs, the above reduction rules allow for an $O(k^3)$-vertex kernel. The number of habitats in the kernel however may still be exponential in k.

Proposition 8 2- REACH GBP *on planar graphs admits a problem kernel with* $O(k^3)$ *vertices and edges and at most* 2^{2k} *habitats.*

Observation 11 *Suppose all reduction rules were applied exhaustively. Then*

 (i) *there are at most* $\binom{2k}{2}$ *vertices of degree two in* $\bar{V}$, *and*
 (ii) *there are at most* $3\binom{2k}{3}$ *vertices of degree at least three in* $\bar{V}$.

Proof (i) By Reduction Rules 3, 2 and 4, every degree-two vertex in $\bar{V}$ has a pairwise different pair of neighbors in $V \setminus \bar{V}$. As there are $\binom{2k}{2}$ (unordered) vertex pairs in $V \setminus \bar{V}$, there are at most $\binom{2k}{2}$ degree-two vertices in $\bar{V}$, otherwise one of the reduction rules was not applied exhaustively.

(ii) Any three vertices u, v, w in a planar graph share at most two neighbors, that is, $|N(u) \cap N(v) \cap N(w)| \leq 2$. Suppose there are more than $3\binom{2k}{3}$ vertices in $\bar{V}$ of degree at least three. Then, by Reduction Rules 3 to 4, there are three vertices $u, v, w \in \bar{V}$ such that $|N(u) \cap N(v) \cap N(w)| \geq 3$, a contradiction to G being planar. □

As $|V \setminus \bar{V}| \leq 2k$ and we deleted all degree-one vertices, Proposition 8 follows.

3.4 At Least Two Hops Between Habitat Patches ($d \geq 3$)

If the data is more sparse, that is, the observed habitats to connect are rather scattered, then the problem becomes significantly harder to solve from the parameterized complexity point of view.
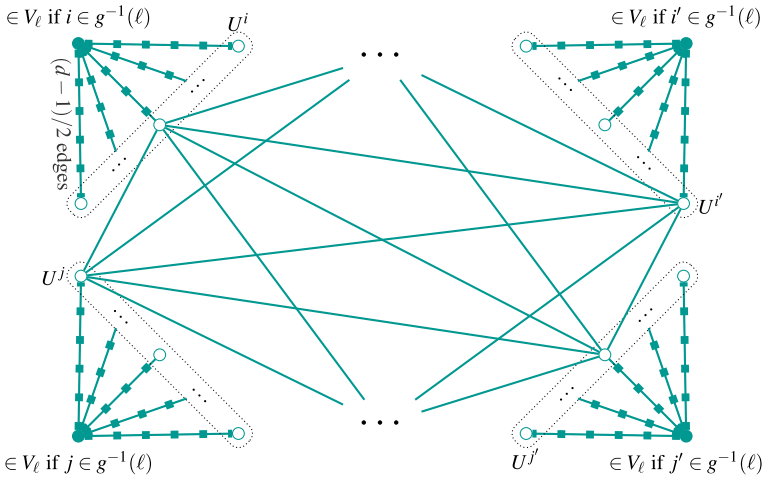
Fig. 6 Illustration to Construction 5 for *d*- REACH GBP for $d \geq 3$

Proposition 9 *d*- REACH GBP *with* $d \geq 3$ *is* NP-*complete and* W[1]-*hard when parameterized by* $k + r$.

We give the construction for *d* being odd. Afterwards, we explain how to adapt the reduction to *d* being even. The reduction is from the MULTICOLORED CLIQUE problem, where, given a *k*-partite graph $G = (U^1, \ldots, U^k, E)$, the question is whether there is a clique containing exactly one vertex from each part. MULTICOLORED CLIQUE is NP-hard and W[1]-hard when parameterized by *k*.

Construction 5 Let (G) with $G = (U^1, \ldots, U^k, E)$ be an instance of MULTICOLORED CLIQUE where $G[U^i]$ forms an independent set for every $i \in \{1, \ldots, k\}$. Assume without loss of generality that $U^i = \{u^i_1, \ldots, u^i_{|V^i|}\}$. Let $k' := \frac{(d-1)}{2}k + \binom{k}{2}$. Construct the instance $(G', \{V_1, \ldots, V_{\binom{k}{2}}\}, k')$ as follows (see Fig. 6 for an illustration).

Let $g \colon \binom{\{1,\ldots,k\}}{2} \to \{1, \ldots, \binom{k}{2}\}$ be a bijective function. Let G' be initially G. For each $i \in \{1, \ldots, k\}$, add a vertex v_i to G', add v_i to each habitat V_ℓ with $i \in g^{-1}(\ell)$, and connect v_i with u^i_j for each $j \in \{1, \ldots, u^i_{|U^i|}\}$ via a path with $\frac{d-1}{2}$ edges, where v_i and u^j_i are the endpoints of the path.

Remark 2 For every even $d \geq 4$, we can adapt the reduction for $d - 1$: At the end of the construction, subdivide each edge between two vertices that are in the original graph *G*.

Observation 12 In the obtained instance, for every $\ell \in \{1, \ldots, \binom{k}{2}\}$, it holds that, $V_\ell = \{v_i, v_j\}$ where $\{i, j\} = g^{-1}(\ell)$, and for every $i, j \in \{1, \ldots, k\}$, $i \neq j$, it holds that $\{\ell' \mid \{v_i, v_j\} \subseteq V_{\ell'}\} = \{\ell\}$ with $\ell = g(\{i, j\})$.

Observation 13 If the obtained instance is a yes-*instance, then in every minimal solution* F, *for every* $i \in \{1, \ldots, k\}$ *there is exactly one* u^i_j *in* $G[F]$.

Proof Note that each v_i must be connected with at least one vertex from U^i in $G'[F]$. Thus, $|V(G'[F]) \cap U^i| \geq 1$. Moreover, from each $i, j \in \{1, \ldots, k\}$, $i \neq j$, F must contain an edge between U^i and U^j, since $\text{dist}_{G'}(v_i, u) + \text{dist}_{G'}(v_j, u') \geq d - 1$ for every $u \in U^i$, $u' \in U^j$. Since additionally $k' = \frac{(d-1)}{2}k + \binom{k}{2}$, it follows that v_i cannot be connected with two vertices from U^i in $G'[F][U^i \cup \{v_i\}]$. Hence, if there are two vertices $u, u' \in U^i \cap F$, with u being connected to v_i in $G'[F][U^i \cup \{v_i\}]$, then u' is not part of an v_a-v_b path in $G'[F]$ of length at most d for every $a, b \in \{1, \ldots, k\}$. It follows that F is not minimal. □

Lemma 14 *Let $\mathscr{I} = (G)$ with $G = (U^1, \ldots, U^k, E)$ be an instance of* MUL-TICOLORED CLIQUE *and let $\mathscr{I}' = (G', \mathscr{H}, k')$ be the instance obtained from $\mathscr{I}$ using Construction 5. Then $\mathscr{I}$ is a* yes*-instance if and only if $\mathscr{I}'$ is a* yes*-instance.*

Proof ($\Rightarrow$)   Let $W \subseteq V(G)$ be a multicolored clique. Let F contain $\binom{W}{2}$ and all edges of a path from v_i to $U^i \cap W$. We claim that F is a solution. Note that $|F| = \binom{k}{2} + k\frac{d-1}{2}$. Since V_ℓ is of size two for all $\ell \in \{1, \ldots, \binom{k}{2}\}$ (Observation 12), we only need to show that v_i, v_j with $\{i, j\} = g^{-1}(\ell)$ is connected by a path of length at most d. We know that v_i is connected to some u_x^i by a path of length $(d-1)/2$, which is adjacent to some u_y^j, which is connected to v_j by a path of length $(d-1)/2$. Thus, v_i and v_j are of distance d.

($\Leftarrow$)   Let F be a solution. Note that $|F| = \binom{k}{2} + k\frac{d-1}{2}$. We claim that $W := V(G'[F]) \cap V(G)$ is a multicolored clique. First, observe that $|W| = k$ since for every v_i there is exactly one $u_{\ell_i}^i$ in $G'[F]$ (Observation 13). Suppose that W is not a multicolored clique, that is, there are U^i and U^j such that there is no edge in F between them. Then v_i and v_j are of distance larger than d in $G'[F]$, contradicting that F is a solution. □

4 Connecting Habitats at Short Pairwise Distance

In the next problem, we require short pairwise reachability.

Problem 6 *d*- CLOSED GREEN BRIDGES PLACEMENT (*d*- CLOSED GBP)

Input: An undirected graph $G = (V, E)$, a set $\mathscr{H} = \{V_1, \ldots, V_r\}$ of habitats where $V_i \subseteq V$ for all $i \in \{1, \ldots, r\}$, and $k \in \mathbb{N}_0$.
Question: Is there a subset $F \subseteq E$ with $|F| \leq k$ such that for every $i \in \{1, \ldots, r\}$ it holds that $V_i \subseteq V(G[F])$ and $G[F]^d[V_i]$ is a clique?

Note that if $G[F]^d[V_i]$ is a clique, then $\text{dist}_{G[F]}(v, w) \leq d$ for all $v, w \in V_i$. Further, *d*- CLOSED GBP is an unweighted variant of the 2NET problem [9].

Theorem 2 *d*- CLOSED GREEN BRIDGES PLACEMENT *is,*

(i) *if $d = 1$, linear-time solvable;*
(ii) *if $d = 2$, NP-hard even on bipartite graphs of diameter three and $r = 1$, and in* FPT *regarding k;*

(iii) if $d \geq 3$, NP-hard and W[1]-hard regarding k even if $r = 1$.

Further, d- CLOSED GBP is linear-time solvable if the number of habitats and the maximum degree are constant.

We first show the linear-time solvability for constant number of habitats and maximum degree. Afterwards we present the results in (i)-(iii).

4.1 Graphs of Constant Maximum Degree

2- REACH GBP is NP-hard if the number r of habitats and the maximum degree Δ are constant (Proposition 5). 2- CLOSED GBP is linear-time solvable in this case:

Proposition 10 *d- CLOSED GBP admits an $O(r\Delta(\Delta - 1)^{3d/2})$-sized problem kernel computable in $O(r(n + m))$ time.*

Proof Let $\mathscr{I} = (G, \mathscr{H}, k)$ be an instance of d- CLOSED GBP. For every $i \in \{1, \ldots, r\}$, fix a vertex $u_i \in V_i$. We assume that we have $V_i \subseteq N_G^d[u_i]$ for all $i \in \{1, \ldots, r\}$, otherwise $\mathscr{I}$ is a no-instance. Now let $W_i = N_G^{\lceil 3d/2 \rceil}[u_i]$ and let $G' := G[\bigcup_{i=1}^r W_i]$. Note that G' contains at most $r\Delta(\Delta - 1)^{\lceil 3d/2 \rceil}$ vertices and can be computed by r breadth-first searches. We claim that G' contains every path of length at most d between every two vertices $v, w \in V_i$, for every $i \in \{1, \ldots, r\}$. Recall that an edge set $F \subseteq E$ is a solution if and only if for every $i \in \{1, \ldots, r\}$ and for every $v, w \in V_i$, the graph $G[F]$ contains a path of length at most d from v to w. As by our claim G' contains any such path, this implies that $\mathscr{I}$ is a yes-instance if and only if $\mathscr{I}' := (G', \mathscr{H}, k)$ is a yes-instance (note that $V_i \subseteq V(G')$ for every $i \in \{1, \ldots, r\}$).

Assuming that $V_i \subseteq N_G^d[u_i]$, $G[W_i]$ contains all paths of length at most d between u_i and any $v \in V_i$. So let $v, w \in V_i$ be two vertices, both distinct from u_i. As $v, w \in N_G^d[u_i]$ and $W_i = N_G^{\lceil 3d/2 \rceil}[u_i]$, the subgraph $G[W_i]$ contains all vertices in $N_G^{\lceil d/2 \rceil}[v]$ and $N_G^{\lceil d/2 \rceil}[w]$. Consider now a path of length at most d between v and w. Suppose it contains a vertex $x \in V(G) \setminus (N_G^{\lceil d/2 \rceil}[v] \cup N_G^{\lceil d/2 \rceil}[w])$. Then $\mathrm{dist}_G(v, x) + \mathrm{dist}_G(w, x) > 2\lceil d/2 \rceil \geq d$, a contradiction to x being on a path from v to w of length at most d. The claim follows. $\square$

4.2 When Every Habitat Must be Complete ($d = 1$)

For $d = 1$, the problem is solvable in linear time: Check whether each habitat induces a clique. If so, check if the union of the cliques is small enough.

Observation 15 *1- CLOSED GBP is solvable in linear time.*

Proof We employ the following algorithm: For each $i \in \{1, \ldots, r\}$, let $G_i := G[V_i]$ and return no if G_i is not a clique. Finally, return yes if $|\bigcup_{i=1}^r E(G_i)| \leq k$, and no otherwise. Clearly, if the algorithm returns yes, then $\mathscr{I}$ is yes-instance. Conversely, let $\mathscr{I}$ be a yes-instance and let F' be a solution to $\mathscr{I}$. We know that for every $i \in$

$\{1, \ldots, r\}$, and any two vertices $v, w \in V_i$, edge $\{v, w\}$ must be in F'. It follows that $\bigcup_{i=1}^{r} E(G_i) \subseteq F'$. Thus, $|\bigcup_{i=1}^{r} E(G_i)| \leq |F'| \leq k$ and the algorithm correctly returns yes. $\qquad\square$

4.3 When Each Part is Just Two Steps Away ($d = 2$)

For $d = 2$, d-CLOSED GBP becomes NP-hard already on quite restrictive inputs. It is however, as we show at the end of this section, still fixed-parameter tractable when parameterized by k.

Proposition 11 2-CLOSED GBP *is NP-complete, even if $r = 1$ and the input graph is bipartite and of diameter three.*

Construction 6 Let $\mathscr{I} = (G, k)$ with $G = (V, E)$ be an instance of VERTEX COVER, and assume without loss of generality that $V = \{1, \ldots, n\}$. Construct an instance of 2-CLOSED GBP with graph $G' = (V', E')$, habitat V_1, and integer $k' := 2|E| + k + 3$ as follows (see Fig. 7 for an illustration).

To construct G' and V_1, add the vertex set $V_E := \{v_e \mid e \in E\}$ and add V_E to V_1. Add two designated vertices y' and y, add y to V_1, and make y' adjacent with y and all vertices in V_E. Add a designated vertex x, add x to V_1, and introduce a path of length two from x to y (call the inner vertex z). Add the vertex set $V_G := \{v_i \mid i \in V\}$, and make each v_i adjacent with x and all edge-vertices corresponding to an edge incident with i, i.e., add the edge set $E_G := \bigcup_{i \in V} \{\{v_i, v_e\} \mid i \in e\}$. $\diamond$

Observation 16 *Let $\mathscr{I}' = (G', \{V_1\}, k')$ be an instance obtained from applying Construction 6 on an instance $\mathscr{I} = (G, k)$ of VERTEX COVER. If $\mathscr{I}'$ is a yes-instance, then for every solution $F \subseteq E(G')$ it holds that $\{\{y, y'\}, \{y, z\}, \{z, x\}\} \cup \{\{y', v_e\} \mid e \in E(G)\} \subseteq F$.*

Lemma 17 *Let $\mathscr{I} = (G, k)$ be an instance of VERTEX COVER. Consider the instance $\mathscr{I}' = (G', \{V_1\}, k')$ obtained from $\mathscr{I}$ using Construction 6. If $\mathscr{I}'$ is a yes-instance, then there is a solution $F \subseteq E(G')$ such that $|N_{G'[F]}(v_e) \cap V_G| = 1$ for all $e \in E(G)$.*
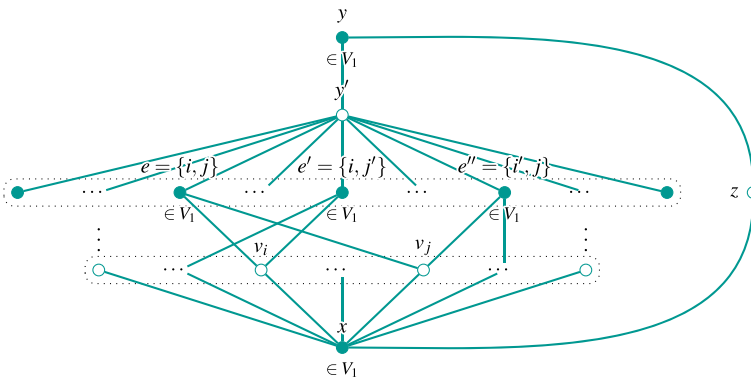

Fig. 7 Illustration to Construction 6 for 2-CLOSED GBP

Proof Note that in every solution, clearly we have $|N_{G'[F]}(v_e) \cap V_G| \geq 1$. Suppose there is a minimal solution F such that there is at least one $e = \{i, j\} \in E$ such that $|N_{G'[F]}(v_e) \cap V_G| = 2$. Let F be a solution with a minimum number of edges incident to vertices in V_E.

Since $\text{dist}_{G'[F]}(v_e, x) = 2$, at least one of the edges $\{v_i, x_i\}$ or $\{v_j, x_j\}$ are in F. If both are present then we can remove one of the edges $\{v_e, v_i\}$ or $\{v_e, v_j\}$ incident with v_e to obtain a solution of smaller size. This yields a contradiction.

Otherwise, assume there is exactly one edge, say $\{v_e, v_i\}$, contained in F. Then exchanging $\{v_e, v_j\}$ with $\{v_j, x\}$ yields a solution with a lower number of edges incident to vertices in V_E. A contradiction. □

Lemma 18 *Let $\mathscr{I} = (G, k)$ be an instance of* VERTEX COVER. *Consider the instance $\mathscr{I}' = (G', \{V_1\}, k')$ obtained from $\mathscr{I}$ using Construction 6. Then $\mathscr{I}$ is a* yes-*instance if and only if $\mathscr{I}'$ is a* yes-*instance.*

Proof $(\Rightarrow)$ Let $W \subseteq V$ be a vertex cover of size at most k in G. We construct a solution $F \subseteq E'$ as follows. Let F' denote the set of all edges required due to Construction 16. Let $F_V := \{\{v_i, x\} \mid i \in W\}$. We define the auxiliary function $g: E \to V'$ with $g(\{i, j\}) = v_{\min(\{i,j\} \cap W)}$. Let $F_E := \bigcup_{e=\{i,j\} \in E} \{v_e, g(e)\}$. Let $F := F' \cup F_V \cup F_E$. Note that $|F| = |F'| + |F_V| + |F_E| \leq |E| + 3 + |E| + k = k'$. Moreover, every $v_e \in V'$ is connected to x via a path (v_e, v_i, z), for some $i \in (e \cap W)$, of length two. Thus all vertex pairs in V_1 are at distance at most two.

$(\Leftarrow)$ Let $\mathscr{I}'$ be a yes-instance. Due to Construction 17, there is a solution $F \subseteq E'$ such that $\deg_{G'[F]}(v_e) = 1$ for all $e \in E$. Let $W := \{i \in V \mid \{v_i, x\} \in F\}$. We claim that W is a vertex cover. Suppose not, that is, there is an edge $e \in E$ such that $e \cap W = \emptyset$. That means that the unique neighbor of v_e, say v_i, is not adjacent with x in $G'[F]$. Then, v_e is not connected with x in $G'[F]^2$, and hence F is no solution, a contradiction. □

We next show fixed-parameter tractability when parameterizing by k. All the reduction rules that worked for 2- REACH GBP also work for 2- CLOSED GBP. It thus follows that 2- CLOSED GBP admits a problem kernel of size exponentially in k. As with 2- REACH GBP, the problem kernel presumably cannot be much improved. This can be shown by combining the constructions of Proposition 7 and 11.

Corollary 1 2- CLOSED GBP *admits a problem kernel of size exponentially in k and, unless* $\text{NP} \subseteq \text{coNP} / \text{poly}$, *none of size polynomial in k, even if $r = 1$.*

4.4 When Reaching Each Part is a Voyage ($d \geq 3$)

For $d \geq 3$, the problem is W[1]-hard regarding the number k of green bridges, even for one habitat. The reduction is similar to the one for Proposition 9.

Proposition 12 *d- CLOSED GBP with $d \geq 3$ is NP-complete and W[1]-hard when parameterized by the number k, even if $r = 1$.*

Proof Let $\mathscr{I} = (G)$ with $G = (U^1, \ldots, U^k, E)$ be an instance of MULTICOLORED CLIQUE. Apply Construction 5 to obtain instance $\mathscr{I}'' = (G', \{V_1, \ldots, V_{\binom{k}{2}}\}, k')$

(recall that $k' = \frac{d-1}{2}k + \binom{k}{2}$). Let $\mathscr{I}' = (G', \{V_1'\}, k')$ with $V_1' := \bigcup_{i=1}^{\binom{k}{2}} V_i = \{v_1, \ldots, v_k\}$ be the finally obtained instance of d- CLOSED GBP. We claim that $\mathscr{I}$ is a yes-instance if and only if $\mathscr{I}'$ is a yes-instance.

($\Rightarrow$) Let C be a multicolored clique in G. Let $z_i := V(C) \cap U^i$. We claim that F, consisting of the edges of each shortest path from v_i to z_i and the edge set $E(C)$, is a solution to $\mathscr{I}'$. Note that $|F| = k'$. Moreover, for any two $i, j \in \{1, \ldots, k\}$, we have that v_i and v_j are of distance $2\frac{d-1}{2} + 1 = d$. Hence, F is a solution.

($\Leftarrow$) Let F be a solution to $\mathscr{I}$. Since F must contain a path from v_i to some $z_i \in U^i$ for every $i \in \{1, \ldots, k\}$, there are at most $\binom{k}{2}$ edges left to connect. Let $Z := \{z_1, \ldots, z_k\}$ be the vertices such that v_i is connected with z_i in $G[F][U^i]$. As

$$d \geq \text{dist}_{G'[F]}(v_i, v_j) = \text{dist}_{G'[F]}(v_i, z_i) + \text{dist}_{G'[F]}(z_i, z_j) + \text{dist}_{G'[F]}(z_j, v_j)$$

and $d - 1 = \text{dist}_{G'[F]}(v_i, z_i) + \text{dist}_{G'[F]}(z_j, v_j)$, it follows that $\text{dist}_{G'[F]}(z_i, z_j) = 1$. Thus, $G[Z]$ forms a multicolored clique. □

5 Connecting Habitats at Small Diameter

Lastly, we consider requiring short pairwise reachability in 1- REACH GBP.

Problem 7 d- DIAMATER GREEN BRIDGES PLACEMENT (d- DIAM GBP)

Input: An undirected graph $G = (V, E)$, a set $\mathscr{H} = \{V_1, \ldots, V_r\}$ of habitats where $V_i \subseteq V$ for all $i \in \{1, \ldots, r\}$, and an integer $k \in \mathbb{N}_0$.
Question: Is there a subset $F \subseteq E$ with $|F| \leq k$ such that for every $i \in \{1, \ldots, r\}$ it holds that $V_i \subseteq V(G[F])$ and $G[F][V_i]$ has diameter d?

In particular, $G[F][V_i]$ is required to be connected. Note that 1- REACH GBP reduces to DIAM GBP (where d is part of the input and then set to the number of vertices in the input instance's graph). We have the following.

Theorem 3 d- DIAM GBP *is,*

  (i) *if $d = 1$, solvable in linear time;*
 (ii) *if $d = 2$, NP-hard even if $r = 1$.*

Moreover, d- DIAM GBP admits a problem kernel with at most $2k$ vertices and at most 2^{2k} habitats.

1- DIAM GBP is equivalent to 1- CLOSED GBP, which is linear-time solvable by Observation 15. Thus, Theorem 3(3) follows. Applying Reduction Rule 2 and 5 and deleting all non-habitat vertices yields the problem kernel. At the end of this section we show that 2- DIAM GBP most likely does not admit a polynomial kernel with respect to k. We now show that 2- DIAM GBP is NP-hard even if there is only one habitat.

Proposition 13 2- DIAM GBP *is NP-hard even if $r = 1$.*

Construction 7 Let $\mathscr{I} = (G, k)$ with $G = (V, E)$ be an instance of VERTEX COVER and assume without loss of generality that $V = \{1, \ldots, n\}$ and $E = \{e_1, \ldots, e_m\}$. Construct an instance $\mathscr{I}' := (G', \{V_1\}, k')$ with $k' := 3m + 2n + 12 + k$ as follows (see Fig. 8 for an illustration). Add the vertex sets $V_E := \{v_e \mid e \in E\}$ and $V_G = \{v_i \mid i \in V\}$, as well as the vertex set $V_A := \{x\} \cup \{y_i \mid i \in \{1, 2, 3\}\} \cup \{z_i \mid i \in \{1, \ldots, 4\}\}$. Add all vertices to V_1. Next, for each $e = \{i, j\} \in E$, connect v_e with v_i, v_j, y_1, and z_3. For each $i \in V$, connect v_i with x, y_1, and y_3. Lastly, add the edge set

$$E^* := \big\{\{y_1, y_2\}, \{y_2, y_3\}, \{y_1, z_1\}, \{y_1, z_2\}, \{y_3, z_3\}, \{y_3, z_4\}, \{y_3, x\},$$
$$\{z_1, z_4\}, \{z_1, z_2\}, \{z_2, z_3\}, \{z_2, x\}, \{z_3, z_4\}\big\}$$

to E'. Let $E_V^1 := \{\{y_1, v_i\} \mid i \in V\}$, $E_V^3 := \{\{y_3, v_i\} \mid i \in V\}$, $E_E^1 := \{\{y_1, v_e\} \mid e \in E\}$, and $E_E^3 := \{\{z_3, v_e\} \mid e \in E\}$. ◇

Observation 19 *Let $\mathscr{I}'$ be the instance obtained from some instance $\mathscr{I}$ using Construction 7. If $\mathscr{I}'$ is a yes-instance, then every solution F for $\mathscr{I}'$ contains the edge set $F' := E^* \cup E_V^1 \cup E_V^3 \cup E_E^1 \cup E_E^3$.*

Proof Let $\mathscr{I}'$ be a yes-instance and let F be a solution. Note that in $G' - \{y_1\}$, there is no path of length at most two from any vertex in $V_E \cup V_G$ to z_1. Hence, $E_V^1 \cup E_E^1 \subseteq F$. In $G' - \{y_3\}$, there is no path of length at most two from any vertex in $V_G \cup \{x\}$ to z_4. Hence, $E_V^3 \subseteq F$. In $G' - \{z_3\}$, there is no path of length at most two from any vertex in V_E to z_4. Hence, $E_E^3 \subseteq F$. In $G' - \{z_2\}$, there is no path of length at most two from x to z_1. Lastly, it is not difficult to see that every edge in E^* must be in F. □

We are set to prove the correctness of Construction 7.

Lemma 20 *Let $\mathscr{I}'$ be the instance obtained from some instance $\mathscr{I}$ using Construction 7. Then, $\mathscr{I}$ is a yes-instance if and only if $\mathscr{I}'$ is a yes-instance.*
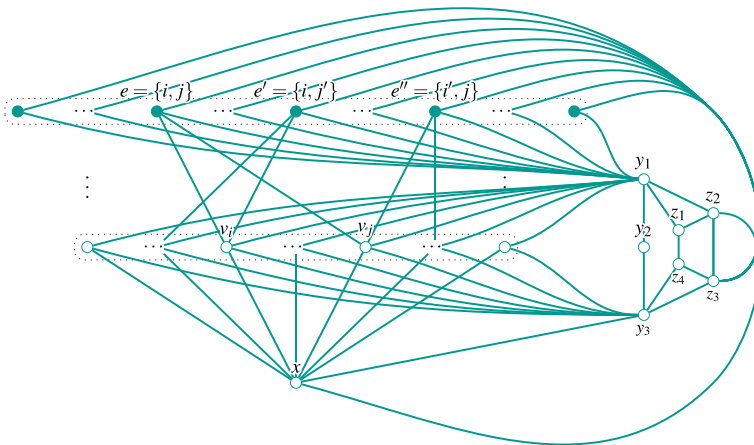

Fig. 8 Illustration for 2- DIAM GBP with $r = 1$

Proof ($\Rightarrow$)   Let $S \subseteq V$ be a vertex cover of size k. Let F' denote the set of all edges required to be in a solution due to Observation 19. Let $F_V := \{\{v_i, x\} \mid i \in S\}$. We define the auxiliary function $g \colon E \to V_G$ with $g(\{i, j\}) = v_{\min(\{i,j\} \cap S)}$. Let $F_E := \bigcup_{e \in E}\{\{v_e, g(e)\}\}$. Let $F := F' \cup F_V \cup F_E$. Note that $|F| = |F'| + |F_V| + |F_E| \leq (2m + 2n + 12) + k + m = k'$. Next consider $G'[F][V_1]$. Observe that $\mathrm{dist}_{G'[F][V_1]}(v, w) \leq 2$ for every vertices $v \in V_G \cup V_E \cup V_A$ and $w \in V_A \setminus \{x\}$, for every vertices $v, w \in V_G$, for every vertices $v, w \in V_E$, and for every vertices $v \in V_G$ and $w = \{x\}$. We claim that for all $e \in E$, $\mathrm{dist}_{G'[F][V_1]}(x, v_e) = 2$. By construction, $\mathrm{dist}_{G'[F][V_1]}(x, v_e) > 1$. Suppose that there is v_e with $e = \{i, j\}$ and $\mathrm{dist}_{G'[F][V_1]}(x, v_e) > 2$. Then there is no path (x, v, v_e) with $v \in \{v_i, v_j\}$. Then $\{i, j\} \cap S = \emptyset$, contradicting the fact that S is a vertex cover.

($\Leftarrow$)   Let F be a solution to $\mathscr{I}'$. Let F' be the set of edges mentioned in Observation 19; so $F' \subseteq F$. Note that $|F'| = 2m + 2n + 12$. Observe that in $G' - V_G$, the distance of x to any $v_e \in V_E$ is larger than two. Hence, for each v_e, there is a path (v_e, v, x) in $G'[F][V_1]$ with $v \in V_G$. We claim that $S := \{i \in V \mid \{v_i, x\} \in F\}$ is a vertex cover for G of size at most k. Suppose not, that is, there is an edge $e = \{i, j\}$ with $e \cap S = \emptyset$. This contradicts the fact that there is a path (v_e, v, x) in $G'[F][V_1]$ with $v \in V_G$. It remains to show that $|S| \leq k$. As F contains an edge $\{v_e, v\}$ with $v \in V_G$ for every $e \in E$, $|S| = |F \cap \{\{v_i, x\} \mid i \in V\}| \leq k' - (|F'| + m) = k$, and the claim follows. $\square$

Additionally, we have the following kernelization lower bound for 2- DIAM GBP.

Proposition 14 *Unless* NP $\subseteq$ coNP / poly, *2- DIAM GBP admits no problem kernel of size polynomial in k.*

Construction 8 Let $\mathscr{I} = (U, \mathscr{F}, k)$ with $U = \{u_1, \ldots, u_n\}$ and $\mathscr{F} = \{F_1, \ldots, F_m\}$ be an instance of HITTING SET. Construct an instance $\mathscr{I}' := (G', \mathscr{H}, k')$ with $k' = n + \binom{n}{2} + k$ as follows (see Fig. 9 for an illustration).

Let $V := V_{\mathscr{F}} \cup V_U \cup \{x\}$, where $V_F := \{v_F \mid F \in \mathscr{F}\}$ and $V_U := \{v_i \mid u_i \in U\}$. Add the edge sets $E' := \{\{v_i, v_{F_j}\} \mid u_i \in F_j\}$, $E_U := \{\{v_i, v_j\} \mid \{i, j\} \in \binom{n}{2}\}$, and $E_x := \{\{x, v_i\} \mid i \in \{1, \ldots, n\}\}$. The habitats $\mathscr{H} = \mathscr{H}_{\mathscr{F}} \cup \mathscr{H}_U$ are defined as follows. For each $F \in \mathscr{F}$, there is the habitat $V_F \in \mathscr{H}_{\mathscr{F}}$ with $V_F := \{x, v_F\} \cup \{v_i \mid u_i \in F\}$. For each $\{i, j\} \in \binom{n}{2}$, there is the habitat $V_{\{i,j\}} \in \mathscr{H}_U$ with $V_{\{i,j\}} := \{v_i, v_j\}$. Finally, let $k' := n + \binom{n}{2} + k$. $\diamond$
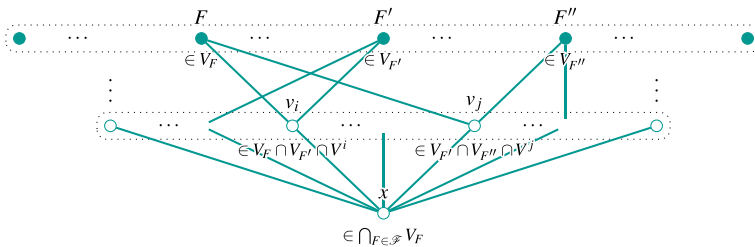

Fig. 9 Illustration for 2- DIAM GBP. Here, V^q denotes all sets $V_{\{q,\cdot\}}$

Due to $\mathcal{H}_U$, we know that every solution needs to contain every edge in the clique induced by V_U.

Observation 21 *Let $\mathcal{I}'$ be the instance obtained from some instance $\mathcal{I}$ using Construction 8. If $\mathcal{I}'$ is a* yes-*instance, then every solution F for $\mathcal{I}'$ contains the edge set E_U.*

Lemma 22 *Let $\mathcal{I}'$ be the instance obtained from some instance $\mathcal{I}$ using Construction 8. Then, $\mathcal{I}$ is a* yes-*instance if and only if $\mathcal{I}'$ is a* yes-*instance.*

Proof ($\Rightarrow$) Let $S \subseteq U$ be a hitting set of size k. We define the auxiliary function $g : \mathcal{F} \to V_U$ with $g(F) = v_{\min\{i \mid u_i \in S \cap F\}}$. Let $X_{\mathcal{F}} := \bigcup_{F \in \mathcal{F}}\{\{v_F, g(F)\}\}$. Then $X = E_U \cup X_{\mathcal{F}} \cup \{\{x, v_i\} \mid u_i \in S\}$ is a solution, as for every $F \in \mathcal{F}$, $G[X][V_F]$ contains as a subgraph a star with center $g(F)$ and leaves x and $V_U \setminus \{g(F)\}$, thus it is of diameter at most two.

($\Leftarrow$) Let X be a solution to I'. Due to Observation 21, we know that $E_U \subseteq X$. Moreover, every vertex in $V_{\mathcal{F}}$ has a neighbor in V_U. We claim that $S := \{u_i \mid \{x, v_i\} \in X\}$ is a solution to I. Suppose not. Then there exists a set $F \in \mathcal{F}$ with $S \cap F = \emptyset$. As $\operatorname{diam}(G[X][V_F]) \leq 2$, we have that the distance between v_F and x is at most two. But then X must contain both $\{v_F, v_i\}$ and $\{v_i, x\}$ for some $i \in \{1, \ldots, n\}$. But then, by construction of E', we have $u_i \in S \cap F$, a contradiction. □

6 Conclusion, Discussion, and Outlook

We modeled the problem of placing wildlife crossings with three different problem families: d- REACH GBP, d- CLOSED GBP, and d- DIAM GBP. We studied the practically desired cases $d = 1$ and $d = 2$, as well as the cases $d \geq 3$. For all three problems, we settled the classic as well as the parameterized complexity (regarding the number k of wildlife crossings and the number r of habitats). All three problems become NP-hard already for $d = 2$, and d- REACH GBP even for $d = 1$, in most of the cases on restricted input graphs and only few habitats. However, all three variants are fixed-parameter tractable regarding k in the case of $d = 2$, whereas, for $d \geq 3$, d- REACH GBP and d- CLOSED GBP turn out to be intractable (yet in XP) for this parameter. Thus, the less desired cases $d \geq 3$ are also algorithmically rather impractical. Moreover, d- CLOSED GBP and d- DIAM GBP are tractable if the number r of habitats and the maximum degree Δ of the graph are small, which is expected to be likely in real-world applications.

Discussion We derived an intriguing interrelation of connection requirements, data quality, and computational and parameterized complexity. While each problem admits its individual complexity fingerprint, each of them depends highly on the value of d, the level of the respective connectivity constraint. This value can reflect the quality of the given data, since naturally we assume that habitats are connected. The worse the data, the stronger are the relaxations according to the connectivity of habitats, and

thus the larger is the value of d. Our results show that having very small ($d = 2$) data gaps already leads to the problems becoming NP-hard, and that even larger gaps ($d \geq 3$) yield W[1]-hardness (when parameterized by k). Hence, knowledge about habitats, connections, and data quality decide which problem models can be applied, thus influencing the computation power required to determine an optimal placement of wildlife crossings. For instance, for larger networks, we recommend to ensure data quality such that one of our proposed problems for $d \leq 2$ becomes applicable. This in turn emphasizes the importance of careful habitat recognition.

In our models, we neglected that different positions possibly lead to *different* costs of building bridges (i.e., edge costs). This neglect is justified when differentiating between types of bridges (and thus their costs) is not necessary (e.g., if the habitat's species share preferred types of green bridges, and the underlying human-made transportation lines are homogeneous). In other scenarios, additionally considering these costs may be beneficial for decision-making.

Outlook and Open Problems As for algorithmic questions to the established problems, there are a few immediate questions that are unanswered in our work. While 1- REACH GBP is NP-hard even if $r \geq 7$ but polynomial-time solvable if $r \leq 2$, its complexity for $2 < r < 7$ remains open. Note that we obtained an $O(rd)$-approximation for d- REACH GBP, which possibly leaves room for improvement and does not directly transfer to the other two problem variants. It may be attractive to find out whether the problems admit FPT approximation algorithms as well. For $d \leq 2$, all our problems allow for problem kernels where the number of vertices only depends on k, but it is presumed impossible to have a polynomial dependence on k. If however the underlying street network is planar, then the input graphs to our problems can be seen as their planar dual. Therefore, it is likely that the input graphs are planar in real-world applications. In a follow-up work [23] we studied 1- REACH GBP with habitats that induce cycles and planar input graphs and analyzed the algorithms (among them the approximation algorithm from Proposition 1) on real-world graphs with synthetic habitats.

We conclude our work with some suggestions for extending our models. Interesting directions here include, for instance, distinguishing types of green bridges to place, taking into account possible movement directions within habitats (connectivity in directed graphs), identifying real-world driven problem parameters leading to tractability, or the problem of maintaining and servicing green bridges over time under a possible seasonal change of wildlife habitats (temporal graph modeling could fit well).

References

1. Agarwal, D., Araújo, J.C.S., Caillouet, C., Cazals, F., Coudert, D., Pérennes, S.: Connectivity inference in mass spectrometry based structure determination. In: Proc. 21st ESA, vol. 8125, pp. 289–300. (2013). https://doi.org/10.1007/978-3-642-40450-4_25

2. Angluin, D., Aspnes, J., Reyzin, L.: Network construction with subgraph connectivity constraints. J. Comb. Optim. **29**(2), 418–432 (2015). https://doi.org/10.1007/s10878-013-9603-2

3. Bateni, M., Hajiaghayi, M.T., Marx, D.: Approximation schemes for steiner forest on planar graphs and graphs of bounded treewidth. J. ACM **58**(5), 21:1–21:37. (2011). https://doi.org/10.1145/2027216.2027219

4. Brandes, U., Cornelsen, S., Pampel, B., Sallaberry, A.: Path-based supports for hypergraphs. J. Discrete Algorithms **14**, 248–261 (2012). https://doi.org/10.1016/j.jda.2011.12.009

5. Chen, J., Komusiewicz, C., Niedermeier, R., Sorge, M., Suchý, O., Weller, M.: Polynomial-time data reduction for the subset interconnection design problem. SIAM J. Discrete Math. **29**(1), 1–25 (2015). https://doi.org/10.1137/140955057

6. Chockler, G.V., Melamed, R., Tock, Y., Vitenberg, R.: Constructing scalable overlays for pub-sub with many topics. In: Proc. of 26th PODC, pp. 109–118. (2007). https://doi.org/10.1145/1281100.1281118

7. Conitzer, V., Derryberry, J., Sandholm, T.: Combinatorial auctions with structured item graphs. In: Proc. of 8th AAAI, pp. 212–218. (2004). http://www.aaai.org/Library/AAAI/2004/aaai04-034.php

8. Cygan, M., Fomin, F.V., Kowalik, Ł., Lokshtanov, D., Marx, D., Pilipczuk, M., Pilipczuk, M., Saurabh, S.: Parameterized Algorithms, Springer (2015)

9. Dahl, G., Johannessen, B.: The 2-path network problem. Networks **43**(3), 190–199 (2004). https://doi.org/10.1002/net.20003

10. Diestel, R.: Graph Theory, 4th Edition, Graduate texts in mathematics, vol. 173, Springer (2012)

11. Ding, W., Qiu, K.: A 2-approximation algorithm and beyond for the minimum diameter k-steiner forest problem. Theor. Comput. Sci. **840**, 1–15 (2020). https://doi.org/10.1016/j.tcs.2019.12.012

12. Dom, M., Lokshtanov, D., Saurabh, S.: Kernelization lower bounds through colors and ids. ACM Trans. Algorithms **11**(2), 13:1–13:20. (2014). https://doi.org/10.1145/2650261

13. Du, D.: An optimization problem on graphs. Discret. Appl. Math. **14**(1), 101–104 (1986). https://doi.org/10.1016/0166-218X(86)90010-7

14. Du, D., Kelley, D.F.: On complexity of subset interconnection designs. J. Glob. Optim. **6**(2), 193–205 (1995). https://doi.org/10.1007/BF01096768

15. Du, D., Miller, Z.: Matroids and subset interconnection design. SIAM J. Discrete Math. **1**(4), 416–424 (1988). https://doi.org/10.1137/0401042

16. Fan, H., Hundt, C., Wu, Y., Ernst, J.: Algorithms and implementation for interconnection graph problem. In: Proc. 2nd COCOA, vol. 5165, pp. 201–210. (2008). https://doi.org/10.1007/978-3-540-85097-7_19

17. Fox, M., Poole, D. (eds.): Proc. of 24th AAAI. AAAI Press (2010)

18. Gassner, E.: The steiner forest problem revisited. J. Discrete Algorithms **8**(2), 154–163 (2010). https://doi.org/10.1016/j.jda.2009.05.002

19. Gionis, A., Rozenshtein, P., Tatti, N., Terzi, E.: Community-aware network sparsification. In: Proc of SDM, pp. 426–434. (2017). https://doi.org/10.1137/1.9781611974973.48

20. Gomes, C.P.: Challenges for constraint reasoning and optimization in computational sustainability. In: I.P. Gent (ed.) Proc. of 15th CP, *Lecture Notes in Computer Science*, vol. 5732, pp. 2–4, Springer. (2009). https://doi.org/10.1007/978-3-642-04244-7_2

21. Gomes, C.P.: Computational sustainability: Computational methods for a sustainable environment, economy, and society. Bridge **39**(4), 5–13 (2009)

22. van der Grift, E., Seiler, A., Rosell, C., Simeonova, V.: Safe roads for wildlife and people: final report of the saferoad project. Tech. Rep., Conference of European Directors of Roads. (2017). https://www.saferoad-cedr.org/upload_mm/4/6/b/80eb4659-f6e6-4327-8358-f692b98f3419_CEDR_finalreport.pdf

23. Herkenrath, M., Fluschnik, T., Grothe, F., Kellerhals, L.: Placing green bridges optimally, with habitats inducing cycles. In: Proc. of 31st IJCAI, pp. 3825–3831. (2022). https://doi.org/10.24963/ijcai.2022/531

24. Herrendorf, E.: On the complexity of community-aware network sparsification. Master's thesis, Universitä Marburg. (2022). https://www.uni-marburg.de/de/fb12/arbeitsgruppen/algorith/forschung/master-emanuel-2.pdf

25. Huijser, M.P., Duffield, J.W., Clevenger, A.P., Ament, R.J., McGowen, P.T.: Cost-benefit analyses of mitigation measures aimed at reducing collisions with large ungulates in the united states and canada: a decision support tool. Ecol. Soc. **14**(2). (2009). http://www.jstor.org/stable/26268301

26. Huijser, M.P., McGowan, P., Hardy, A., Kociolek, A., Clevenger, A., Smith, D., Ament, R., et al.: Wildlife-vehicle collision reduction study: Report to congress. (2008). https://www.fhwa.dot.gov/publications/research/safety/08034/08034.pdf

27. Kim, E.J., Milanic, M., Monnot, J., Picouleau, C.: Complexity and algorithms for constant diameter augmentation problems. Theor. Comput. Sci. **904**, 15–26 (2022). https://doi.org/10.1016/j.tcs.2021.05.020

28. Kruskal, J.B.: On the shortest spanning subtree of a graph and the traveling salesman problem. Proc. Amer. Math. Soc. **7**, 48–50. (1956). https://doi.org/10.1090/S0002-9939-1956-0078686-7

29. Lai, K.J., Gomes, C.P., Schwartz, M.K., McKelvey, K.S., Calkin, D.E., Montgomery, C.A.: The steiner multigraph problem: Wildlife corridor design for multiple species. In: Proc. of 25th AAAI. AAAI Press. (2011). http://www.aaai.org/ocs/index.php/AAAI/AAAI11/paper/view/3768

30. LeBras, R., Dilkina, B., Xue, Y., Gomes, C.P., McKelvey, K.S., Schwartz, M.K., Montgomery, C.A.: Robust network design for multispecies conservation. In: Proc. of 27th AAAI, pp. 1305–1312. (2013). https://doi.org/10.1609/aaai.v27i1.8491

31. Plesník, J.: The complexity of designing a network with minimum diameter. Netw. **11**(1), 77–85 (1981). https://doi.org/10.1002/net.3230110110

32. Ravi, R., Sundaram, R., Marathe, M.V., Rosenkrantz, D.J., Ravi, S.S.: Spanning trees - short or small. SIAM J. Discret. Math. **9**(2), 178–200 (1996). https://doi.org/10.1137/S0895480194266331

33. Van der Ree, R., Heinze, D., McCarthy, M., Mansergh, I.: Wildlife tunnel enhances population viability. Ecol. Soc. **14**(2). (2009). http://www.ecologyandsociety.org/vol14/iss2/art7/

34. Richey, M.B., Parker, R.G.: On multiple steiner subgraph problems. Netw. **16**(4), 423–438 (1986). https://doi.org/10.1002/net.3230160408

35. Rossi, F. (ed.): IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013. IJCAI/AAAI. (2013). http://ijcai.org/proceedings/2013

36. Sawaya, M.A., Kalinowski, S.T., Clevenger, A.P.: Genetic connectivity for two bear species at wildlife crossing structures in banff national park. Proceedings of the Royal Society B: Biological Sciences **281**(1780), 20131705. (2014). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4027379/

37. Shilling, F., Waetjen, D., Porter, G., Short, C., Karcs, M., Honigman, T., Mejrano, M., Mohabir, G., Jyaw, M., Jones, A., Vickers, W., Harrold, K.: From wildlife-vehicle conflict to solutions for california drivers & animals. Tech. Rep., Road Ecology Center, UC Davis. (2021). https://roadecology.ucdavis.edu/resources/california-wildlife-vehicle-collision-hotspots-2021

38. Tarjan, R.E.: Efficiency of a good but not linear set union algorithm. J. ACM **22**(2), 215–225 (1975). https://doi.org/10.1145/321879.321884

39. Zheng, R., Luo, Z., Yan, B.: Exploiting time-series image-to-image translation to expand the range of wildlife habitat analysis. In: Proc. of 33rd AAAI, pp. 825–832. AAAI Press. (2019). https://doi.org/10.1609/aaai.v33i01.3301825

Springer